

## ORIGINAL ARTICLE

# Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity

Atul Malhotra, M.D., Ronald R. Grunstein, M.D., Ph.D., Ingo Fietze, M.D., Terri E. Weaver, Ph.D., Susan Redline, M.D., M.P.H., Ali Azarbarzin, Ph.D., Scott A. Sands, Ph.D., Richard J. Schwab, M.D., Julia P. Dunn, M.D., Sujatro Chakladar, Ph.D., Mathijs C. Bunck, M.D., Ph.D., and Josef Bednarik, M.D., for the SURMOUNT-OSA Investigators\*

## ABSTRACT

**BACKGROUND**

Obstructive sleep apnea is characterized by disordered breathing during sleep and is associated with major cardiovascular complications; excess adiposity is an etiologic risk factor. Tirzepatide may be a potential treatment.

**METHODS**

We conducted two phase 3, double-blind, randomized, controlled trials involving adults with moderate-to-severe obstructive sleep apnea and obesity. Participants who were not receiving treatment with positive airway pressure (PAP) at baseline were enrolled in trial 1, and those who were receiving PAP therapy at baseline were enrolled in trial 2. The participants were assigned in a 1:1 ratio to receive either the maximum tolerated dose of tirzepatide (10 mg or 15 mg) or placebo for 52 weeks. The primary end point was the change in the apnea–hypopnea index (AHI, the number of apneas and hypopneas during an hour of sleep) from baseline. Key multiplicity-controlled secondary end points included the percent change in AHI and body weight and changes in hypoxic burden, patient-reported sleep impairment and disturbance, high-sensitivity C-reactive protein (hsCRP) concentration, and systolic blood pressure.

**RESULTS**

At baseline, the mean AHI was 51.5 events per hour in trial 1 and 49.5 events per hour in trial 2, and the mean body-mass index (BMI, the weight in kilograms divided by the square of the height in meters) was 39.1 and 38.7, respectively. In trial 1, the mean change in AHI at week 52 was  $-25.3$  events per hour (95% confidence interval [CI],  $-29.3$  to  $-21.2$ ) with tirzepatide and  $-5.3$  events per hour (95% CI,  $-9.4$  to  $-1.1$ ) with placebo, for an estimated treatment difference of  $-20.0$  events per hour (95% CI,  $-25.8$  to  $-14.2$ ) ( $P < 0.001$ ). In trial 2, the mean change in AHI at week 52 was  $-29.3$  events per hour (95% CI,  $-33.2$  to  $-25.4$ ) with tirzepatide and  $-5.5$  events per hour (95% CI,  $-9.9$  to  $-1.2$ ) with placebo, for an estimated treatment difference of  $-23.8$  events per hour (95% CI,  $-29.6$  to  $-17.9$ ) ( $P < 0.001$ ). Significant improvements in the measurements for all prespecified key secondary end points were observed with tirzepatide as compared with placebo. The most frequently reported adverse events with tirzepatide were gastrointestinal in nature and mostly mild to moderate in severity.

**CONCLUSIONS**

Among persons with moderate-to-severe obstructive sleep apnea and obesity, tirzepatide reduced the AHI, body weight, hypoxic burden, hsCRP concentration, and systolic blood pressure and improved sleep-related patient-reported outcomes. (Funded by Eli Lilly; SURMOUNT-OSA ClinicalTrials.gov number, NCT05412004.)

From the University of California, San Diego, La Jolla (A.M.); Woolcock Institute of Medical Research, Macquarie University, Royal Prince Alfred Hospital, and the University of Sydney — all in Sydney (R.R.G.); the Center of Sleep Medicine, Charité University Hospital Berlin, Berlin (I.F.); the College of Nursing, University of Illinois Chicago, Chicago (T.E.W.); the School of Nursing (T.E.W.) and Perelman School of Medicine (R.J.S.), University of Pennsylvania, Philadelphia; the Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, and Harvard Medical School — both in Boston (S.R., A.A., S.A.S.); and Eli Lilly, Indianapolis (J.P.D., S.C., M.C.B., J.B.). Dr. Malhotra can be contacted at [amalhotra@health.ucsd.edu](mailto:amalhotra@health.ucsd.edu) or at the University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92037.

\*A complete list of the SURMOUNT-OSA trial investigators is provided in the Supplementary Appendix, available at [NEJM.org](http://NEJM.org).

This article was published on June 21, 2024, at [NEJM.org](http://NEJM.org).

DOI: 10.1056/NEJMoa2404881

Copyright © 2024 Massachusetts Medical Society.

**O**BSTRUCTIVE SLEEP APNEA IS CHARACTERIZED by repetitive pharyngeal collapse during sleep resulting in apneas and hypopneas, with consequent hypoxemia, hypercapnia, and recurrent arousals.<sup>1</sup> Obstructive sleep apnea is accompanied by clinically relevant symptoms, such as excessive daytime sleepiness, and is an independent risk factor for cardiovascular disease.<sup>1,2</sup> The disease is common and has major medical and economic effects; more than 900 million persons are affected worldwide, approximately 40% of whom have moderate-to-severe disease.<sup>3</sup>

The treatment of patients with obstructive sleep apnea has historically focused on mechanical support during sleep. Positive airway pressure (PAP) therapy improves the apnea–hypopnea index (AHI, the number of apneas and hypopneas during an hour of sleep) and reduces symptoms related to obstructive sleep apnea, but its overall effectiveness can be affected by varying adherence to therapy. Randomized, controlled trials have failed to show that PAP reduces occurrences of adverse cardiovascular outcomes and death.<sup>4,6</sup> Mandibular advancement therapy is predominantly used in patients who are unable or unwilling to adhere to treatment with PAP, but it is not universally efficacious.<sup>7</sup> Upper-airway surgery, including stimulation of the hypoglossal nerve, may be effective but is an invasive option that may be appropriate for selected patients. At present, there is no pharmaceutical intervention that has been approved for the treatment of obstructive sleep apnea.

Excess adiposity is a major reversible etiologic risk factor for obstructive sleep apnea and its complications.<sup>8,9</sup> The benefit of substantial weight reduction in the treatment of patients with obstructive sleep apnea is well recognized, and clinical guidelines recommend treatment of obesity in patients with obstructive sleep apnea.<sup>9</sup> Thus, a pharmacologic intervention that targets obesity and its downstream effects on obstructive sleep apnea, symptoms, blood pressure, and low-grade systemic inflammation may facilitate a holistic approach that is not fully attained with the aforementioned mechanical treatments.<sup>10-12</sup>

Tirzepatide is a long-acting glucose-dependent insulinotropic polypeptide (GIP) receptor and glucagon-like peptide-1 (GLP-1) receptor agonist that selectively binds to and activates both the GIP and GLP-1 receptors. It is an amino acid sequence that includes a C20 fatty diacid moiety that enables albumin binding, which prolongs the half-

life.<sup>13</sup> Treatment with tirzepatide has led to significant reductions in excess body weight, improvements in blood pressure, and reductions in markers of inflammation and vascular endothelial dysfunction, and may have the potential to be efficacious in persons with obstructive sleep apnea.<sup>14,15</sup> Here we report the results of the SURMOUNT-OSA phase 3 trials evaluating the safety and efficacy of tirzepatide for the treatment of adults with obstructive sleep apnea and obesity.

## METHODS

### TRIAL DESIGN

The SURMOUNT-OSA trials were two 52-week, phase 3, multicenter, parallel-group, double-blind, randomized, controlled trials that were conducted at 60 sites across nine countries to evaluate the efficacy and safety of the maximum tolerated dose of weekly tirzepatide (10 mg or 15 mg) in adults with moderate-to-severe obstructive sleep apnea and obesity (Fig. S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org). Two participant populations were included in the SURMOUNT-OSA master protocol (available at NEJM.org): trial 1 included participants who were unable or unwilling to use PAP therapy, and trial 2 included participants who had been using PAP therapy for at least 3 consecutive months at the time of screening and who planned to continue PAP therapy during the trial. The master protocol rationale and design have been previously published.<sup>16</sup> The protocols were designed by the sponsor (Eli Lilly) and members of the trial steering committee. The protocol was approved by the relevant institutional review boards, and all the participants provided written informed consent.

The trials were conducted in accordance with the Good Clinical Practice guidelines of the International Council for Harmonisation and the principles of the Declaration of Helsinki. Statistical analyses were performed by employees of the sponsor. The first and last authors prepared the first draft of the manuscript, which was reviewed, edited, and approved by all the authors. A medical writer employed by the sponsor provided medical-writing assistance. The investigators and steering committee worked under confidentiality agreements with the sponsor. The sponsor was involved in the collection, management, analyses, and in-

terpretation of the data; the preparation, review, and approval of an earlier version of the manuscript; and the decision to submit the manuscript for publication. Final decisions on preparation of the manuscript for submission were made by the authors, some of whom were employees of the sponsor. The authors vouch for the completeness and accuracy of the data and for the fidelity of the trials to the protocols.

#### PARTICIPANTS

Adults who had received a diagnosis of moderate-to-severe obstructive sleep apnea (AHI  $\geq 15$  events per hour) and obesity (body-mass index [BMI, the weight in kilograms divided by the square of the height in meters] of  $\geq 30$  [ $\geq 27$  in Japan]) were eligible. Key exclusion criteria were the presence of type 1 or type 2 diabetes, a participant-reported change in body weight of more than 5 kg in the 3 months before screening, planned surgery for sleep apnea or obesity, a diagnosis of central or mixed sleep apnea, and major craniofacial abnormalities. A full list of eligibility criteria is available in the Supplementary Appendix.

#### TRIAL PROCEDURES

After a 4-week screening period, participants were assigned to trial 1 or trial 2 and randomly assigned in a 1:1 ratio to receive tirzepatide or placebo subcutaneously once weekly with the use of a single-dose pen autoinjector. All the participants received regular lifestyle counseling sessions regarding the maintenance of healthy nutrition while adhering to a 500 kilocalorie per day deficit and at least 150 minutes per week of physical activity.

Randomization was conducted by means of a Web-based interactive response system and stratified according to trial, country or geographic region, baseline AHI severity category, and sex. Enrollment of men was limited to 70% to ensure adequate representation of women. Participants, investigators, and the sponsor were unaware of trial-group assignment. Participants were required to receive tirzepatide or placebo during a planned 52-week period that included a dose-escalation period of up to 20 weeks and a 4-week safety follow-up. The initial dose of tirzepatide was 2.5 mg once weekly and was increased by 2.5 mg every 4 weeks during the dose-escalation period until the participant reached the maximum tolerated dose of 10 mg or 15 mg in week 20. Par-

ticipants in whom doses of 10 mg or more produced unacceptable side effects discontinued tirzepatide or placebo but were encouraged to remain in the trial.

The AHI was measured by laboratory polysomnography at screening, week 20, and week 52. Data from polysomnographic studies were scored centrally with the use of the American Academy of Sleep Medicine rule 1B for identification of hypopneas (which specifies a  $\geq 30\%$  reduction in airflow for  $\geq 10$  seconds and oxygen desaturation of  $\geq 4\%$ ).<sup>17</sup>

#### END POINTS AND ASSESSMENTS

The primary end point was the change in the AHI from baseline. Key secondary end points that were controlled for type 1 errors included the percent change in AHI; the percentage of participants with an AHI reduction of at least 50%; the percentage of participants with an AHI of less than 5 events per hour or with an AHI of 5 to 14 events per hour and a score of 10 or less on the Epworth Sleepiness Scale (ESS; range, 0 to 24, with higher scores indicating greater daytime sleepiness); the percent change in body weight; the change in high-sensitivity C-reactive protein (hsCRP) concentration; the change in sleep apnea-specific hypoxic burden (a measure calculated from a polysomnographic study that comprises frequency, duration, and depth of oxygen saturation related to the respiratory event)<sup>18</sup>; the change in scores on the Patient-Reported Outcomes Measurement Information System (PROMIS) Short Form Sleep-related Impairment 8a (PROMIS-SRI) and PROMIS Short Form Sleep Disturbance 8b (PROMIS-SD) scales (higher scores indicate more sleep impairment or sleep disturbance, respectively); and the change in systolic blood pressure. Participants in trial 2 were instructed to suspend PAP therapy for 7 days before polysomnographic and patient-reported outcome (PRO) assessments at baseline, week 20, and week 52 to minimize the confounding effect of PAP therapy on sleep-disordered breathing and other breathing-related and PRO assessments. All end points were assessed from baseline to week 52 except for blood pressure, which was assessed at week 48 to prevent suspension of PAP therapy in trial 2 from confounding the assessment.

In response to a recommendation by a regulatory body, key secondary PRO end points were changed from the hierarchical combination of

change in scores on the Functional Outcomes of Sleep Questionnaire to change in PROMIS-SRI and PROMIS-SD. Owing to the timing of regulatory advice and because the change did not affect how clinical trial investigators conducted the trials, the change in key secondary PRO end points was captured in the final statistical analysis plan, available with the protocol, before the time of the data unblinding, the database lock, and data analyses (details are provided in the Supplement-

**Table 1. Demographic and Clinical Characteristics of the Participants at Baseline.\***

Characteristic	Trial 1			Trial 2		
	Tirzepatide (N=114)	Placebo (N=120)	Total (N=234)	Tirzepatide (N=120)	Placebo (N=115)	Total (N=235)
Age — yr	47.3±11.0	48.4±11.9	47.9±11.5	50.8±10.7	52.7±11.3	51.7±11.0
<50 yr	63 (55.3)	62 (51.7)	125 (53.4)	54 (45.0)	45 (39.1)	99 (42.1)
≥50 yr	51 (44.7)	58 (48.3)	109 (46.6)	66 (55.0)	70 (60.9)	136 (57.9)
Female sex — no. (%)	36 (31.6)	41 (34.2)	77 (32.9)	33 (27.5)	32 (27.8)	65 (27.7)
Race or ethnic group — no. (%)						
American Indian or Alaska Native	9 (7.9)	9 (7.5)	18 (7.7)	10 (8.3)	9 (7.9)	19 (8.1)
Asian	23 (20.2)	24 (20.0)	47 (20.1)	17 (14.2)	16 (14.0)	33 (14.1)
Black or African American	6 (5.3)	7 (5.8)	13 (5.6)	8 (6.7)	3 (2.6)	11 (4.7)
White	74 (64.9)	80 (66.7)	154 (65.8)	85 (70.8)	86 (75.4)	171 (73.1)
Multiple	2 (1.8)	0	2 (0.9)	—	—	—
Hispanic or Latino	51 (44.7)	47 (39.2)	98 (41.9)	38 (31.7)	38 (33.0)	76 (32.3)
Body weight — kg	116.7±24.6	112.8±22.6	114.7±23.7	115.8±21.5	115.1±22.7	115.5±22.0
Body-mass index						
Mean value	39.7±7.3	38.6±6.7	39.1±7.0	38.6±6.1	38.7±6.0	38.7±6.0
Distribution — no. (%) †						
<35	33 (28.9)	44 (36.7)	77 (32.9)	33 (27.7)	33 (28.9)	66 (28.3)
≥35 to <40	39 (34.2)	35 (29.2)	74 (31.6)	47 (39.5)	41 (36.0)	88 (37.8)
≥40	42 (36.8)	41 (34.2)	83 (35.5)	39 (32.8)	40 (35.1)	79 (33.9)
Waist circumference — cm	122.6±16.6	119.8±14.8	121.2±15.7	120.7±13.1	121.0±14.0	120.9±13.5
AHI — events/hr	52.9±30.5	50.1±31.5	51.5±31.0	46.1±22.4	53.1±30.2	49.5±26.7
Obstructive sleep apnea severity — no. (%) ‡						
No apnea	0	1 (0.8)	1 (0.4)	—	—	—
Mild: AHI <15 events/hr	1 (0.9)	2 (1.7)	3 (1.3)	0	2 (1.8)	2 (0.9)
Moderate: AHI ≥15 events/hr	39 (34.2)	43 (36.1)	82 (35.2)	35 (29.4)	37 (32.5)	72 (30.9)
Severe: AHI ≥30 events/hr	74 (64.9)	73 (61.3)	147 (63.1)	84 (70.6)	75 (65.8)	159 (68.2)
Missing data	0	1 (0.8)	1 (0.4)	1 (0.8)	1 (0.9)	2 (0.9)
PROMIS Sleep-related Impairment T score §	53.2±7.5	54.3±8.5	53.8±8.1	55.3±8.4	55.0±9.5	55.2±8.9
PROMIS Sleep Disturbance T score ¶	53.8±6.0	53.5±7.4	53.6±6.7	56.0±7.6	55.7±7.6	55.9±7.6
ESS score	10.3±5.3	10.8±5.2	10.6±5.3	10.8±4.6	9.5±4.4	10.2±4.5
Sleep apnea-specific hypoxic burden — % min/hr**	153.6 (102.7)	137.8 (104.1)	145.3 (103.4)	132.2 (83.4)	142.1 (112.5)	137.0 (97.5)
Systolic blood pressure — mm Hg	128.4±12.2	130.3±10.7	129.4±11.5	130.5±14.3	130.5±12.8	130.5±13.5
Diastolic blood pressure — mm Hg	83.7±8.9	84.0±8.6	83.8±8.7	83.2±8.2	80.5±8.6	81.8±8.5
Hypertension — no. (%)	84 (73.7)	93 (77.5)	177 (75.6)	91 (75.8)	91 (79.1)	182 (77.4)

**Table 1. (Continued.)**

Characteristic	Trial 1			Trial 2		
	Tirzepatide (N=114)	Placebo (N=120)	Total (N=234)	Tirzepatide (N=120)	Placebo (N=115)	Total (N=235)
hsCRP concentration — mg/liter <sup>††</sup>	3.5 (120.0)	3.6 (124.6)	3.5 (122.0)	3.0 (124.3)	2.7 (127.5)	2.8 (125.8)
Prediabetes — no. (%)	74 (64.9)	78 (65.0)	152 (65.0)	69 (57.5)	64 (55.7)	133 (56.6)
Glycated hemoglobin — %	5.69±0.37	5.64±0.35	5.67±0.36	5.62±0.37	5.65±0.44	5.63±0.41
Dyslipidemia — no. (%)	91 (79.8)	98 (81.7)	189 (80.8)	100 (83.3)	97 (84.3)	197 (83.8)

\* Plus-minus values are mean ±SD. Categories include all participants who underwent randomization unless otherwise noted.

<sup>†</sup> Trial 2 had one missing participant value for body-mass index for each of the two trial groups.

<sup>‡</sup> Participants with an apnea-hypopnea index (AHI, the number of apneas and hypopneas during an hour of sleep) of less than 15 events per hour were determined to have been enrolled in error and were withdrawn from the trial.

<sup>§</sup> The PROMIS Short Form Sleep-related Impairment 8a consists of eight factors that the participant can recall in the past 7 days, with each factor rated on a 5-point scale from “not at all” to “very much.” Scores for individual factors were totaled to obtain a raw score that was then converted to a T score (with the use of response-pattern scoring), with a mean score of 50 and a standard deviation of 10, with higher scores indicating more sleep-related impairment.<sup>23</sup>

<sup>¶</sup> The PROMIS Short Form Sleep Disturbance 8b consists of eight factors that the participant can recall in the past 7 days, with each factor rated on a 5-point scale from “not at all” to “very much,” “never” to “always,” or “very poor” to “very good.” Scores for individual factors were totaled to obtain a raw score that was then converted to a T score (with the use of response-pattern scoring), with a mean score of 50 and a standard deviation of 10, with higher scores indicating more sleep disturbance.<sup>23</sup>

<sup>||</sup> The Epworth Sleepiness Scale (ESS) is an eight-factor participant-reporting measure that asks the participant to rate, on a scale of 0 (would never doze) to 3 (high chance of dozing), their recent typical likelihood of dozing in eight different daytime situations. The ESS total score is the sum of the eight factor scores and ranges from 0 to 24, with higher scores indicating greater daytime sleepiness.

<sup>\*\*</sup> Hypoxic burden is defined as the total respiratory-event-related area under the oxygen-desaturation curve from a pre-event baseline and is expressed as % min/hr — the time (in minutes) spent in oxygen desaturation (%) per hour of sleep. This measure is calculated from a polysomnographic study that encapsulated frequency, duration, and depth of respiratory event-related oxygen desaturation, and data are geometric means (coefficient of variation, %).

<sup>††</sup> High-sensitivity C-reactive protein (hsCRP) data are geometric means (coefficient of variation, %).

tary Appendix). Safety assessments included adverse events and serious adverse events that occurred during the reporting period.

#### STATISTICAL ANALYSIS

The statistical analysis plan, including all the end points and assessments, was finalized and submitted before the time of the database lock and data unblinding. We calculated that a sample size of 206 participants per trial would provide the trial with at least 90% power to show the superiority of tirzepatide to placebo relative to the primary end point at a two-sided significance level of 0.05. In calculating the sample size for the primary end point, we assumed a mean 50% reduction in AHI, with a common standard deviation of 50% and a dropout rate of up to 25%. The updated primary end point was also deemed to have sufficient power; therefore, no adjustments to the sample-size calculation were made.

Data from all the participants who received at least one dose of tirzepatide or placebo (the intention-to-treat population) were used to analyze the efficacy and safety end points. For each trial, two estimands — the treatment-regimen estimand

and efficacy estimand — were used to assess the primary and key secondary end points from different perspectives, and the two estimands accounted for intercurrent events differently. The treatment-regimen estimand represented the average treatment effect of tirzepatide relative to placebo for all participants who had received at least one dose of tirzepatide or placebo regardless of whether they discontinued trial treatment or placebo for any reason. The efficacy estimand represented the average treatment effect of tirzepatide relative to placebo for all the participants if the treatment or placebo was administered as intended for the entire planned 52-week trial duration. All results are reported with the use of the treatment-regimen estimand unless otherwise specified. For the primary and key secondary end points, the type 1 error rate was controlled at a two-sided alpha level of 0.05 within each estimand and within each trial by means of a graphical testing procedure.<sup>19</sup> PROMIS-SRI and PROMIS-SD end points from both trials that were controlled for type 1 errors were pooled and tested with the use of a distinct graphical testing scheme to provide relevant power for



analysis under the submission-wise type 1 error-rate-control strategy (Fig. S2).<sup>20,21</sup> The populations of the two trials were suitable for pooling because of similar baseline PRO characteristics. The potential issue of PAP confounding of the PROMIS outcomes in trial 2 was minimized by a 7-day PAP washout period.<sup>22</sup>

Statistical analyses were conducted with the use of an analysis of covariance model, with the end point as a response variable, trial-group assignment and randomization strata as fixed effects (except for the severity of obstructive sleep apnea for AHI-related end points), and the baseline value as a covariate. Categorical variables of the proportion of participants who had at least 50% reduction in AHI and an AHI of less than 5 events per hour or an AHI of 5 to 14 events per hour with an ESS of 10 or less were evaluated with the use of logistic regression analysis with trial-group assignment, geographic region, baseline AHI, and sex as covariates. Baseline or post-baseline data were assumed to be missing at random or not at random depending on the reason for missingness. Data that were missing at random were handled through a multiple-imputation approach with the use of data from the same trial group. Data that were missing but not at random were imputed with the use of a placebo-based multiple-imputation approach. Full details on the estimands, handling of missing values, and statistical analysis methods are provided in the Supplementary Appendix.

## RESULTS

### PARTICIPANTS

The trials were conducted from June 21, 2022, through March 29, 2024. A total of 469 participants were randomly assigned to receive tirzepatide or placebo in trial 1 (234 participants) or trial 2 (235 participants) (Table 1). Overall, 82.9% of the participants completed the trial (91.5% in the tirzepatide groups and 74.4% in the placebo groups) and 79.7% adhered to the assigned regimen (87.6% in the tirzepatide groups and 71.9% in the placebo groups) (Fig. S3).

Demographic and baseline characteristics of the participants are shown in Table 1. In trial 1, the mean age of the participants was 47.9 years; most were male (67.1%) and White (65.8%), with a mean BMI of 39.1 and a mean AHI of 51.5 events per hour. In trial 2, the mean age was 51.7

years; most were male (72.3%) and White (73.1%), with a mean BMI of 38.7 and a mean AHI of 49.5 events per hour. Details regarding the geographic distribution and representativeness of the trial participants are shown in Tables S1 and S2.

### SLEEP-DISORDERED BREATHING-RELATED END POINTS

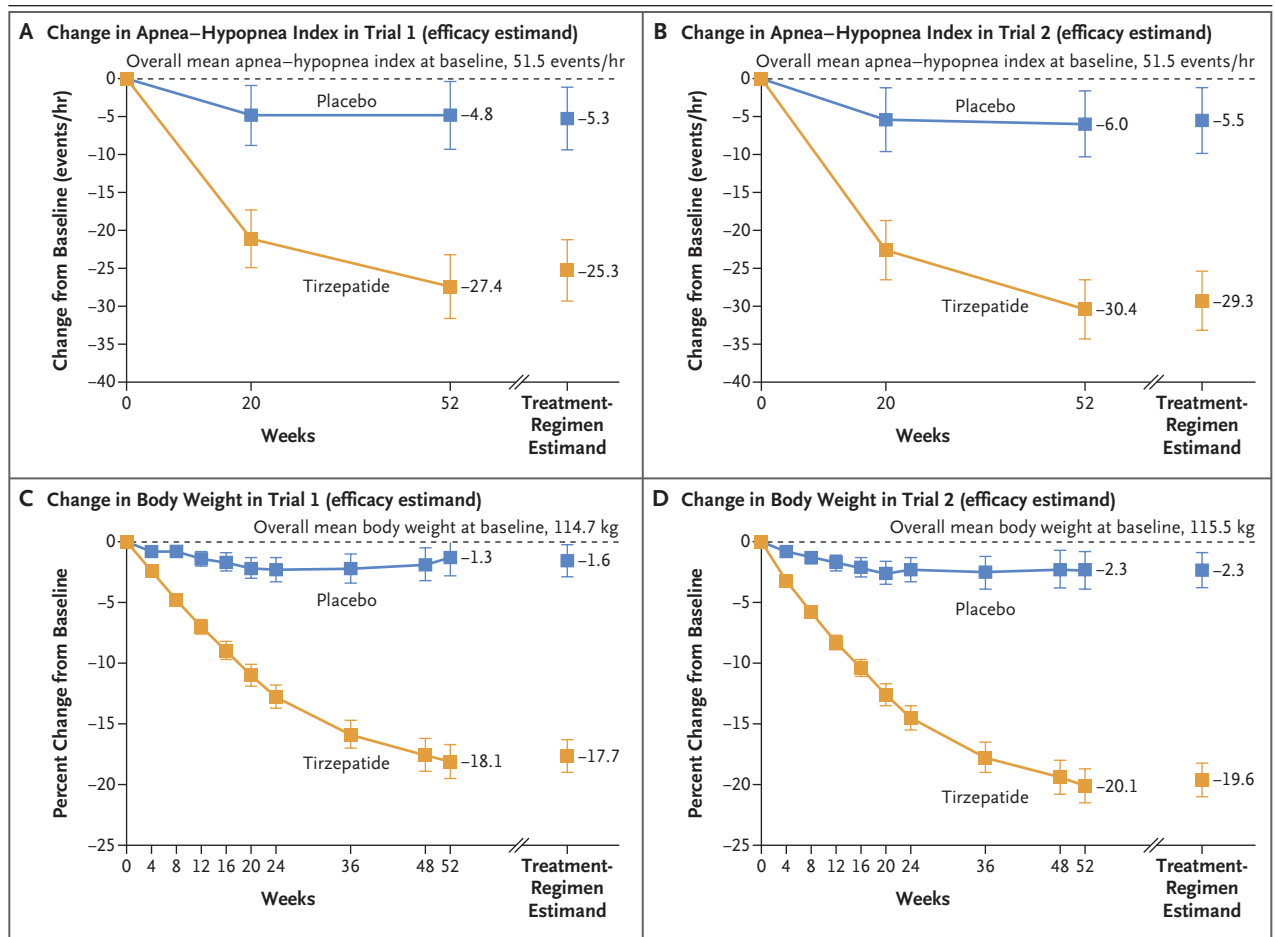
For the trial 1 treatment-regimen estimand, the change in AHI at week 52 was  $-25.3$  events per hour (95% confidence interval [CI],  $-29.3$  to  $-21.2$ ) with tirzepatide and  $-5.3$  events per hour (95% CI,  $-9.4$  to  $-1.1$ ) with placebo, for an estimated treatment difference of  $-20.0$  events per hour (95% CI,  $-25.8$  to  $-14.2$ ), ( $P < 0.001$ ) (Fig. 1A and Table 2). For the efficacy estimand, the change in AHI at week 52 was  $-27.4$  events per hour (95% CI,  $-31.6$  to  $-23.2$ ) with tirzepatide and  $-4.8$  events per hour (95% CI,  $-9.3$  to  $-0.3$ ) with placebo, for an estimated treatment difference of  $-22.5$  events per hour (95% CI,  $-28.7$  to  $-16.4$ ).

For the trial 2 treatment-regimen estimand, the change in AHI at week 52 was  $-29.3$  events per hour (95% CI,  $-33.2$  to  $-25.4$ ) with tirzepatide and  $-5.5$  events per hour (95% CI,  $-9.9$  to  $-1.2$ ) with placebo, for an estimated treatment difference of  $-23.8$  events per hour (95% CI,  $-29.6$  to  $-17.9$ ), ( $P < 0.001$ ) (Fig. 1B and Table 2). For the efficacy estimand, the change in AHI at week 52 with tirzepatide was  $-30.4$  events per hour (95% CI,  $-34.3$  to  $-26.5$ ) with tirzepatide and  $-6.0$  events per hour (95% CI,  $-10.3$  to  $-1.6$ ) with placebo, for an estimated treatment difference of  $-24.4$  events per hour (95% CI,  $-30.3$  to  $-18.6$ ). The change over time in AHI in the efficacy estimand is shown in Figures 1A and 1B.

Participants in both trials who received tirzepatide had significant reductions in AHI and in the sleep apnea-specific hypoxic burden (Table 2). The percentages of participants who had a reduction in the AHI of 50% or more at week 52 and the percentages who had an AHI of less than 5 events per hour or an AHI of 5 to 14 events per hour and an ESS of 10 or less at week 52 are reported in Table 2.

### CHANGE IN PROS AND CARDIOVASCULAR RISK FACTORS

In a pooled trial 1 and trial 2 analysis, participants who received tirzepatide had significant reductions in PROMIS-SRI and PROMIS-SD T scores (Table 3). PROMIS data that were analyzed separately for



**Figure 1. Change in AHI and Body Weight.**

The change in the apnea–hypopnea index (AHI, the number of apneas and hypopneas during an hour of sleep) (Panels A and B) and body weight (Panels C and D) from baseline to week 52 for trial 1 and trial 2 are shown according to the weeks since randomization, derived from a mixed-model-for-repeated-measures analysis for the efficacy estimand, and no explicit imputations were performed for missing data. Week 52 estimates for the treatment-regimen estimand are also shown. For the treatment-regimen estimand, missing data at week 52 due to coronavirus disease 2019, missing data at week 52 after trial discontinuation due to the participant having undergone randomization in error, or missing data at baseline were assumed to be missing at random and were imputed with the use of multiple imputation from the same trial group. All other missing data at week 52 were considered to be not missing at random, and a placebo-based multiple imputation method was implemented. Least-squares means are shown unless otherwise noted. I bars indicate 95% confidence intervals.

each trial showed reductions similar to those shown in analysis of the pooled data (Table S4). Participants in both trials who received tirzepatide had significant reductions in body weight (Figs. 1C and 1D), systolic blood pressure, and hsCRP concentrations (Table 2).

#### SAFETY

Adverse events that occurred during the receipt of tirzepatide or placebo were reported by 79.8% of the participants who received tirzepatide and 76.7% of those who received placebo in trial 1

and by 83.2% of the participants who received tirzepatide and 72.8% of those who received placebo in trial 2 (Table 4). The most frequently reported adverse events were generally gastrointestinal and occurred more frequently in the participants who received tirzepatide. These events were generally mild-to-moderate in severity and occurred most frequently during the dose-escalation phase.

Serious adverse events were reported by 35 participants (7.5%) overall. Similar percentages of participants in the tirzepatide and placebo groups

**Table 2. Primary and Key Secondary End Points According to Trial Group for the Treatment-Regimen Estimand.\***

End Point	Trial 1		Trial 2	
	Tirzepatide N = 114	Placebo N = 120	Tirzepatide N = 120	Placebo N = 115
<b>Primary end point</b>				
Change in AHI (95% CI) — no. of events/hr	-25.3 (-29.3 to -21.2)	-5.3 (-9.4 to -1.1)	-29.3 (-33.2 to -25.4)	-5.5 (-9.9 to -1.2)
<b>Key secondary end points</b>				
Percent change in AHI (95% CI)	-50.7 (-62.3 to -39.1)	-3.0 (-16.9 to 10.9)	-58.7 (-69.1 to -48.4)	-2.5 (-16.2 to 11.2)
Reduction of ≥50% in AHI events at wk 52 — no. (%)	70 (61.2)	23 (19.0)	86 (72.4)	27 (23.3)
AHI of <5 or AHI of 5 to 14 with ESS ≤10 at wk 52 — no. (%)	48 (42.2)	19 (15.9)	60 (50.2)	16 (14.3)
Percent change in body weight (95% CI)	-17.7 (-19.0 to -16.3)	-1.6 (-2.9 to -0.2)	-19.6 (-21.0 to -18.2)	-2.3 (-3.8 to -0.9)
Change in hsCRP concentration at wk 52 (95% CI) — mg/dl	-1.4 (-1.7 to -1.1)	-0.7 (-1.1 to -0.3)	-1.4 (-1.6 to -1.1)	-0.3 (-0.8 to 0.1)
Change in sleep apnea-specific hypoxic burden at wk 52 (95% CI) — % min/hr	-95.2 (-103.2 to -87.2)	-25.1 (-44.3 to -5.9)	-103.0 (-110.3 to -95.6)	-41.7 (-63.9 to -19.5)
Change in systolic blood pres- sure at wk 48 (95% CI) — mm Hg	-9.5 (-11.5 to -7.5)	-1.8 (-3.9 to 0.2)	-7.6 (-9.7 to -5.6)	-3.9 (-6.3 to -1.6)
<b>Additional secondary end point†:</b>				
Change in diastolic blood pres- sure at wk 48 (95% CI) — mm Hg	-4.9 (-6.4 to -3.5)	-2.1 (-3.6 to -0.6)	-3.3 (-4.7 to -1.9)	-2.2 (-3.8 to -0.6)

\* Data are least-squares means with 95% confidence intervals or numbers and percents of patients, unless otherwise stated. Relative risks are calculated using g-computation methods<sup>24</sup> from logistic regression. P values for categorical end points are based on a logistic regression model. All changes are from baseline to week 52 with the exception of blood pressure, which was change from baseline to week 48 to prevent suspension of PAP therapy in trial 2 from confounding the assessment.

† Differences between the groups are presented as the estimated treatment difference with the exception of the week-52 categories of reduction of ≥50% in AHI events and AHI of <5 or 5 to 14 with ESS ≤10, which are shown as relative risk. Estimated treatment differences for the secondary end points are the differences in the least-squares mean changes. P < 0.001 for the primary and key secondary end points with the exceptions of the change in hsCRP concentration at week 52 in trial 1 (P = 0.004) and the change in systolic blood pressure at week 48 in trial 2 (P = 0.02).

‡ The confidence intervals for this end point have not been adjusted for multiplicity and should not be used to make inferences.



**Table 3. Pooled Trial 1 and Trial 2 Patient-Reported Outcomes.\***

Variable	Tirzepatide (N = 234)	Placebo (N = 233)	Estimated Treatment Difference (95% CI)	P Value
Change in PROMIS Sleep-related Impairment T score	-7.5 (-8.8 to -6.3)	-3.6 (-4.9 to -2.3)	-3.9 (-5.7 to -2.2)	<0.001
Change in PROMIS Sleep Disturbance T score	-5.7 (-6.8 to -4.7)	-2.7 (-3.8 to -1.6)	-3.1 (-4.5 to -1.5)	<0.001

\* Data are least-squares means (95% confidence interval) unless otherwise stated. All changes shown are from baseline to week 52 in the modified intention-to-treat population.

reported serious adverse events. There were two adjudicated confirmed cases of acute pancreatitis in the trial 2 tirzepatide group. No cases of medullary thyroid cancer were reported. There were five cases of severe or serious depressive disorder or suicidal ideation or behavior events across both trials (two with tirzepatide and three with placebo). There were no deaths reported in either trial.

## DISCUSSION

In the present trials involving adults with moderate-to-severe obstructive sleep apnea and obesity, the AHI decreased significantly by up to 29.3 events per hour (a 58.7% change from baseline) among the participants who received tirzepatide, as compared with a decrease of up to 5.3 events per hour (a 3.0% change from baseline) among those who received placebo. This change is considered clinically relevant; the American Academy of Sleep Medicine defines the clinical significance threshold for the AHI as 15 or more events per hour,<sup>25</sup> and other sources have proposed a 50% improvement in AHI as clinically relevant.<sup>26,27</sup> A meaningful percentage of participants who received tirzepatide (up to 50.2%) in both SURMOUNT-OSA trials met the combined key secondary end-point criteria of fewer than 5 AHI events per hour or 5 to 14 AHI events per hour and an ESS of 10 or less, which is relevant because these thresholds for disease severity represent a level at which PAP therapy may not be recommended.<sup>2,3,28,29</sup> The reductions in AHI were also accompanied by meaningful improvements in hypoxic burden, which better captures the obstructive sleep apnea-related risk of cardiovascular complications and death.<sup>18,30</sup> These reductions were consistent in both trials regardless of concomitant PAP therapy and may inform

treatment decisions about patients with or without PAP therapy. Patients with obstructive sleep apnea are sometimes unable or unwilling to adhere to PAP treatment, and PAP has not been shown to affect cardiovascular complications and death in obstructive sleep apnea; therefore, there is a need for additional treatment options.<sup>4-6,10</sup>

Symptoms of obstructive sleep apnea represent a substantial disease burden and increased risk of injury, including increased risk of motor vehicle accidents and work-related injuries.<sup>31</sup> The symptom severity in obstructive sleep apnea may also be a predictor of increased risk of cardiovascular complications.<sup>32</sup> Therefore, it is clinically relevant that in the current trials, positive effects of tirzepatide on the participants' sleep-related functioning and sleep disturbance were detected on the basis of PROMIS-SRI and PROMIS-SD scores.

Obstructive sleep apnea and obesity are two distinct but closely related diseases, and both have independent etiologic roles in the development of cardiovascular complications.<sup>10</sup> Current guidelines recommend weight reduction of 7 to 11% for patients with obstructive sleep apnea<sup>8,33</sup>; however, a recent meta-analysis reports additional weight reduction can further reduce the AHI.<sup>34</sup> This level of weight reduction has been difficult to accomplish with lifestyle intervention alone.<sup>35</sup> Bariatric surgery has shown benefits in adults with obstructive sleep apnea<sup>36</sup>; however, owing to the invasive nature of surgery, it is not a feasible approach for many persons with obstructive sleep apnea and obesity. In SURMOUNT-OSA, tirzepatide reduced blood pressure and inflammation, which are important risk factors for cardiovascular complications of obstructive sleep apnea with obesity.

The safety profile of tirzepatide was consistent with that observed in previous trials.<sup>14,37-39</sup> As

Table 4. Adverse Events and Safety.*				
Variable	Trial 1		Trial 2	
	Tirzepatide N=114	Placebo N=120	Tirzepatide N=119	Placebo N=114
	number (percent)			
≥1 Adverse event while receiving tirzepatide or placebo	91 (79.8)	92 (76.7)	99 (83.2)	83 (72.8)
Death	0	0	0	0
Serious adverse events	9 (7.9)	7 (5.8)	7 (5.9)	12 (10.5)
Adverse events leading to discontinuation of trial drug or placebo	5 (4.4)	2 (1.7)	4 (3.4)	8 (7.0)
Adverse events occurring in ≥5% of participants in any trial group				
Diarrhea	30 (26.3)	15 (12.5)	26 (21.8)	10 (8.8)
Nausea	29 (25.4)	12 (10.0)	26 (21.8)	6 (5.3)
Vomiting	20 (17.5)	5 (4.2)	11 (9.2)	1 (0.9)
Constipation	18 (15.8)	3 (2.5)	18 (15.1)	5 (4.4)
Eructation	9 (7.9)	0	10 (8.4)	1 (0.9)
Gastroesophageal reflux disease	9 (7.9)	1 (0.8)	6 (5.0)	0
Injection-site reaction	8 (7.0)	1 (0.8)	6 (5.0)	0
Abdominal pain	7 (6.1)	4 (3.3)	5 (4.2)	2 (1.8)
Upper respiratory tract infection	7 (6.1)	10 (8.3)	5 (4.2)	8 (7.0)
Coronavirus disease 2019	6 (5.3)	10 (8.3)	8 (6.7)	11 (9.6)
Nasopharyngitis	3 (2.6)	8 (6.7)	15 (12.6)	12 (10.5)
Dyspepsia	5 (4.4)	2 (1.7)	11 (9.2)	1 (0.9)
Gastroenteritis	3 (2.6)	4 (3.3)	8 (6.7)	1 (0.9)
Upper abdominal pain	4 (3.5)	2 (1.7)	7 (5.9)	2 (1.8)
Influenza	4 (3.5)	8 (6.7)	3 (2.5)	3 (2.6)
Arthralgia	3 (2.6)	6 (5.0)	4 (3.4)	5 (4.4)
Bronchitis	0	0	3 (2.5)	7 (6.1)
Hypertension	1 (0.9)	8 (6.7)	2 (1.7)	2 (1.8)
Other adverse events of special interest				
Severe hypoglycemia	0	1 (0.8)	0	1 (0.9)
Adjudication-confirmed MACE†	0	0	0	1 (0.9)
Arrhythmias or cardiac conduction disorders	7 (6.1)	9 (7.5)	6 (5.0)	2 (1.8)
Severe or serious gastrointestinal events‡	4 (3.5)	0	4 (3.4)	0
Severe or serious hepatic events	0	0	0	0
Severe or serious acute renal events	0	0	1 (0.8)	0
Adjudication-confirmed acute pancreatitis	0	0	2 (1.7)	0
C-cell hyperplasia or thyroid cancer	0	0	0	0
Severe or serious major depressive disorder or suicidal behavior and ideation events	2 (1.8)	1 (0.8)	0	2 (1.8)
Severe or serious allergic or hypersensitivity reactions§	0	0	0	0

\* Adverse events are classified according to the preferred terms in the *Medical Dictionary for Regulatory Activities*, version 26.1.

† Key major adverse cardiac events (MACE) were cardiovascular death, nonfatal myocardial infarction, nonfatal stroke, hospitalization for unstable angina, and hospitalization for heart failure.

‡ In trial 1, two participants had diarrhea, one had gastroesophageal reflux disease, and one had nausea. In trial 2, three participants had diarrhea, two had nausea, and one had acute pancreatitis. Participants could be counted in more than one category.

§ This category includes injection-site reactions and antidrug antibody formation.

typically observed with tirzepatide and GLP-1 receptor agonists, mild-to-moderate gastrointestinal events were the most frequently reported adverse events, occurring primarily during the dose-escalation period. There were no differences observed between tirzepatide recipients and placebo recipients with regard to reported gallbladder-related events or hepatic and renal events.

The current trials have several strengths. They were conducted globally, were adequately sized, ensured relevant representation of women (who typically represent a minority in obstructive sleep apnea trials), and assessed multiple obstructive sleep apnea–related, patient-reported, and cardiovascular-related end points that covered the burden of obstructive sleep apnea. Therefore, we believe that our findings are generalizable. The weight reduction among the participants who received placebo was similar to the results with placebo in other antiobesity medication studies, a finding that suggests a similar level of adherence to the lifestyle intervention. On the basis of the current evidence regarding PAP treatment, the American Academy of Sleep Medicine guidelines prioritize treatment with PAP for patients with symptoms of obstructive sleep apnea.<sup>25</sup> Future treatments of obstructive sleep apnea should also address obstructive sleep apnea–related cardiovascular risk, which is associated with moderate-to-severe obstructive sleep apnea regardless of symptoms.<sup>40</sup> Therefore, it is important that the enrollment in our trials was not limited to participants with current symptoms, and the results may inform broader treatment decisions in future clinical practice. Finally, the design of two independent trials involving participants with and without current PAP therapy provides insights into the effect of tirzepatide treatment in these patient populations that are prevalent in clinical practice.

Interpretation of the current findings should take into account the potential limitations of our

trials. First, the design and shorter duration of the current trials does not support the assessment of long-term cardiovascular outcomes. The ongoing SURMOUNT-Morbidity and Mortality in Obesity trial (ClinicalTrials.gov number, NCT05556512) may provide additional information. Second, our trials excluded participants who did not have obesity and did not analyze the effect of the intervention in people with overweight or normal BMI. Third, trial 2 was not designed to investigate the potential effect of treatment interventions on adherence to PAP treatment, and this was not prespecified as an end point for analysis in trial 2. Fourth, the trials were not designed to assess whether the results differed according to the presence of participant’s symptoms at baseline. Fifth, the thresholds for the minimum clinically important changes for PROMIS-SRI and PROMIS-SD have not been established in clinical practice yet. Therefore, the clinical importance of the observed improvements remains to be evaluated. Finally, although obstructive sleep apnea affects patients’ lives over a period of many years, the trials did not investigate a period of treatment with tirzepatide longer than 52 weeks.

In two trials, the participants who received tirzepatide had a clinically meaningful change in sleep-disordered breathing and alleviation of perceived sleep disturbance and sleep-related impairment, as well as reductions in common obstructive sleep apnea-related cardiovascular risk factors.

Presented in part at the 84th Scientific Sessions of the American Diabetes Association, Orlando, FL, June 21–24, 2024.

Supported by Eli Lilly.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

We thank the participants and caregivers involved in the trial, as well as Eli Lilly employees Casey J. Mast, Pharm.D., for medical writing and editorial assistance; Bram Brouwers, Ph.D., and Meri Kay Scott, Ph.D., for trial design advice; and Angel Rodriguez, M.D., Ph.D., for a critical review of an earlier version of the manuscript.

## REFERENCES

- Jordan AS, McSharry DG, Malhotra A. Adult obstructive sleep apnoea. *Lancet* 2014;383:736-47.
- Veasey SC, Rosen IM. Obstructive sleep apnea in adults. *N Engl J Med* 2019;380:1442-9.
- Benjafield AV, Ayas NT, Eastwood PR, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. *Lancet Respir Med* 2019;7:687-98.
- McEvoy RD, Antic NA, Heeley E, et al. CPAP for prevention of cardiovascular events in obstructive sleep apnea. *N Engl J Med* 2016;375:919-31.
- Peker Y, Glantz H, Eulenburg C, Wegscheider K, Herlitz J, Thunström E. Effect of positive airway pressure on cardiovascular outcomes in coronary artery disease patients with nonsleepy obstructive sleep apnea: the RICCADSA randomized controlled trial. *Am J Respir Crit Care Med* 2016;194:613-20.
- Sánchez-de-la-Torre M, Sánchez-de-la-Torre A, Bertran S, et al. Effect of obstructive sleep apnoea and its treatment with continuous positive airway pressure on

- the prevalence of cardiovascular events in patients with acute coronary syndrome (ISAACC study): a randomised controlled trial. *Lancet Respir Med* 2020;8:359-67.
7. Francis CE, Quinnell T. Mandibular advancement devices for OSA: an alternative to CPAP? *Pulm Ther* 2021;7:25-36.
  8. Hudgel DW, Patel SR, Ahasic AM, et al. The role of weight management in the treatment of adult obstructive sleep apnea: an official American Thoracic Society clinical practice guideline. *Am J Respir Crit Care Med* 2018;198(6):e70-e87.
  9. Morgenthaler TI, Kapen S, Lee-Chiong T, et al. Practice parameters for the medical therapy of obstructive sleep apnea. *Sleep* 2006;29:1031-5.
  10. Javaheri S, Barbe F, Campos-Rodriguez F, et al. Sleep apnea: types, mechanisms, and clinical cardiovascular consequences. *J Am Coll Cardiol* 2017;69:841-58.
  11. Heinzer R, Eckert D. Treatment for obstructive sleep apnoea and cardiovascular diseases: are we aiming at the wrong target? *Lancet Respir Med* 2020;8:323-5.
  12. Yeghiazarians Y, Neid H, Tietjens JR, et al. Obstructive sleep apnea and cardiovascular disease: a scientific statement from the American Heart Association. *Circulation* 2021;144(3):e56-e67.
  13. Coskun T, Sloop KW, Loghin C, et al. LY3298176, a novel dual GIP and GLP-1 receptor agonist for the treatment of type 2 diabetes mellitus: from discovery to clinical proof of concept. *Mol Metab* 2018;18:3-14.
  14. Jastreboff AM, Aronne LJ, Ahmad NN, et al. Tirzepatide once weekly for the treatment of obesity. *N Engl J Med* 2022;387:205-16.
  15. Wilson JM, Lin Y, Luo MJ, et al. The dual glucose-dependent insulinotropic polypeptide and glucagon-like peptide-1 receptor agonist tirzepatide improves cardiovascular risk biomarkers in patients with type 2 diabetes: a post hoc analysis. *Diabetes Obes Metab* 2022;24:148-53.
  16. Malhotra A, Bednarik J, Chakladar S, et al. Tirzepatide for the treatment of obstructive sleep apnea: rationale, design, and sample baseline characteristics of the SURMOUNT-OSA phase 3 trial. *Contemp Clin Trials* 2024;141:107516.
  17. Berry R, Brooks R, Gamaldo C, et al. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, version 2.0.2. Darien, IL: American Academy of Sleep Medicine, 2012.
  18. Azarbarzin A, Sands SA, Stone KL, et al. The hypoxic burden of sleep apnoea predicts cardiovascular disease-related mortality: the Osteoporotic Fractures in Men Study and the Sleep Heart Health Study. *Eur Heart J* 2019;40:1149-57.
  19. Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Stat Med* 2009;28:586-604.
  20. Bretz F, Xi D. Commentary on "Statistics at FDA: reflections on the past six years." *Stat Biopharm Res* 2019;11:20-5 (<https://www.tandfonline.com/doi/full/10.1080/19466315.2018.1562369>).
  21. Vandemeulebroecke M, Häring DA, Hua E, Wei X, Xi D. New strategies for confirmatory testing of secondary hypotheses on combined data from multiple trials. *Clin Trials* 2024;21:171-9.
  22. Schwarz EI, Stradling JR, Kohler M. Physiological consequences of CPAP therapy withdrawal in patients with obstructive sleep apnoea—an opportunity for an efficient experimental model. *J Thorac Dis* 2018;10:Suppl 1:S24-S32.
  23. Donovan LM, Yu L, Bertisch SM, Buysse DJ, Rueschman M, Patel SR. Responsiveness of patient-reported outcomes to treatment among patients with type 2 diabetes mellitus and OSA. *Chest* 2020;157:665-72.
  24. Ye T, Bannick M, Yi Y, Shao J. Robust variance estimation for covariate-adjusted unconditional treatment effect in randomized clinical trials with binary outcomes. *Stat Theory Relat Fields* 2023;7:159-63.
  25. Patil SP, Ayappa IA, Caples SM, Kimoff RJ, Patel SR, Harrod CG. Treatment of adult obstructive sleep apnea with positive airway pressure: an American Academy of Sleep Medicine systematic review, meta-analysis, and GRADE assessment. *J Clin Sleep Med* 2019;15:301-34.
  26. Ramar K, Dort LC, Katz SG, et al. Clinical practice guideline for the treatment of obstructive sleep apnea and snoring with oral appliance therapy: an update for 2015. *J Clin Sleep Med* 2015;11:773-827.
  27. Lin H-C, Friedman M, Chang H-W, Gurpinar B. The efficacy of multilevel surgery of the upper airway in adults with obstructive sleep apnea/hypopnea syndrome. *Laryngoscope* 2008;118:902-8.
  28. Epstein LJ, Kristo D, Strollo PJ Jr, et al. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J Clin Sleep Med* 2009;5:263-76.
  29. Centers for Medicare and Medicaid Services. Positive airway pressure (PAP) devices: complying with documentation & coverage requirements. October 2016 ([https://www.cms.gov/files/document/papdoccvfactsheeticn905064\\_textonly.pdf](https://www.cms.gov/files/document/papdoccvfactsheeticn905064_textonly.pdf)).
  30. Azarbarzin A, Sands SA, Taranto-Montemurro L, et al. The sleep apnea-specific hypoxic burden predicts incident heart failure. *Chest* 2020;158:739-50.
  31. Bonsignore MR, Randerath W, Schiza S, et al. European Respiratory Society statement on sleep apnoea, sleepiness and driving risk. *Eur Respir J* 2021;57:2001272.
  32. Baratta F, Pastori D, Fabiani M, et al. Severity of OSAS, CPAP and cardiovascular events: a follow-up study. *Eur J Clin Invest* 2018;48(5):e12908.
  33. Garvey WT, Mechanick JI, Brett EM, et al. American Association of Clinical Endocrinologists and American College of Endocrinology comprehensive clinical practice guidelines for medical care of patients with obesity. *Endocr Pract* 2016;22:Suppl 3:1-203.
  34. Malhotra A, Curtis A, Dunn J, Bunck MC, Bednarik J. Weight reduction and the impact on apnea-hypopnea index: a meta-analysis. *Sleep* 2023;46:Suppl 1:A251-A252 ([https://academic.oup.com/sleep/article/46/Supplement\\_1/A251/7182261](https://academic.oup.com/sleep/article/46/Supplement_1/A251/7182261)).
  35. Dombrowski SU, Knittle K, Avenell A, Araújo-Soares V, Sniehotta FF. Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials. *BMJ* 2014;348:g2646.
  36. Zhang Y, Wang W, Yang C, Shen J, Shi M, Wang B. Improvement in nocturnal hypoxemia in obese patients with obstructive sleep apnea after bariatric surgery: a meta-analysis. *Obes Surg* 2019;29:601-8.
  37. Garvey WT, Frias JP, Jastreboff AM, et al. Tirzepatide once weekly for the treatment of obesity in people with type 2 diabetes (SURMOUNT-2): a double-blind, randomised, multicentre, placebo-controlled, phase 3 trial. *Lancet* 2023;402:613-26.
  38. Wadden TA, Chao AM, Machineni S, et al. Tirzepatide after intensive lifestyle intervention in adults with overweight or obesity: the SURMOUNT-3 phase 3 trial. *Nat Med* 2023;29:2909-18.
  39. Aronne LJ, Sattar N, Horn DB, et al. Continued treatment with tirzepatide for maintenance of weight reduction in adults with obesity: the SURMOUNT-4 randomized clinical trial. *JAMA* 2024;331:38-48.
  40. Young T, Palta M, Dempsey J, Peppard PE, Nieto FJ, Hla KM. Burden of sleep apnea: rationale, design, and major findings of the Wisconsin Sleep Cohort study. *WMJ* 2009;108:246-9.

Copyright © 2024 Massachusetts Medical Society.

## Supplementary Appendix

Supplement to: Malhotra A, Grunstein RR, Fietze I, et al. Tirzepatide for the treatment of obstructive sleep apnea and obesity. *N Engl J Med*. DOI: 10.1056/NEJMoa2404881

This appendix has been provided by the authors to give readers additional information about the work.



Supplementary Appendix  
SURMOUNT-OSA TRIAL

Contents

LIST OF INVESTIGATORS .....	4
SUPPLEMENTAL METHODS .....	8
<b>INCLUSION AND EXCLUSION CRITERIA</b> .....	8
<b>STATISTICAL ANALYSIS METHODS</b> .....	13
<b>Integrated Efficacy Analysis Plan (IEAP)</b> .....	18
Rationale for Change of Key Secondary Endpoints .....	21
Pre-specified Endpoints .....	22
<b>Master Protocol Endpoints</b> .....	23
<b>Study 1 Endpoints</b> .....	25
<b>Study 2 Endpoints</b> .....	25
AUTHOR CONTRIBUTIONS .....	26
SUPPLEMENTARY FIGURES .....	27
<b>Figure S1. SURMOUNT-OSA study design.</b> .....	27
<b>Figure S2. Pre-Specified Multiplicity Control for SURMOUNT-OSA Master Protocol.</b> .....	28
<b>Figure S3. Participant disposition from randomization to primary end point.</b> .....	29
SUPPLEMENTARY TABLES .....	30
<b>Table S1. Patient Geography</b> .....	30

<b>Table S2. Representativeness of Study</b> .....	31
<b>Table S3. Summary of Missing Data for Primary and Key Secondary Endpoints</b> .....	32
<b>Table S4. PROMIS Outcomes by Study</b> .....	33
<b>REFERENCES</b> .....	34

## LIST OF INVESTIGATORS

1. Prof. Danny J. Eckert PhD. Adelaide Institute for Sleep Health Loc. 1, 5 Laffer Drive, Bedford Park, South Australia, 5042 Australia
2. Prof. Ronald R. Grunstein MD, PhD. Sleep & Circadian Research Group, Woolcock Institute of Medical Research Loc. 1, 431 Glebe Point Road, Sydney, New South Wales, 2037 Australia
3. Dr. Luis Henrique Santos Canani MD. Núcleo de Pesquisa Clínica do Rio Grande do Sul, Mostardeiro, 5, Room 310, Porto Alegre, Rio Grande do Sul, 90430-001 Brazil
4. Dr. Bruno Halpern PhD. BR Trials - Ensaios Clínicos e Consultoria, 81 Praça Tomás Morus, 9º andar, São Paulo, São Paulo, 05003-090 Brazil
5. Ms. Cintia Cercato PhD. Hospital das Clínicas FMUSP, 155 Avenida Doutor Enéas Carvalho de Aguiar, PAMP- 4º floor - 15 Block, São Paulo, São Paulo, 05403-000 Brazil
6. Dr. Mariana Delgado Fernandes MD. CPQuali Pesquisa Clínica, 916, Avenida Angélica, conjunto 506 Santa Cecília, São Paulo, São Paulo, 01228-000 Brazil
7. Dr. Breno Balabem Alves MD. CEPIC - Centro Paulista de Investigação Clínica, Rua Moreira e Costa 342, São Paulo, São Paulo, 04266-010 Brazil
8. Dr. Jose Francisco Kerr Saraiva MD. Instituto de Pesquisa clinica de Campinas, 1209 Rua Oswaldo Oscar Barthelson, Campinas, São Paulo, 13060-080 Brazil
9. Dr. Denise Reis Franco MD. CPCLIN, Av. Angélica nº2162, Consolação, São Paulo, São Paulo, 01228-200 Brazil
10. Dr. Zhenmei An MD. Endocrinology Loc. 1, No.37 Guoxue Alley, The third hospitalization building Wuhou District, Cheng Du, Sichuan, 610041 China
11. Prof. Ming Liu MD. Endocrinology & Metabolism Loc. 1, No.154 Anshan Avenue, Tianjin, Tianjin, 300052 China
12. Prof. Xiaoying Li MD, PhD. Endocrinology & Metabolism Loc. 1, No.180 Fenglin Road, Xuhui District, Shanghai, 4th Floor, 9th Building, Shanghai, Shanghai, 200032 China
13. Dr. Guixia Wang MD. The First Hospital of Jilin University, No. 71 Xin Min Road, Changchun, Jilin, 130021 China
14. Dr. Lixin Guo MD. Beijing Hospital, No. 1 Da Hua Road, Dongcheng District, Beijing, Beijing, 100730 China
15. Dr. Lan Xu MD. Endocrinology Loc. 1, No. 299 Qingyang Road, Wuxi, Jiangsu, 214023 China
16. Dr. Yibing Lu MD. Endocrinology Loc. 1, No. 121 Jiangjiayuan Road, Nanjing, Jiangsu, 210011 China
17. Dr. Vera Klokocnikova MD. Praglandia s.r.o, 253 Ostrovského, Zenske domovy, 3. patro, Praha 5, Praha 5, 150 00 Czech Republic
18. Dr. Martin Matoulek MD, PhD. III. Interni lkinika - Obezitologicke centrum Loc. 1, U Nemocnice 504/1, Praha 5, Praha 5, 12808 Czech Republic
19. Ms. Katharina Lederer Master of Neuroscience and Approbation. Advanced Sleep Research, Luisenstrasse 54-55, Third Floor, Berlin, Berlin, 10117 Germany
20. Prof. Jens Aberle MD. Diabeteszentrum Hamburg West, Beselerstrasse 2a, Hamburg, Hamburg, 22607 Germany

21. Dr. Thomas Schürholz Dr. Med. InnoDiab Forschung GmbH, Eleonorastraße 42, Essen, Nordrhein-Westfalen, 45136 Germany
22. Dr. Andreas Deimling Dr. Med. Lungenpraxis Schleswig, Auf der Freiheit 4, Schleswig, Schleswig-Holstein, 24837 Germany
23. Dr. Thomas Schaum Dr. Med. RED-Institut GmbH, Markt 15, Oldenburg, Schleswig-Holstein, 23758 Germany
24. Mr. Sascha Brinkmann. Siteworks GmbH, 21 Niemeyerstraße, Zentrum für klinische Studien Hannover, Hannover, Niedersachsen, 30449 Germany
25. Dr. Ludger Rose MD. Institut für Diabetesforschung GmbH Münster, Hohenzollernring 70, Münster, Nordrhein-Westfalen, 48145 Germany
26. Dr. Yoshimitsu Yamasaki MD, PhD. AMC Nishiumeda Clinic, 3-3-45 Umeda, Osaka, Osaka, 530-0001 Japan
27. Dr. Iwao Gohma MD. Sakai City Medical Center, 1-1-1, Ebaraji-cho, Nishi-ku, Sakai, Osaka, 593-8304 Japan
28. Dr. Tetsuya Kawano MD. Kirigaokatsuda Hospital, 3-9-20 Kirigaoka Kokourakita-ku, Kitakyushu, Fukuoka, 802-0052 Japan
29. Dr. Akinori Ishihara MD. Koujunkai Daido Clinic, 8 Hakusuicho\_Minami-ku, Nagoya, Aichi, 457-8511 Japan
30. Dr. Yasushi Fukushima MD, PhD. Fukuwa Clinic, 1-1-6 Kyobashi, Chuo-ku, Tokyo, 104-0031 Japan
31. Dr. Ryutaro Shirahama MD, PhD. RESM Respiratory and Sleep Medical Care Clinic, 3-8-12 Shinyokohama, Yokohama, Kanagawa, 222-0033 Japan
32. Dr. Arihiro Kiyosue MD, PhD. Tokyo-Eki Center-Building Clinic, 3-3-14, Nihombashi, Chuo-ku, Tokyo, 103-0027 Japan
33. Dr. Mitsutaka Taniguchi MD. Osaka Kaisei Hospital, 1-6-10 Miyahara, Yodogawa-ku, Osaka, Osaka, 532-0003 Japan
34. Dr. Luis Alejandro Nevarez Ruiz MD, Internal Medicine. Investigacion En Salud Y Metabolismo S.C., Haciendas del Valle #7120, Consultorio #408 Col. Haciendas del Valle, Chihuahua, Chihuahua, 31217 Mexico
35. Dr. Hector Glenn Valdez Lopez MD, Pneumology. Servicios Integrales Nova de Monterrey S.A. de C.V., Avenida Del Bosque No. 139, Colonia Cuauhtémoc, San Nicolás de los Garza, San Nicolas de los Garza, Nuevo León, 66450 Mexico
36. Dr. Amado Ramirez Hernandez MD. Arké SMO S.A de C.V, Alacio Perez # 928 int. 101-104, 109 and 111 Colonia Ignaci Zaragoza, Veracruz, Veracruz, 91910 Mexico
37. Dr. Marisol Herrera Marmolejo MD. Unidad de Investigación Clínica y Atención Médica HEPA S.C, Tierra de Fuego 3262 piso 2 y piso 3, Col. Prados Providencia, Guadalajara, Jalisco, 44670 Mexico
38. Dr. Ramiro Banda Elizondo PhD. Unidad Médica para la Salud Integral, Calle Río Conchos, 440 Col. Riberas de las Puentes, San Nicolás de los Garza, Nuevo León, 66465 Mexico
39. Dr. Maria Arechavaleta Granell MD. Arechavaleta Granell Maria del Rosario, Avenida México 2343 Col. Ladrón de Guevara, Guadalajara, Jalisco, 44670 Mexico

40. Dr. Melchor Alpizar Sallazar MD. Centro Especializado En Diabetes, Obesidad Y Prevencion De Enfermedades Cardiovasculares S.C., Calle 3 No 7 Colonia Reforma Social, Mexico City, Distrito Federal, 11650 Mexico
41. Dr. Mariana Zolandi Crespo Hernández MD. RM Pharma Specialists S.A. de C.V., Amores 734, Col. Del Valle, Mexico City, Distrito Federal, 3100 Mexico
42. Dr. Yi-Ching Yang MD. National Cheng Kung University Hospital, 138, SHENG-LI RD., TAINAN, Tainan, 704 Taiwan
43. Dr. Liang-Wen Hang MD, PhD. China Medical University Hospital, 2, Yude Road North Dist., Taichung, Taichung, 404332 Taiwan
44. Dr. Ryan S Drake DO. NeuroScience Research Center, 4048 Dressler Road Northwest, Canton, OH, 44718 United States
45. Dr. John Douglas Hudson MD. FutureSearch Trials of Neurology, 1600 W 38th Street, Ste 404, Austin, TX, 78731 United States
46. Dr. Haresh Boghara MD. Epic Medical Research, 106 Plaza Drive, Red Oak, TX, 75154 United States
47. Dr. James Michael Andry MD. Sleep Therapy Research Center, 5290 Medical Drive, San Antonio, TX, 78229 United States
48. Dr. Blair Brengle MD. Brengle Family Medicine, 8803 North Meridian Street, Indianapolis, IN, 46260 United States
49. Dr. Mira Baron Dr. Med. Palm Beach Research Center, 2277 Palm Beach Lakes Boulevard, West Palm Beach, FL, 33409 United States
50. Dr. Raj Karunakara MD. Renstar Medical Research, 21 NE 1st Avenue, Ocala, FL, 34470 United States
51. Dr. Kevin T. Prier MD. Rocky Mountain Clinical Research, 3910 Washington Parkway, Idaho Falls, ID, 83404 United States
52. Dr. Michael J Lillestol MD. Lillestol Research, 4450 31st AVE S, STE 101, Fargo, ND, 58104 United States
53. Dr. Dennis Lacey MD. NeuroTrials Research Inc., 5887 Glenridge Drive, Suite 400 Fulton County, Atlanta, GA, 30328 United States
54. Dr. Frances Eileen Broyles MD. Rainier Clinical Research Center Loc. 14, 800 Southwest 39th Street #110, Renton, WA, 98057 United States
55. Dr. Bryce A. Palchick MD. Preferred Primary Care Physicians, Preferred Clinical Research (Ofc 18) Loc. 3, 140 Curry Hollow Road, Suite 3, Pittsburgh, PA, 15236 United States
56. Dr. Aamr Aamr Herekar MD. ANESC Research, 7100 Westwood Drive, 300C, El Paso, TX, 79912 United States
57. Dr. Daniel G. Lorch MD. Teradan Clinical Trials, LLC Loc. 11, 910 Oakfield Drive, Suite 201, Brandon, FL, 33511 United States
58. Dr. James Paul Maynard MD. CTI-CRC Loc. 5, 1775 Lexington Ave., Suite 200, Cincinnati, OH, 45212 United States
59. Dr. Paul A. Hartley MD. Preferred Clinical Research Loc. 1, 202 Jacob Murphy Lane, Suite 201, Uniontown, PA, 15401 United States



60. Dr. Stacey Layle MD. Artemis Institute for Clinical Research, 770 Washington Street, Suite 300,  
San Diego, CA, 92103 United States

## **SUPPLEMENTAL METHODS**

### **INCLUSION AND EXCLUSION CRITERIA**

#### **Inclusion Criteria**

Participants are eligible to be included in the study only if all of the following criteria apply:

#### **Age**

1. Participant must be at least 18 (or the legal age of consent in the jurisdiction in which the study is taking place) years of age, inclusive, at the time of signing the informed consent.

#### **Type of Participant and Disease Characteristics**

2. Previously diagnosed moderate-to-severe OSA with an AHI  $\geq 15$ , as diagnosed with PSG, home sleep apnea test (HSAT), or other method that meets local guidelines prior to Visit 1.
3. AHI  $\geq 15$  on PSG as part of the trial at Visit 1.
4. In the investigator's opinion, are well-motivated, capable, and willing to
  - a. learn how to self-inject study intervention, as required for this protocol (visually impaired persons who are not able to perform the injections must have the assistance of a sighted individual trained to inject study intervention; persons with physical limitations who are not able to perform the injections must have the assistance of an individual trained to inject study intervention)
  - b. inject study intervention (or receive an injection from a trained individual if visually impaired or with physical limitations), and
  - c. follow study procedures for the duration of the study, including, but not limited to: follow lifestyle advice (for example, dietary restrictions and exercise plan), maintain a study diary, and complete required questionnaires.

#### **Weight**

5. BMI  $\geq 30$  kg/m<sup>2</sup>.
6. Have a history of at least 1 self-reported unsuccessful dietary effort to lose body weight.

#### **Sex and Contraceptive/Barrier Requirements**

7. Males and females may participate in this trial.

Female participants must not be pregnant, intending to be pregnant, breastfeeding, or intending to breastfeed.

Contraceptive use by participants should be consistent with local regulations regarding the methods of contraception for those participating in clinical studies.

#### **Informed Consent**

8. Capable of giving signed informed consent, which includes compliance with the requirements and restrictions listed in the ICF and in the protocol.

#### **Exclusion Criteria**

Participants are excluded from the study if any of the following criteria apply:

#### **Medical Conditions**

### ***Diabetes-related***

9. Have T1DM or T2DM, history of ketoacidosis, or hyperosmolar state/coma.
10. HbA1c  $\geq$  6.5% ( $\geq$  48 mmol/mol) at Visit 1.

### ***OSA-related***

11. Any previous or planned surgery for sleep apnea or major ear, nose or throat surgery, including tonsillectomy and adenoidectomy that still may affect breathing at time of Visit 1. Inclusion of a participant with more minor ear, nose or throat surgery (for example, deviated septum) will be at the investigator's discretion.
12. Significant craniofacial abnormalities that may affect breathing at time of Visit 1.
13. Diagnosis of central or Mixed Sleep Apnea with % of mixed or central apneas/hypopneas  $\geq$ 50%, or diagnosis of Cheyne Stokes Respiration.
14. Diagnosis of Obesity Hypoventilation Syndrome or daytime hypercapnia.
15. Active device treatment of OSA other than PAP therapy (for example, dental appliance), or other treatment, that in the opinion of the investigator, may interfere with study outcomes, unless willing to stop treatment at Visit 1 and throughout the study.
16. Respiratory and neuromuscular diseases that could interfere with the results of the trial in the opinion of the investigator.

### ***Obesity-related***

17. Have a self-reported change in body weight  $>$ 5 kg within 3 months prior to screening.
18. Have a prior or planned surgical treatment for obesity (excluding liposuction or abdominoplasty if performed more than 1 year prior to screening).
19. Have or plan to have endoscopic and/or device-based therapy for obesity or have had device removal within the last 6 months (for example, mucosal ablation, gastric artery embolization, intragastric balloon, and duodenal-jejunal bypass sleeve).

### ***Other medical***

20. History of clinically relevant medical, behavioral, or psychiatric disorder, other than OSA, that is associated with insomnia or excessive sleepiness.
21. Impaired renal function, defined as eGFR  $<$ 30 mL/min/1.73 m<sup>2</sup>.
22. Have a known clinically significant gastric emptying abnormality (for example, severe gastroparesis or gastric outlet obstruction) or chronically take drugs that directly affect GI motility.
23. History of chronic or acute pancreatitis.
24. Thyroid-stimulating hormone outside of the range of 0.4 to 6.0 mIU/L at the screening visit.  
*Note:* Participants receiving treatment for hypothyroidism may be included, provided their thyroid hormone replacement dose has been stable for at least 3 months.  
*Note:* TSH values above the normal range can, in some patients, suggest subclinical hypothyroidism. If, in the investigator's opinion, the participant has subclinical hypothyroidism and may require initiation of thyroid hormone replacement during the course of the study, the participant should be excluded from the study.
25. Have obesity induced by other endocrinologic disorders (for example, Cushing Syndrome) or diagnosed monogenetic or syndromic forms of obesity (for example, Melanocortin 4 Receptor deficiency or Prader-Willi Syndrome).
26. Are, in the judgment of the investigator, actively suicidal and therefore deemed to be at significant risk for suicide.

27. Have answered “yes” to either Question 4 or Question 5 on the “Suicidal Ideation” portion of the C-SSRS or have answered “yes” to any of the suicide-related behaviors on the “suicidal behavior” portion of the C-SSRS, and the ideation or behavior occurred within the past month.
28. PHQ-9 score of 15 or more at Visit 1 or 2, prior to randomization.
29. Uncontrolled hypertension (SBP  $\geq$ 160 mmHg and/or DBP  $\geq$ 100 mmHg) at Visit 1.
30. Any of the following CV conditions less than 3 months prior to randomization: acute MI, cerebrovascular accident (stroke), unstable angina, or hospitalization due to congestive heart failure.
31. History of (less than 3 months prior to Visit 1) or planned CV procedure.
32. Heart failure, including New York Heart Association Functional Classification Class IV
33. Have acute or chronic hepatitis, signs and symptoms of any other liver disease other than nonalcoholic fatty liver disease, or any of the following, as determined by the central laboratory during screening:
  - ALT level  $>$ 3.0X the ULN for the reference range
  - ALP level  $>$ 1.5X the ULN for the reference range, or
  - TBL level  $>$ 1.2X the ULN for the reference range (except for cases of known Gilbert’s Syndrome).

*Note:* Participants with nonalcoholic fatty liver disease are eligible to participate in this trial if their ALT level is  $\leq$ 3.0X the ULN for the reference range.

34. Have a calcitonin level (at Visit 1) of:
  - a.  $\geq$ 20 ng/L at Visit 1, if eGFR  $\geq$ 60 mL/min/1.73 m<sup>2</sup>
  - b.  $\geq$ 35 ng/L at Visit 1, if eGFR  $<$ 60 mL/min/1.73 m<sup>2</sup>
35. Have a family or personal history of medullary thyroid carcinoma or multiple endocrine neoplasia syndrome type 2.
36. Have a history of an active or untreated malignancy or are in remission from a clinically significant malignancy (other than basal- or squamous-cell skin cancer, in situ carcinomas of the cervix, or in situ prostate cancer) for less than 5 years.
37. Have any other condition not listed in this section (for example, hypersensitivity or intolerance) that is a contraindication to GLP-1R agonists.
38. Have a history of any other condition (such as known drug or alcohol abuse, diagnosed eating disorder, or other psychiatric disorder) that, in the opinion of the investigator, may preclude the participant from following and completing the protocol.
39. Have history of use of marijuana less than 3 months of V1 and unwillingness to abstain from marijuana use during the trial. Participants should also refrain from use of cannabidiol oil for the duration of the study.
40. Have had a transplanted organ (corneal transplants [keratoplasty] allowed) or awaiting an organ transplant.
41. Requires the use of supplemental oxygen.

### **Prior/Concomitant Therapy**

42. Are receiving or have received within 3 months prior to screening
  - a. chronic ( $>$ 2 weeks or 14 days) systemic glucocorticoid therapy (excluding topical, intraocular, intranasal, intra-articular, or inhaled preparations) or
  - b. have evidence of a significant, active autoimmune abnormality (for example, lupus or rheumatoid arthritis) that has required (within the last 3 months) or

- c. is likely to require, in the opinion of the investigator, concurrent treatment with systemic glucocorticoids (excluding topical, intraocular, intranasal, intra-articular or inhaled preparations) in the next 12 months.
43. Have current or history of (less than 3 months prior to Visit 1) treatment with medications that may cause significant weight gain, including but not limited to: tricyclic antidepressants, atypical antipsychotic and mood stabilizers, for example:
- Imipramine
  - Amitriptyline
  - Mirtazapine
  - Paroxetine
  - Phenelzine
  - Chlorpromazine
  - Thioridazine
  - Clozapine
  - Olanzapine
  - valproic acid and its derivatives
  - lithium

*Note:* Selective serotonin reuptake inhibitors are permitted, except for paroxetine.

44. Have taken, less than 3 months prior to Visit 1, medications (prescribed or over-the-counter) or alternative remedies intended to promote weight loss. Examples include, but are not limited to:
- Saxenda® (liraglutide 3.0 mg)
  - Xenical®/Alli® (orlistat)
  - Meridia® (sibutramine)
  - Acutrim® (phenylpropanolamine)
  - Sanorex® (mazindol)
  - Adipex® (phentermine)
  - BELVIQ® (lorcaserin)
  - Qsymia® (phentermine/topiramate combination)
  - Contrave® (naltrexone/bupropion)
  - Pramlintide
  - Zonisamide
  - Topiramate
  - Wegovy®

*Note:* Use of metformin or any other glucose-lowering medication, whether prescribed for polycystic ovary syndrome or diabetes prevention is not permitted.

45. Use of stimulants less than 3 months prior to Visit 1 (for example, modafinil, armodafinil, solriamfetol, pitolisant, amphetamine, dextroamphetamine, dexmethylphenidate, methylphenidate, and lisdexamfetamine).
46. Use of hypnotics, mirtazapine, opioids, trazodone less than 3 months prior to Visit 1.
47. Use of GLP-1 RA less than 3 months prior to Visit 1.
48. Use of any over-the-counter or prescription medications that could affect the evaluation of excessive sleepiness, per investigator discretion.



49. Use of any over-the-counter or prescription medications that could affect the evaluation of excessive sleepiness, per investigator discretion.
50. Unwillingness to discontinue over-the-counter (herbal or supplemental) medication that, in the opinion of the investigator, can interfere with the study.

### **Prior/Concurrent Clinical Study Experience**

51. Are currently enrolled in any other clinical study involving an investigational product or any other type of medical research judged not to be scientifically or medically compatible with this study.
52. Previously randomly assigned to study intervention in this study or any other study investigating tirzepatide.
53. Have participated, within the last 30 days (4 months for studies conducted in Japan; 3 months for studies conducted in the United Kingdom), in a clinical trial involving a study intervention. If the previous study intervention is scientifically or medically incompatible with this study and has a long half-life, 3 months or 5 half-lives (whichever is longer) should have passed prior to screening (participation in observational studies may be permitted upon review of the observational study protocol and approval by the sponsor).
54. Are Lilly employees or employees of third-party organizations involved with the study.
55. Are investigator site personnel directly affiliated with this study and/or their immediate families. Immediate family is defined as a spouse, parent, child, or sibling, whether biological or legally adopted.
56. Have any medical condition that, in the opinion of the investigator, would be a contraindication to participation in the trial

### **Additional Criteria for Study 1**

#### **Inclusion Criteria**

Participants who are unable or unwilling to use PAP therapy. Participants must not have used PAP for at least 4 weeks prior to Visit 1.

#### **Exclusion Criteria**

No additional exclusion criteria apply.

### **Additional Criteria for Study 2**

#### **Inclusion Criteria**

Have been on PAP therapy for at least 3 consecutive months prior to Visit 1 and plan to continue PAP therapy during the study.

#### **Exclusion Criteria**

Have personal or job-related responsibilities, or in the opinion of the investigator have any situation, that would make it unsafe to stop PAP therapy for 7 days prior to PSG testing during the course of the study.

Are unwilling to stop PAP therapy for 7 days prior to PSG testing during the course of the study.

## STATISTICAL ANALYSIS METHODS

### Primary estimands

The primary and each key secondary efficacy analysis will be guided by the “treatment regimen” estimand and the “efficacy” estimand to support global regulatory submissions and publications. The “efficacy” estimand provides an on-treatment assessment of efficacy without confounding the treatment effect from the data collected after treatment discontinuation. It represents on-treatment efficacy. The “treatment regimen” estimand estimates the treatment effect, including the effect of intervention discontinuation to reflect clinical practice. It represents the efficacy irrespective of adherence to study intervention. The “treatment-regimen” estimand will be used as the primary estimand to support a marketing application for the FDA.

### Efficacy estimand

The clinical question of interest for the efficacy estimand is the treatment difference between tirzepatide and placebo after 52 weeks of intervention in treated participants with obesity and OSA, prior to study intervention discontinuation for any reason.

### Efficacy estimand attributes

- *Population:* Adult participants with obesity and OSA who received at least 1 dose of study treatment.
- *Treatment condition:* On randomized treatment.
- *Endpoints:* The primary and key secondary endpoints will be studied. Further details on the endpoints can be found in the Objectives and Endpoints table (Section 1.1 of the Statistical Analysis Plan).
- *Population level summary:* The difference in mean change from baseline to 52 weeks will be used for continuous endpoints; the difference in proportion (absolute or relative, as appropriate) will be used for dichotomous endpoints. The population level summary will be conducted using the EAS described in Section 3 of the Statistical Analysis Plan.
- *Handling of intercurrent events:* The intercurrent events of treatment discontinuation and use of PAP therapy for participants in ISA1 is addressed by the hypothetical strategy. The potential outcome of interest is the response in the efficacy measurement if participants would remain on their randomly assigned treatment for 52 weeks and would not initiate PAP therapy during the study.
- *Rationale:* The efficacy estimand provides an on-treatment assessment without confounding the treatment effect from off-treatment data.

### Treatment-regimen estimand

The clinical question of interest for the treatment-regimen estimand is the treatment difference between tirzepatide and placebo after 52 weeks of intervention in treated participants with obesity and OSA, regardless of intervention discontinuation for any reason.

### Treatment-regimen estimand attributes

- *Population:* Adult participants with obesity and OSA who received at least 1 dose of study treatment.
- *Treatment condition:* On- or off-randomized-treatment.
- *Endpoints:* The primary and key secondary endpoints will be studied. Further details on the endpoints are in the Objectives and Endpoints table (Section 1.1 of the Statistical Analysis Plan).

- *Population level summary:* The difference in mean change from baseline to 52 weeks will be used for continuous endpoints and the difference in proportion (absolute or relative, as appropriate) will be used for dichotomous endpoints. The population level summary will be conducted using the FAS described in Section 3 of the Statistical Analysis Plan.
- *Handling of intercurrent events:* No intercurrent events since treatment adherence and the initiation of PAP therapy are part of the treatment condition. Methods to handle missing data are described in detail in Section 4.1.2 of the Statistical Analysis Plan.
- *Rationale:* The treatment-regimen estimand estimates treatment effect, including the effect of intervention discontinuation to reflect clinical practice. It is used for submission and registration purpose with regulatory agencies.

Efficacy and treatment-regimen estimands will be evaluated for key secondary objectives similarly to the primary objectives.

### Handling of Dropouts or Missing Data

For analyses aligned to the “efficacy” estimand, missing data will be considered missing at random and hence no explicit imputation will be performed for continuous endpoints. For categorical endpoints, the corresponding continuous variable associated with the missing categorical data will be considered missing at random, and multiple imputation assuming the data to be missing at random will be performed.

For the primary and key secondary efficacy endpoint analyses aligned to the treatment-regimen estimand and subject to Type 1 error rate control, missing data will be imputed based on the reason for the missing values, as described in Table GPIF.4.2.

**Table GPIF.4.2. Imputation Approaches to Handle Missing/Invalid Data for Treatment-Regimen Estimand**

Missing/Invalid Data	Strategy to Handle Missing/Invalid Data	Assumptions for Missing Values	Methods to Handle Missing Values
Data missing at baseline, invalid data collected or missing data after treatment DC due to the COVID-19 pandemic (after other reasons for missing data are ruled out), technical issues (that is, sensor error on PSG) leading to invalid measurements ascertained while on treatment, or missing data after study DC due to inadvertent enrollment.	Hypothetical	MAR	Multiple imputation assuming MAR
Missing data due to any other reason (for example, study DC due to any reason other than COVID-19 or inadvertent enrollment).	Treatment policy	MNAR	Retrieved dropout imputation <sup>a</sup> . If there are not enough retrieved dropouts to provide a reliable imputation model, placebo-based multiple imputation will be used.

Abbreviations: COVID-19 = coronavirus disease-2019; DC = discontinuation; MAR = missing at random; MNAR = missing not at random; PSG = polysomnography; SBP = systolic blood pressure.

<sup>a</sup> Retrieved dropout imputation utilizes observed data from participants in the same treatment group who had outcome measures at Week 52 (or Week 48 for SBP) after early DC of study drug to impute the missing value.

Multiple imputation was implemented using a Bayesian Markov chain Monte Carlo (MCMC) method with Jeffrey’s non informative prior assuming a multivariate Normal distribution for the complete data. For each variable, the model includes AHI severity, Sex, Region, and value of the variable at all scheduled visits including baseline. 100 multiple imputed datasets were generated, and the statistical inference over multiple imputations was guided by the method proposed by Rubin (1987).<sup>1</sup>

**Table GPIF.0.1. Description of Analysis Population**

<b>Analysis Population</b>	<b>Description</b>
Entered	All participants who sign informed consent.
Randomized	All participants who are randomly assigned a study treatment (double-blind).
Modified intent-to-treat (mITT)	All randomized participants who are exposed to at least 1 dose of study intervention.

**Table GPIF.0.2. Description of Analysis Data Point Sets**

<b>Analysis Set</b>	<b>Description</b>
Full analysis set (FAS)	Data obtained during treatment period of set of participants from the mITT population, regardless of adherence to study intervention.  For AHI related endpoints and PROs associated with PSG visits (ESS, FOSQ, PROMIS, PGIS, PGIC, SF-36v2 acute form, and EQ-5D-5L), data obtained outside the anticipated window through end of study are included as part of treatment period.
Efficacy analysis set (EAS)	Data obtained during treatment period of set of participants from the mITT population, excluding data after discontinuation of study intervention (last dose + 7 days) and for ISA1, excluding data after initiating PAP therapy.  For AHI related endpoints and PROs associated with PSG visits (ESS, FOSQ, PROMIS, PGIS, PGIC, SF-36v2 acute form, and EQ-5D-5L), data obtained outside the anticipated window through end of study are included as part of treatment period.
Safety analysis set (SS)	Data obtained during treatment and safety follow-up period of set of participants from the mITT population, regardless of adherence to study intervention.

Abbreviations: EAS = efficacy analysis set; FAS = full analysis set; ISA = intervention-specific appendix; mITT = modified intent-to-treat; PAP = positive airway pressure; SS = safety analysis set.

**Primary Endpoint Analysis**

The primary objective of this study is to test the hypothesis that tirzepatide at the MTD (10 mg or 15 mg) is superior to placebo for participants with moderate-to-severe OSA and obesity on the mean AHI reduction from baseline to Week 52. The primary and key secondary efficacy analyses will be guided by 2 estimands,

the “treatment regimen” estimand and the “efficacy” estimand to support global regulatory submissions and publications.

### **Analysis Related to the Efficacy Estimand**

The primary analysis guided by the “efficacy” estimand will be conducted using the EAS. This analysis will be based on the contrast between tirzepatide at the MTD (10 mg or 15 mg) and placebo at Week 52 (Visit 11) from the MMRM analysis of mean change from baseline in AHI. All the longitudinal observations at each scheduled postbaseline visit will be included in the analysis. REML will be used to obtain model parameter estimates and the Kenward-Roger option will be used to estimate the denominator degrees of freedom. The response variable of the MMRM will be the change in AHI from baseline values obtained at each scheduled postbaseline AHI measurement.

The model will include the fixed class effects of treatment, strata (geographic region [US/OUS] and gender), visit, and treatment-by-visit interaction, as well as the continuous, fixed covariate of baseline AHI. An unstructured covariance structure will be used to model the within-participant errors. Significance tests will be based on least-squares means and Type III tests. If this analysis fails to converge, the following covariance structures will be tested in order until convergence is achieved:

- Toeplitz with heterogeneity
- autoregressive with heterogeneity
- compound symmetry with heterogeneous variances
- Toeplitz
- autoregressive, and
- compound symmetry without heterogeneous variances.

Analysis aligned to each estimand will be evaluated at the full significance level of 0.05.

### **Analysis Related to the Treatment-Regimen Estimand**

For the primary analysis guided by the “treatment regimen” estimand, the analysis will be conducted using the FAS. Missing values will be imputed based on the strategy to handle intercurrent events described in Section 4.1.2 of the Statistical Analysis Plan. After imputation, the primary efficacy comparison will be based on the contrast between tirzepatide at the MTD (10 mg or 15 mg) and placebo from the ANCOVA analysis of mean change from baseline to Week 52 in AHI using FAS. The ANCOVA model will include treatment and strata (geographic region [US/OUS] and gender) as fixed effects and baseline AHI as a fixed covariate. Statistical inference over multiple imputed data sets will be guided by Rubin (1987).<sup>1</sup>

### **Key Secondary Endpoint Analysis**

#### **Main Analytical Approaches**

Analysis of percent change in AHI, percent change from baseline to Week 52 in body weight, change from baseline to Week 52 in log of high-sensitivity C-reactive protein, change from baseline to Week 48 in SBP, change from baseline to Week 52 in PROMIS-SRI, change from baseline to Week 52 in PROMIS-SD, and change from baseline to Week 52 in log of hypoxic burden will be conducted in a manner similar to the primary efficacy analyses using an ANCOVA model with treatment, strata (geographic region [US/OUS], AHI stratum [not severe (AHI <30), severe (AHI ≥30)], and gender), and baseline of the corresponding variable as a covariate for the treatment-regimen estimand. If the hypoxic burden is reported to be 0, log (1) will be used in place of the log of hypoxic burden. The analysis method utilizing data from both ISAs for

change from baseline to Week 52 in PROMIS-SRI and PROMIS-SD is described in the integrated efficacy analysis plan.

For the efficacy estimand, the MMRM analyses will be conducted as described in Section 4.1 of the Statistical Analysis Plan. For both estimands, analysis of percent change in AHI will adjust for the continuous, fixed baseline value of AHI instead of the baseline AHI stratum (not severe, severe). Comparisons at the 52-week visit between the treatments relative to the proportion of participants achieving  $\geq 50\%$  AHI reduction and  $AHI < 5$  or ( $AHI$  5 through 14 and  $ESS \leq 10$ ) will be conducted using logistic regression analysis including the following terms as a covariate: treatment geographic region (US/OUS) baseline AHI, and gender. Analysis aligned to each estimand will be evaluated at the full significance level of 0.05 contingent on reaching statistical significance of the primary objective.

### **PROMIS-SRI and PROMIS-SD endpoints**

Analysis for change in PROMIS Short Form Sleep-Related Impairment 8a (PROMIS-SRI) and PROMIS Short Form Sleep Disturbance 8b (PROMIS-SD) is specified in the Integrated Efficacy Analysis Plan (IEAP) to be tested subject to the submission wise error rate control strategy<sup>2,3</sup> by conducting a pooled analysis across the 2 studies. The pooled analysis across the two studies with PROMIS key secondary endpoints aims to maintain rigorous control of the family wise error rate (FWER) within each study for the primary endpoints while facilitating efficient statistical evaluation of secondary endpoints.

Study 1 and Study 2 were different in presence of the background PAP therapy. The potential issue of PAP confounding the PRO outcomes in Study 2 was minimized by 7 days PAP wash-out period. The washout period, consistently implemented at baseline, week 20, and week 52 assessment, safeguards that the changes detected represent the effect of the therapeutic intervention of tirzepatide or placebo with minimal to no PAP confounding. Therefore, the populations of the two studies were seen as homogenous and suitable for pooling in the context of the endpoints analyzed.

## **Integrated Efficacy Analysis Plan (IEAP)**

### **Objective and Study Designs**

The objective of the integrated efficacy analysis plan (IEAP) is to specify the planned integrated efficacy analysis of tirzepatide (TZP) for the indication of obstructive sleep apnea (OSA). This document describes the integrated analysis intended to be included in Summary of Clinical Efficacy (SCE)/Integrated Summary of Efficacy as part of the Common Technical Document for submission.

This document includes

- strategies for combining studies
- analyses to be performed, and
- statistical methods to be used

Lilly does not plan to pool any Intervention-Specific Appendices (ISAs) for the primary efficacy assessment and closely related key secondary endpoints because the individual ISAs are adequately powered to assess the primary efficacy objective. Lilly is planning an integrated analysis for selected efficacy parameters potentially not adequately powered within each ISA using pooled data from Study GPI1 (ISA1) and Study GPI2 (ISA2), while strongly controlling the submission-wise Type 1 error rate (SWER), that is, the probability to make a false claim of success for an endpoint at the submission level.<sup>3</sup>

The side-by-side presentation of primary efficacy results across individual ISAs are planned to be included in the SCE. This document will not detail the methods to be used for analyses in individual ISAs. Those analyses will be described in the statistical analysis plan (SAP) for the study.

### **Overview of Individual Studies Evaluating Clinical Efficacy**

Phase 3 study/ISAs for patients with OSA that will be included in the SCE are briefly described below.

Study I8F-MC-GPIF (GPIF) is a multicenter, randomized, parallel-arm, double-blind, placebo- controlled Phase 3 study to evaluate the efficacy and safety of TZP at the maximum tolerated dose (MTD) (10 mg or 15 mg) once weekly (QW) versus placebo in participants who have obesity and moderate-to-severe OSA.

This basket-type master protocol investigates 2 participant populations, described in 2 ISAs:

- Study GPI1 includes participants who are unwilling or are unable to use positive airway pressure (PAP) therapy.
- Study GPI2 includes participants who are on PAP therapy for at least 3 months at time of screening and plan to continue PAP therapy during the study.

Participants are to be assigned to whichever ISA they qualify for. Participants will then be randomly assigned to

- TZP at the maximum tolerated dose (10 mg or 15 mg) subcutaneously (SC) once weekly (QW), or
- placebo.

The expected total duration of study participation for each participant, including screening and the posttreatment follow-up periods, is 60 weeks across the following study periods:

- Screening: 4 weeks
- Treatment: 52 weeks
- Posttreatment follow-up: 4 weeks

The maximum duration of treatment is 52 weeks.

Efficacy measures presented in the SCE for the primary efficacy endpoint and closely related key secondary endpoints will be derived from individual ISAs for Study GPIF, without patient-level integration of data from individual ISAs. The details of these analyses are provided in the respective Study GPIF SAP.

The primary and key secondary endpoints from individual ISAs for the Study GPIF are provided in the SAP.

### Analysis Sets

For individual ISAs, the analysis sets are specified in the Study GPIF SAP. Pooled analyses will be conducted using the same analysis set as the set used in the individual ISAs specified in the Study GPIF SAP.

## A Priori Statistical Methods

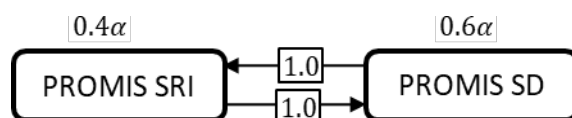
### General Considerations

Statistical analysis will be the responsibility of Eli Lilly and Company or their appointed contract research organization.

Treatment group will be based on assignment at randomization. The multiplicity adjustment within each ISA is described in the SAP for Study GPIF. The sample size of individual ISAs was guided by the anticipated efficacy relative to the primary endpoint. Due to potentially inadequate power within each ISA, a pooled analysis will be conducted for the following key secondary endpoints:

- change from baseline in Patient-Reported Outcomes Measurement Information System (PROMIS) Short Form Sleep-Related Impairment 8a (PROMIS-SRI) score at Week 52
- change from baseline in PROMIS Short Form Sleep Disturbance 8b (PROMIS-SD) Score at Week 52

The pooled analyses described above are subject to the SWER<sup>2,3</sup> of  $\alpha = 0.04875 = 2 \times (0.025 - 0.025^2)$  if the primary endpoint is achieved under the family wise error rate (FWER) control in both ISAs, and  $\alpha = 0.05$  if all endpoints subject to the FWER control of endpoint is achieved in both ISAs. Figure GPIF 5.1 illustrates the graphical testing strategy for evaluating the key secondary endpoints subject to the SWER control.



Abbreviations: PROMIS = Patient-Reported Outcomes Measurement Information System; SD = Sleep Disturbance; SRI = Sleep-Related Impairment.

Figure GPIF 5.1 Graphical testing scheme for pooled analysis.

### Handling of Dropouts or Missing Data

The details are described in the SAP for Study GPIF. The retrieved dropout imputation will be performed within the pooled treatment group across ISAs.

### Comparison of Results of Individual Studies



There is no statistical comparison of results of individual ISAs.

### **Disposition**

Patient disposition will be summarized by treatment group for each ISA. The details are specified in the Study GPIF SAP, which is aligned with Program Safety Analysis Plan (PSAP).

### **Demographics and Baseline Characteristics**

Demographics and other baseline characteristics will be summarized by treatment group for each ISA. The details are specified in the Study GPIF SAP, which is aligned with PSAP.

### **Primary and Important Secondary Endpoints**

The statistical analysis method for the primary and key secondary efficacy endpoints are specified in the Study GPIF SAP.

The primary and key secondary efficacy results from each individual ISAs will be presented side-by-side to support the SCE. Selected key secondary endpoints to conduct the pooled analysis specified in Section 5.1 of the Integrated Efficacy Analysis Plan will also be presented in the SCE.

### **Pooled Analysis of Data from More than One Study**

All analyses will be performed and align to the “treatment regimen” estimand as described in GPIF SAP. Analysis methods and imputation approaches for handling missing/invalid data are in line with the individual ISAs specified in the Study GPIF SAP. Following imputation, all endpoints will be analyzed from the analysis of covariance model with treatment, ISA [ISA1/ISA2], geographic region [US/OUS], Apnea-Hypopnea Index (AHI) stratum (not severe [AHI <30]/severe [AHI ≥30]), and gender as fixed effects, with baseline as a covariate, using the pooled full analysis set (FAS) in each ISA. Statistical inference over multiple imputed data sets will be guided by Rubin’s rule (1987).<sup>1</sup>

### **Comparison of Results in Subpopulations**

The subgroup analyses at the ISA level will be used to support the SCE. The details are specified in the Study GPIF SAP.

### **Analysis of Clinical Information Relevant to Dosing Recommendation**

The relationship between efficacy outcomes and pharmacokinetics (PK) may be presented. The details of these analyses are included in the population PK/pharmacodynamic analysis plan.

### **Time Course of Effect, Persistence of Effect, and/or Tolerance, Distribution of Responses**

Therapeutic effects of a treatment can decline over time because of tolerability issues (patients who experience adverse events and refuse treatment), from the development of drug resistance or tolerance, or because the disease tends to resolve spontaneously (FDA 2015).

Studies GPI1 and GPI2 have primary objectives at 1 year of treatment. These studies will provide evidence of the long-term sustainability of effect for TZP.

## **Rationale for Change of Key Secondary Endpoints**

All final pre-specified primary and key secondary endpoints are found in the Statistical Analysis Plan (SAP) v4 dated April 5, 2024. Key secondary patient-reported outcomes (PRO) endpoints were changed per regulatory body recommendation from the hierarchical combination of change in FOSQ scores to change in PROMIS-SRI and PROMIS-SD. Due to timing of regulatory advice and because the change did not impact how clinical trial investigators conducted the studies, the change in key secondary PRO endpoints was captured in the final statistical analysis plan and not in the protocol. Given the global nature of the studies, the local IRBs would not have been able to complete the approval review process for a protocol amendment prior to the end of the studies. The database lock occurred after the final SAP was submitted with no possibility of unblinding prior to data lock. The sponsor, participants, and investigators remained blinded to the data while these decisions were made.

## **Pre-specified Endpoints**

Primary and key secondary endpoints controlled for Type 1 error are reported in the manuscript.

## Master Protocol Endpoints

Objective	Endpoints
<b>Primary</b>	
To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for decrease in AHI.	Change in AHI from baseline to Week 52.
<b>Key Secondary (controlled for Type 1 error)</b>	
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Percent change in AHI</li> <li>• Clinically meaningful change in AHI</li> <li>• Achieving OSA remission or mild non-symptomatic OSA</li> <li>• Change in body weight</li> <li>• Change in inflammatory status</li> <li>• Hypoxic burden</li> <li>• Change in PROs</li> </ul> <p>• Change in SBP</p>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Percent change in AHI</li> <li>• Percent of participants with <math>\geq 50\%</math> AHI reduction</li> <li>• Percent of participants with               <ul style="list-style-type: none"> <li>○ AHI <math>&lt; 5</math> or</li> <li>○ AHI 5-14 with ESS <math>\leq 10</math></li> </ul> </li> <li>• Percent change in body weight</li> <li>• Change in hsCRP concentration</li> <li>• Change in SASHB (% min/hour)</li> <li>• Change in<sup>a</sup>:               <ul style="list-style-type: none"> <li>○ PROMIS Sleep-related impairment short form 8a</li> <li>○ PROMIS Sleep disturbance short form 8b</li> </ul> </li> </ul> <p>From baseline to Week 48<sup>b</sup></p> <ul style="list-style-type: none"> <li>• Change in SBP</li> </ul>
<b>Other Secondary</b>	
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in excessive daytime sleepiness</li> <li>• Change in patient-reported functional status as assessed by FOSQ (30 items)</li> </ul> <p>• Change in body weight</p> <p>• Change in lipid parameters</p>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in ESS score</li> <li>• Change in FOSQ-10 score</li> <li>• Change in FOSQ (30 items) Score</li> <li>• Change in all FOSQ domain scores, specifically               <ul style="list-style-type: none"> <li>○ General Productivity</li> <li>○ Activity level</li> <li>○ Vigilance</li> <li>○ Social outcomes</li> <li>○ Intimate and sexual relationships</li> </ul> </li> <li>• Percent of participants who achieve               <ul style="list-style-type: none"> <li>○ <math>\geq 10\%</math> body weight reduction</li> <li>○ <math>\geq 15\%</math> body weight reduction</li> <li>○ <math>\geq 20\%</math> body weight reduction</li> </ul> </li> <li>• Change in               <ul style="list-style-type: none"> <li>○ HDL-cholesterol</li> <li>○ non-HDL-cholesterol</li> <li>○ triglycerides</li> </ul> </li> </ul>

Objective	Endpoints
<ul style="list-style-type: none"> <li>• A hierarchical assessment of PRO change</li> <li>• Change in supportive secondary PROs</li> <li>• Insulin</li> <li>• Change in DBP</li> </ul>	<ul style="list-style-type: none"> <li>• A hierarchical combination of the following: <ul style="list-style-type: none"> <li>○ Change in PROMIS Sleep-related impairment short form 8a</li> <li>○ Change in PROMIS Sleep disturbance short form 8b</li> </ul> </li> <li>• Change in: <ul style="list-style-type: none"> <li>○ SF-36v2 acute form domain and summary scores</li> </ul> </li> <li>• Percent of participants with improved categorical shift in: <ul style="list-style-type: none"> <li>○ PGIS-OSA Sleepiness</li> <li>○ PGIS-OSA Fatigue</li> <li>○ PGIS-OSA Snoring</li> </ul> </li> <li>• Proportion of participants achieving clinically meaningful within-patient change in: <ul style="list-style-type: none"> <li>○ PROMIS Sleep-related impairment</li> <li>○ PROMIS Sleep disturbance</li> </ul> </li> <li>• Change in fasting insulin From baseline to Week 48<sup>a</sup></li> <li>• Change in DBP</li> </ul>
<b>Exploratory</b>	
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in exploratory PROs</li> </ul> <p>• To evaluate the effect of tirzepatide on sleep parameters as measured by Actigraphy (AX6)</p>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in <ul style="list-style-type: none"> <li>○ EQ-5D-5L utility index</li> <li>○ EQ-VAS scores</li> </ul> </li> <li>• Percent of participants with improved categorical shift in: <ul style="list-style-type: none"> <li>○ PGIC-OSA Sleepiness</li> <li>○ PGIC-OSA Fatigue</li> <li>○ PGIC-OSA Sleep quality</li> <li>○ PGIC-OSA Snoring</li> </ul> </li> <li>• Change from baseline to endpoint assessment in <ul style="list-style-type: none"> <li>○ Daytime sleep duration</li> <li>○ Daily step counts</li> <li>○ Average acceleration</li> </ul> </li> </ul>

Abbreviations: AHI = Apnea-Hypopnea Index; AX6 = Axivity 6; BP = blood pressure; DBP = diastolic blood pressure; ESS = Epworth Sleepiness Scale; EQ-5D-5L = EuroQol-5 Dimension-5-Level; EQ-VAS = EuroQol Visual Analogue Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; HDL = high-density lipoprotein; hsCRP = high-sensitivity C-reactive protein; MTD = maximum tolerated dose; OSA = obstructive sleep apnea; PAP = positive airway pressure; PGIC-OSA = Patient Global Impression of Change – Obstructive Sleep Apnea; PGIS-OSA = Patient Global Impression of Status – Obstructive Sleep Apnea; PRO = patient-reported outcome; PROMIS = Patient-Reported Outcomes Measurement Information System; QW = once weekly; SASHB = sleep apnea-specific hypoxic burden; SBP = systolic blood pressure; SF-36v2 = Short-Form 36 version 2.

<sup>a</sup> Subject to submission wide type 1 error rate control (Vandemeulebroecke et al. 2024).

<sup>b</sup> BP will be assessed at Week 48 because PAP withdrawal at Week 52 may confound BP assessment.

## Study 1 Endpoints

Objective	Endpoints
<b>Exploratory</b>	
<ul style="list-style-type: none"><li>To evaluate the effect of tirzepatide on sleep apnea as measured by WatchPAT300</li></ul>	<ul style="list-style-type: none"><li>Change from baseline to Week 52 in PAT-based device determinations of:<ul style="list-style-type: none"><li>pAHI</li><li>SASHB</li></ul></li></ul>

Abbreviations: pAHI = peripheral apnea-hypopnea index; PAT = peripheral arterial tone; SASHB = sleep apnea-specific hypoxic burden.

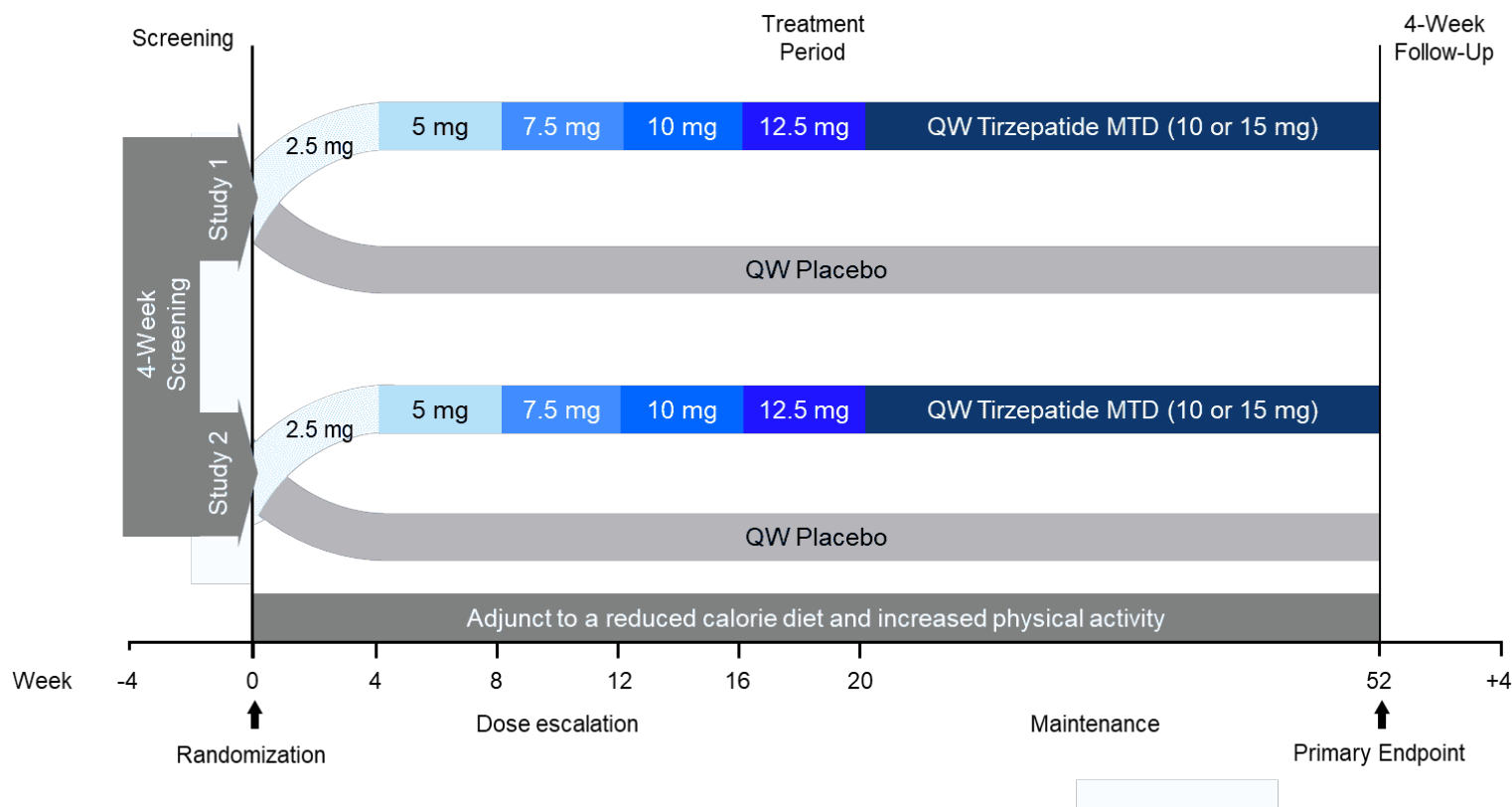
## Study 2 Endpoints

There are no additional objectives and endpoints for Study 2.

## **AUTHOR CONTRIBUTIONS**

All authors contributed to the study design. JB and MCB provided medical oversight during the studies. SC was responsible for the statistical analyses. AM and JB wrote the first draft of the manuscript. AM, JB, and MCB are the guarantors of this work and, as such, take responsibility for the integrity of the data and the accuracy of the data analysis. All authors participated in data interpretation and critical review of the manuscript, had full access to all the data in the study, and approved of this manuscript to be submitted for publication.

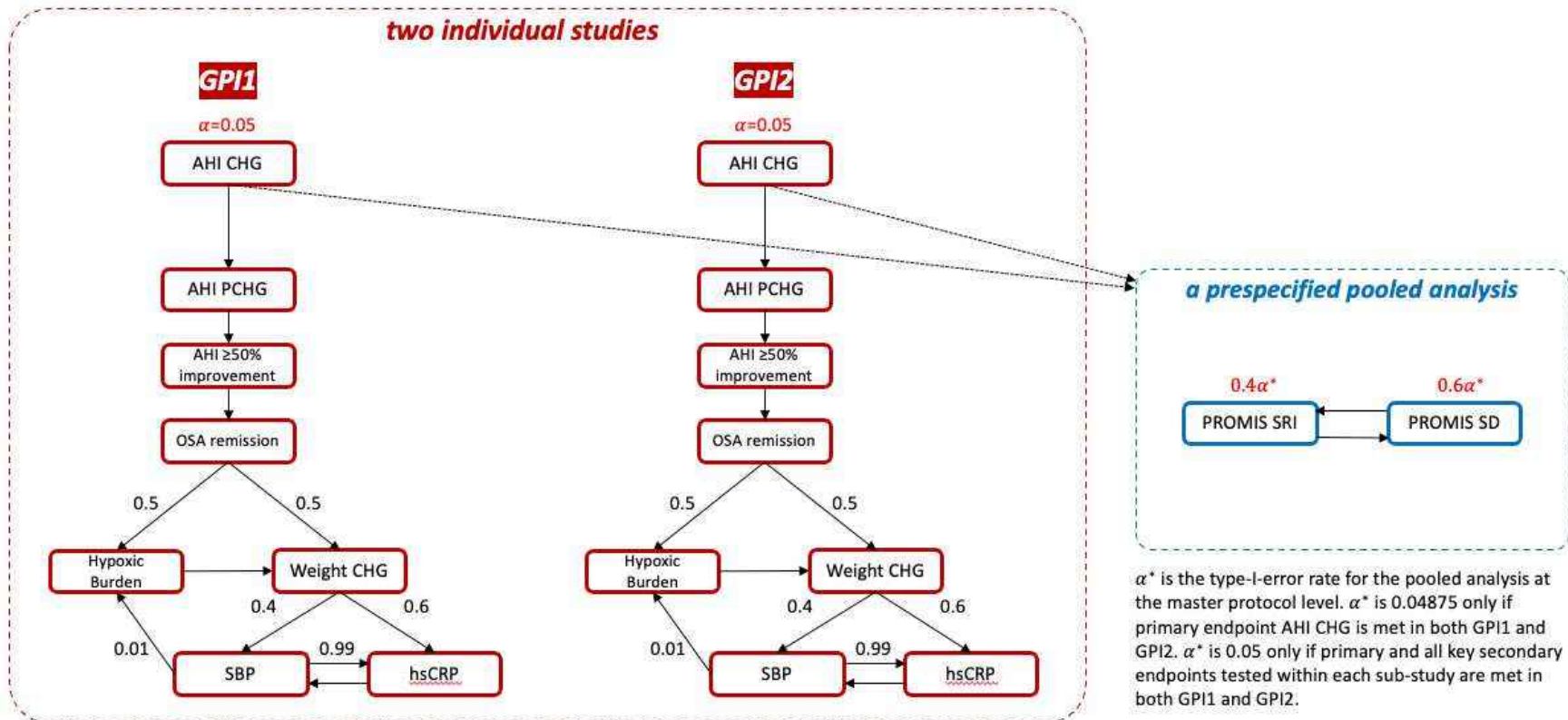
## SUPPLEMENTARY FIGURES



**Figure S1. SURMOUNT-OSA study design.**

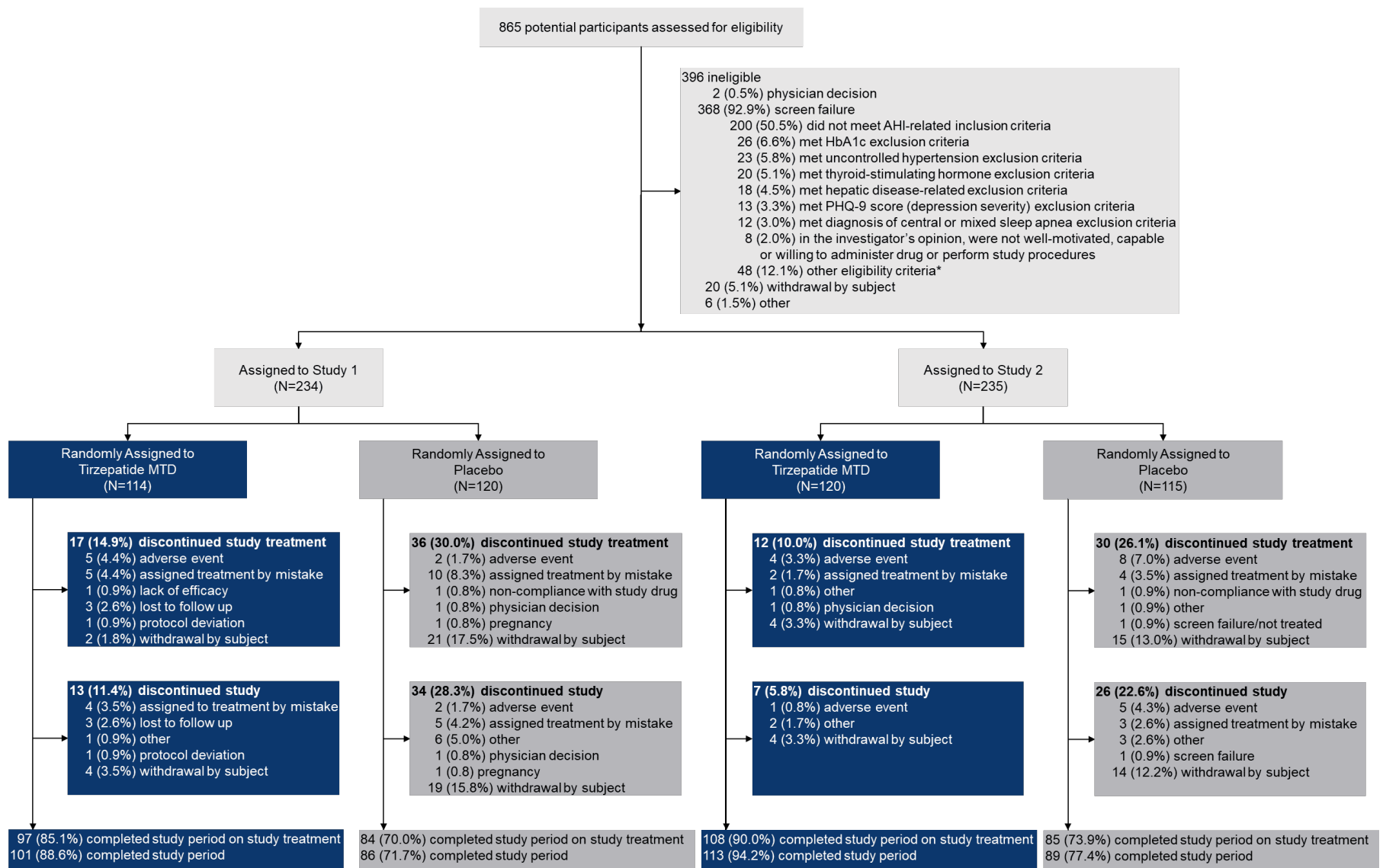
This is a multicenter, randomized, parallel-arm, double-blind, placebo-controlled, Phase 3 study with 52-week treatment duration conducted under a basket-design, which investigated the effects of treatment with once weekly (QW) tirzepatide at the maximum-tolerated dose (MTD) (10 mg or 15 mg), compared with placebo in participants who have moderate-to-severe OSA and obesity. The master protocol supported 2 studies: Study 1 included participants who are unable or unwilling to use PAP therapy and Study 2 included participants who were on PAP therapy for at least 3 months at time of screening and planned to continue PAP therapy during the study. Participants were assigned to the Study which reflected their PAP usage at screening. Participants were then randomly assigned 1:1 to treatment or placebo.





**Figure S2. Pre-Specified Multiplicity Control for SURMOUNT-OSA Master Protocol**

Details of the graphical multiple testing procedure. Abbreviations: AHI = Apnea-Hypopnea Index; CHG = change; GPI1 = Study 1; GPI2 = Study 2; hsCRP = high-sensitivity C-reactive protein; OSA = obstructive sleep apnea; PCHG = percentage change; PROMIS = Patient-Reported Outcomes Measurement Information; SBP = systolic blood pressure; SD = Sleep Disturbance; SRI = Sleep-Related Impairment.



**Figure S3. Participant disposition from randomization to primary end point.**

\*Other represents all eligibility criteria screen failures individually occurring in less than 2% of screened individuals.

## SUPPLEMENTARY TABLES

**Table S1. Patient Geography**

	Study 1			Study 2		
	Tirzepatide N=114	Placebo N=120	Total N=234	Tirzepatide N=120	Placebo N=115	Total N=235
<b>Country/Region</b>						
Australia	1 (0.9)	5 (4.2)	6 (2.6)	10 (8.3)	9 (7.8)	19 (8.1)
Brazil	24 (21.1)	25 (20.8)	49 (20.9)	18 (15.0)	17 (14.8)	35 (14.9)
China	14 (12.3)	14 (11.7)	28 (12.0)	4 (3.3)	5 (4.3)	9 (3.8)
Czech Republic	5 (4.4)	6 (5.0)	11 (4.7)	6 (5.0)	6 (5.2)	12 (5.1)
Germany	7 (6.1)	7 (5.8)	14 (6.0)	16 (13.3)	16 (13.9)	32 (13.6)
Japan	3 (2.6)	4 (3.3)	7 (3.0)	7 (5.8)	6 (5.2)	13 (5.5)
Mexico	20 (17.5)	19 (15.8)	39 (16.7)	14 (11.7)	16 (13.9)	30 (12.8)
Taiwan	5 (4.4)	4 (3.3)	9 (3.8)	4 (3.3)	3 (2.6)	7 (3.0)
United States	35 (30.7)	36 (30.0)	71 (30.3)	41 (34.2)	37 (32.2)	78 (33.2)

Data are n (%) and include all randomized participants.

**Table S2. Representativeness of Study**

<b>Category</b>	<b>Details</b>
<b>Disease under investigation</b>	Obstructive Sleep Apnea (OSA)
Special considerations related to:	
<b>Sex and gender</b>	The prevalence of OSA is greater in men than women (male-to-female ratio between 3:1 to 5:1). In the United States, it is estimated that 13% of men and 6% of women have moderate-to-severe OSA.
<b>Age</b>	The prevalence of OSA increases with age. In the United States, the prevalence of OSA is highest among adults aged 50 to 70.
<b>Race or ethnic group</b>	African Americans, Hispanics, and Native Americans have been found to have a higher rate of OSA.
<b>Geography</b>	OSA is a global health care challenge. Countries with the highest OSA prevalence estimates include China, United States, India, Brazil, Pakistan Russia, Nigeria, Germany, France and Japan.
<b>Other considerations</b>	Prevalence of OSA is highly correlated to the prevalence of obesity.
<b>Overall representativeness of these studies</b>	The SURMOUNT-OSA studies were conducted in adults $\geq 18$ years of age in 9 countries: Australia, Brazil, China, Czech Republic, Germany, Japan, Mexico, Taiwan, and the United States. The United States, which has a high prevalence of OSA and obesity, represented 32% of the participant population. Male enrollment was limited to 70% to ensure adequate representation of females. Men represented 69.7% of the study population. The mean age was 49.7 years, with 51.9% representing those aged 50 years or older. This is in line with the higher prevalence of OSA in men and the 50-to-70-year age group. All demographic measures were self-reported. In the SURMOUNT-OSA studies, the US study population was representative of the prevailing demographics in the United States. Overall, the trial was representative of the global and US prevalence of OSA by age and race or ethnic group.

**Table S3. Summary of Missing Data for Primary and Key Secondary Endpoints**

Variables	Study 1		Study 2	
	Baseline	Endpoint	Baseline	Endpoint
<b>Treatments</b>				
AHI (Change, Percentage Change, >= 50% Improvement)				
Placebo	1 (0.8)	34 (28.3)	0 (0.0)	27 (23.7)
Tirzepatide	0 (0.0)	14 (12.3)	0 (0.0)	9 (7.6)
AHI < 5 or AHI 5-14 with ESS <= 10				
Placebo	--	35 (29.2)	--	27 (23.7)
Tirzepatide	--	19 (16.7)	--	11 (9.2)
Body Weight				
Placebo	0 (0.0)	33 (27.5)	0 (0.0)	24 (21.1)
Tirzepatide	0 (0.0)	13 (11.4)	0 (0.0)	6 (5.0)
hsCRP				
Placebo	4 (3.3)	35 (29.2)	7 (6.1)	25 (21.9)
Tirzepatide	7 (6.1)	13 (11.4)	8 (6.7)	10 (8.4)
Hypoxic Burden				
Placebo	1 (0.8)	34 (28.3)	0 (0.0)	27 (23.7)
Tirzepatide	0 (0.0)	14 (12.3)	0 (0.0)	9 (7.6)
PROMIS Sleep Related Impairment				
Placebo	17 (14.2)	41 (34.2)	19 (16.7)	36 (31.6)
Tirzepatide	20 (17.5)	24 (21.1)	17 (14.3)	26 (21.8)
PROMIS Sleep Disturbance				
Placebo	16 (13.3)	41 (34.2)	16 (14.0)	35 (30.7)
Tirzepatide	20 (17.5)	24 (21.1)	14 (11.8)	26 (21.8)
Systolic Blood Pressure				
Placebo	0 (0.0)	31 (25.8)	0 (0.0)	23 (20.2)
Tirzepatide	0 (0.0)	14 (12.3)	0 (0.0)	7 (5.9)

Data are n (%). Endpoint is Week 52 for all variables except for systolic blood pressure. The endpoint for systolic blood pressure is Week 48. AHI = apnea-hypopnea index; ESS = Epworth Sleepiness Scale; hsCRP = high sensitivity C-reactive protein.

**Table S4. PROMIS Outcomes by Study**

	Study 1			Study 2		
	Tirzepatide N=114	Placebo N=120	Estimated Treatment Difference (95% CI)	Tirzepatide N=120	Placebo N=115	Estimated Treatment (95% CI); p value
Change in PROMIS Sleep- Related Impairment T Score	-6.6 (-8.2 to -4.9)	-3.1 (-4.7 to -1.6)	-3.4 (-5.7 to -1.2)	-8.2 (-10.0 to -6.3)	-3.9 (-5.9 to -1.9)	-4.3 (-7.0 to -1.6)
Change in PROMIS Sleep Disturbance T Score	-4.5 (-5.8 to -3.1)	-2.4 (-3.8 to -1.1)	-2.0 (-4.0 to -0.1)	-7.0 (-8.6 to -5.4)	-3.1 (-4.8 to -1.4)	-3.9 (-6.2 to -1.6)

Treatment-regimen estimand least-squares means (95% confidence interval) are presented. The PROMIS Short Form Sleep-related Impairment 8a consists of 8 items each rated on a 5-point scale ranging from “not at all” to “very much.” Items have a recall period of “in the past 7 days.” Individual item scores were totaled to obtain a raw score which was then converted to a T-score (using response pattern scoring) with a mean of 50 and a SD of 10. Higher scores indicate more sleep-related impairment. The PROMIS Short Form Sleep Disturbance 8b consists of 8 items each rated on a 5-point scale ranging from “not at all” to “very much,” “never” to “always,” or “very poor” to “very good.” Items have a recall period of “in the past 7 days.” Individual item scores were totaled to obtain a raw score which was then converted to a T-score (using response pattern scoring) with a mean of 50 and a SD of 10. Higher scores indicate more sleep disturbance.

## REFERENCES

1. Rubin DB. Multiple Imputation for Nonresponse in Surveys: John Wiley & Sons Inc.; 1987.
2. Bretz F, Xi D. Commentary on “Statistics at FDA: Reflections on the Past Six Years”. *Statistics in Biopharmaceutical Research* 2019;11(1):20-25. DOI: 10.1080/19466315.2018.1562369.
3. Vandemeulebroecke M, Häring DA, Hua E, Wei X, Xi D. New strategies for confirmatory testing of secondary hypotheses on combined data from multiple trials. *Clinical Trials* 2024;21(2):171-179. DOI: 10.1177/17407745231214382.

## Data Sharing Statement

Malhotra A, Grunstein RR, Fietze I, et al. Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity. *N Engl J Med*. DOI: 10.1056/NEJMoa2404881.

Question	Authors' Response
Will the data collected for your study be made available to others?	Yes
Would you like to offer context for your decision?	—
Which data?	Complete de-identified patient data set
Additional information about data	Lilly provides access to all individual participant data collected during the trial, after anonymization, with the exception of pharmacokinetic or genetic data.
How or where can the data be obtained?	Access is provided after a proposal has been approved by an independent review committee identified for this purpose and after receipt of a signed data sharing agreement. Data and documents, including the study protocol, statistical analysis plan, clinical study report, blank or annotated case report forms, will be provided in a secure data sharing environment. For details on submitting a request, see the instructions provided at <a href="http://www.vivli.org">www.vivli.org</a> .
When will data availability begin?	Data are available to request 6 months after indication studied has been approved in US & EU and after primary publication acceptance, whichever is later.
When will data availability end?	No expiration date of data requests is currently set once data are made available.
Will any supporting documents be available?	—
Which supporting documents?	—
Additional information about supporting documents	—
How or where can supporting documents be obtained?	—
When will supporting documents availability begin?	—
When will supporting documents availability end?	—
To whom will data be available?	—
For what type of analysis or purpose?	—
By what mechanism?	—
Any other restrictions?	—
Additional information	—

This statement was posted on June 21, 2024, at [NEJM.org](http://NEJM.org).



Terri Weaver

**Discloser Identifier:** 1184233

**Disclosure Purpose:** 24-04881

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
Bayer AG	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consultation on patient reported outcomes for sleep disorders and use of the Functional Outcomes of Sleep Questionnaire <b>Additional Information:</b>		
Eli Lilly and Company	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Steering committee for clinical trial as well as patient reported outcomes <b>Additional Information:</b>		
IDORSIA Pharmaceuticals US, Inc.	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Steering Committee, Alliance for Sleep <b>Additional Information:</b>		

### Intellectual Property

Type	Is Licensed	Interest Held By
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self
<b>Description:</b> Functional Outcomes of Sleep Questionnaire <b>Filing Jurisdiction:</b> <b>Licenses:</b>		
<b>Copyright Number:</b> <b>Copyright Holder:</b> Discloser		
Licensee	Title	Date
Alkermes, Inc.		
<b>Additional Information:</b>		
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self
<b>Description:</b> Functional Outcomes of Sleep Questionnaire <b>Filing Jurisdiction:</b> <b>Licenses:</b>		
<b>Copyright Number:</b> <b>Copyright Holder:</b> Discloser		
Licensee	Title	Date
Philips		
<b>Additional Information:</b>		

Type	Is Licensed	Interest Held By						
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <p><b>Copyright Number:</b>  <b>Copyright Holder:</b> Discloser</p> <table border="1"> <thead> <tr> <th>Licensee</th> <th>Title</th> <th>Date</th> </tr> </thead> <tbody> <tr> <td>Cook Research Incorporated</td> <td></td> <td></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Cook Research Incorporated		
Licensee	Title	Date						
Cook Research Incorporated								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <p><b>Copyright Number:</b>  <b>Copyright Holder:</b> Discloser</p> <table border="1"> <thead> <tr> <th>Licensee</th> <th>Title</th> <th>Date</th> </tr> </thead> <tbody> <tr> <td>Resmed Corp</td> <td></td> <td></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Resmed Corp		
Licensee	Title	Date						
Resmed Corp								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <p><b>Copyright Number:</b>  <b>Copyright Holder:</b> Discloser</p> <table border="1"> <thead> <tr> <th>Licensee</th> <th>Title</th> <th>Date</th> </tr> </thead> <tbody> <tr> <td>Inspire Medical Systems, Inc.</td> <td></td> <td></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Inspire Medical Systems, Inc.		
Licensee	Title	Date						
Inspire Medical Systems, Inc.								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <p><b>Copyright Number:</b>  <b>Copyright Holder:</b> Discloser</p> <table border="1"> <thead> <tr> <th>Licensee</th> <th>Title</th> <th>Date</th> </tr> </thead> <tbody> <tr> <td>Signifier Medical Technologies LLC</td> <td></td> <td></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Signifier Medical Technologies LLC		
Licensee	Title	Date						
Signifier Medical Technologies LLC								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <p><b>Copyright Number:</b>  <b>Copyright Holder:</b> Discloser</p> <table border="1"> <thead> <tr> <th>Licensee</th> <th>Title</th> <th>Date</th> </tr> </thead> <tbody> <tr> <td>Eli Lilly and Company</td> <td></td> <td></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Eli Lilly and Company		
Licensee	Title	Date						
Eli Lilly and Company								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						

Type	Is Licensed	Interest Held By						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <p style="text-align: right;"><b>Copyright Number:</b>  <b>Copyright Holder:</b> Discloser</p> <table border="1" data-bbox="107 218 1515 327"> <thead> <tr> <th data-bbox="107 218 1053 275">Licensee</th> <th data-bbox="1053 218 1279 275">Title</th> <th data-bbox="1279 218 1515 275">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 275 1053 327">Bioproject Deutschland GmbH</td> <td data-bbox="1053 275 1279 327"></td> <td data-bbox="1279 275 1515 327"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Bioproject Deutschland GmbH		
Licensee	Title	Date						
Bioproject Deutschland GmbH								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <p style="text-align: right;"><b>Copyright Number:</b>  <b>Copyright Holder:</b> Discloser</p> <table border="1" data-bbox="107 575 1515 684"> <thead> <tr> <th data-bbox="107 575 1094 632">Licensee</th> <th data-bbox="1094 575 1299 632">Title</th> <th data-bbox="1299 575 1515 632">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 632 1094 684">Ignis Therapeutics (Shanghai) Lmt</td> <td data-bbox="1094 632 1299 684"></td> <td data-bbox="1299 632 1515 684"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Ignis Therapeutics (Shanghai) Lmt		
Licensee	Title	Date						
Ignis Therapeutics (Shanghai) Lmt								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <p style="text-align: right;"><b>Copyright Number:</b>  <b>Copyright Holder:</b> Discloser</p> <table border="1" data-bbox="107 932 1515 1041"> <thead> <tr> <th data-bbox="107 932 753 989">Licensee</th> <th data-bbox="753 932 1127 989">Title</th> <th data-bbox="1127 932 1515 989">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 989 753 1041">Nyxoah SA</td> <td data-bbox="753 989 1127 1041"></td> <td data-bbox="1127 989 1515 1041"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Nyxoah SA		
Licensee	Title	Date						
Nyxoah SA								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <p style="text-align: right;"><b>Copyright Number:</b>  <b>Copyright Holder:</b> Discloser</p> <table border="1" data-bbox="107 1289 1515 1398"> <thead> <tr> <th data-bbox="107 1289 566 1346">Licensee</th> <th data-bbox="566 1289 1279 1346">Title</th> <th data-bbox="1279 1289 1515 1346">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 1346 566 1398">Icannex Healthcare Limited</td> <td data-bbox="566 1346 1279 1398">Functional Outcomes of Sleep Questionnaire</td> <td data-bbox="1279 1346 1515 1398">10/11/2023</td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Icannex Healthcare Limited	Functional Outcomes of Sleep Questionnaire	10/11/2023
Licensee	Title	Date						
Icannex Healthcare Limited	Functional Outcomes of Sleep Questionnaire	10/11/2023						
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <p style="text-align: right;"><b>Copyright Number:</b>  <b>Copyright Holder:</b> Discloser</p> <table border="1" data-bbox="107 1646 1515 1755"> <thead> <tr> <th data-bbox="107 1646 899 1703">Licensee</th> <th data-bbox="899 1646 1200 1703">Title</th> <th data-bbox="1200 1646 1515 1703">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 1703 899 1755">Valis Biosciences</td> <td data-bbox="899 1703 1200 1755"></td> <td data-bbox="1200 1703 1515 1755"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Valis Biosciences		
Licensee	Title	Date						
Valis Biosciences								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						

Type	Is Licensed	Interest Held By						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <table border="1" data-bbox="107 218 1516 327"> <thead> <tr> <th data-bbox="107 218 919 275">Licensee</th> <th data-bbox="919 218 1208 275">Title</th> <th data-bbox="1208 218 1516 275">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 275 919 327">RWS Life Sciences</td> <td data-bbox="919 275 1208 327"></td> <td data-bbox="1208 275 1516 327"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	RWS Life Sciences		
Licensee	Title	Date						
RWS Life Sciences								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <table border="1" data-bbox="107 575 1516 684"> <thead> <tr> <th data-bbox="107 575 919 632">Licensee</th> <th data-bbox="919 575 1208 632">Title</th> <th data-bbox="1208 575 1516 632">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 632 919 684">ResMed Germany</td> <td data-bbox="919 632 1208 684"></td> <td data-bbox="1208 632 1516 684"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	ResMed Germany		
Licensee	Title	Date						
ResMed Germany								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <table border="1" data-bbox="107 932 1516 1041"> <thead> <tr> <th data-bbox="107 932 1013 989">Licensee</th> <th data-bbox="1013 932 1256 989">Title</th> <th data-bbox="1256 932 1516 989">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 989 1013 1041">Harmony Biosciences LLC</td> <td data-bbox="1013 989 1256 1041"></td> <td data-bbox="1256 989 1516 1041"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Harmony Biosciences LLC		
Licensee	Title	Date						
Harmony Biosciences LLC								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <table border="1" data-bbox="107 1289 1516 1398"> <thead> <tr> <th data-bbox="107 1289 724 1346">Licensee</th> <th data-bbox="724 1289 1110 1346">Title</th> <th data-bbox="1110 1289 1516 1346">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 1346 724 1398">Axsome</td> <td data-bbox="724 1346 1110 1398"></td> <td data-bbox="1110 1346 1516 1398"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Axsome		
Licensee	Title	Date						
Axsome								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <table border="1" data-bbox="107 1646 1516 1755"> <thead> <tr> <th data-bbox="107 1646 919 1703">Licensee</th> <th data-bbox="919 1646 1208 1703">Title</th> <th data-bbox="1208 1646 1516 1703">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 1703 919 1755">LivaNova USA, Inc.</td> <td data-bbox="919 1703 1208 1755"></td> <td data-bbox="1208 1703 1516 1755"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	LivaNova USA, Inc.		
Licensee	Title	Date						
LivaNova USA, Inc.								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						

Type	Is Licensed	Interest Held By						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <table border="1" data-bbox="107 218 1515 327"> <thead> <tr> <th data-bbox="107 218 1036 270">Licensee</th> <th data-bbox="1036 218 1255 270">Title</th> <th data-bbox="1255 218 1515 270">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 270 1036 327">Clinical Outcomes Solutions</td> <td data-bbox="1036 270 1255 327"></td> <td data-bbox="1255 270 1515 327"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Clinical Outcomes Solutions		
Licensee	Title	Date						
Clinical Outcomes Solutions								
Copyright - Functional Outcomes of Sleep Questionnaire and Se	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire and Self-Efficacy Measure for Sleep Apnea  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <table border="1" data-bbox="107 596 1515 705"> <thead> <tr> <th data-bbox="107 596 993 651">Licensee</th> <th data-bbox="993 596 1248 651">Title</th> <th data-bbox="1248 596 1515 651">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 651 993 705">Verily Life Sciences, LLC</td> <td data-bbox="993 651 1248 705"></td> <td data-bbox="1248 651 1515 705"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Verily Life Sciences, LLC		
Licensee	Title	Date						
Verily Life Sciences, LLC								
Copyright - Functional Outcome of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcome of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <table border="1" data-bbox="107 953 1515 1062"> <thead> <tr> <th data-bbox="107 953 852 1008">Licensee</th> <th data-bbox="852 953 1175 1008">Title</th> <th data-bbox="1175 953 1515 1008">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 1008 852 1062">IQVIA RDS Inc.</td> <td data-bbox="852 1008 1175 1062"></td> <td data-bbox="1175 1008 1515 1062"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	IQVIA RDS Inc.		
Licensee	Title	Date						
IQVIA RDS Inc.								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <table border="1" data-bbox="107 1310 1515 1419"> <thead> <tr> <th data-bbox="107 1310 1021 1365">Licensee</th> <th data-bbox="1021 1310 1263 1365">Title</th> <th data-bbox="1263 1310 1515 1365">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 1365 1021 1419">Jazz Pharmaceuticals Inc.</td> <td data-bbox="1021 1365 1263 1419"></td> <td data-bbox="1263 1365 1515 1419"></td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Jazz Pharmaceuticals Inc.		
Licensee	Title	Date						
Jazz Pharmaceuticals Inc.								
Copyright - Functional Outcomes of Sleep Questionnaire	-	Self						
<p><b>Description:</b> Functional Outcomes of Sleep Questionnaire  <b>Filing Jurisdiction:</b>  <b>Licensees:</b></p> <table border="1" data-bbox="107 1667 1515 1776"> <thead> <tr> <th data-bbox="107 1667 380 1722">Licensee</th> <th data-bbox="380 1667 1235 1722">Title</th> <th data-bbox="1235 1667 1515 1722">Date</th> </tr> </thead> <tbody> <tr> <td data-bbox="107 1722 380 1776">Cryosa, Inc.</td> <td data-bbox="380 1722 1235 1776">Functional Outcomes of Sleep Questionnaire</td> <td data-bbox="1235 1722 1515 1776">09/29/2023</td> </tr> </tbody> </table> <p><b>Additional Information:</b></p>			Licensee	Title	Date	Cryosa, Inc.	Functional Outcomes of Sleep Questionnaire	09/29/2023
Licensee	Title	Date						
Cryosa, Inc.	Functional Outcomes of Sleep Questionnaire	09/29/2023						

## Additional Questions

1. **Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?**

Yes.

- a. **Please describe those relationships.**

Payment for writing manuscript for UpToDate website

2. **What is the manuscript title?**

Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity: The SURMOUNT-OSA Randomized Clinical Trials

3. **Are you the corresponding author?**

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.



**Discloser Identifier:** 1016453

**Disclosure Purpose:** 24-04881

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
Eli Lilly and Company	Employment	Self
<b>Title:</b> Senior Advisor-statistics <b>Additional Information:</b> Full time job		<b>Position Description:</b> Provide statistical leadership and expertise to clinical trials
Eli Lilly and Company	Stock	Self
<b>Additional Information:</b> Restricted stock units		

## Additional Questions

1. **Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?**

No.

2. **What is the manuscript title?**

Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity

3. **Are you the corresponding author?**

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.

Julia Dunn

**Discloser Identifier:** 1184235

**Disclosure Purpose:** 24-04881

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
Eli Lilly and Company	Employment	Self
<b>Title:</b> Executive Director, Global Medical Affairs <b>Additional Information:</b>		<b>Position Description:</b>
Eli Lilly and Company	Stock	Self
<b>Additional Information:</b>		

## Additional Questions

1. **Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?**

Yes.

a. **Please describe those relationships.**

Eli Lilly Employee and stock holder

2. **What is the manuscript title?**

Tirzepatide for the Treatment of OSA and Obesity

3. **Are you the corresponding author?**

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.



**Discloser Identifier:** 493122

**Disclosure Purpose:** 24-04881

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
AGB Pharma	Other	Self
<b>Category:</b> Other <b>Description:</b> Lecture <b>Additional Information:</b>		
Bioprojet	Other	Self
<b>Category:</b> Other <b>Description:</b> Lecture <b>Additional Information:</b>		
Eli Lilly and Company	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> OSA and Tirzepatide <b>Additional Information:</b>		
Hennig	Other	Self
<b>Category:</b> Other <b>Description:</b> Lectures <b>Additional Information:</b>		
Idorsia	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> OSA and Daridorexant <b>Additional Information:</b>		
Löwenstein	Grant / Contract	Self
<b>Recipient Name:</b> Charite - Universitätsmedizin Berlin <b>Grant / Contract Description:</b> Research Grant - Telemedicine <b>Additional Information:</b>		
<b>Recipient Type:</b> Institution <b>Grant / Contract Purpose:</b> Research		
Proctor & Gamble	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consultant and Lectures <b>Additional Information:</b>		
ResMed	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Telemedicine in Sleep Medicine <b>Additional Information:</b>		

Entity	Type	Interest Held By
Vanda Pharmaceuticals Inc.	Other	Self
<b>Category:</b> Other <b>Description:</b> Lecture <b>Additional Information:</b>		

## Additional Questions

1. **Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?**

No.

2. **What is the manuscript title?**

Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity

3. **Are you the corresponding author?**

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.

Discloser 699232  
Identifier:

Disclosure 24-04881  
Purpose:

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
Alkermes, Inc.	Grant / Contract	Self
<b>Recipient Name:</b> Woolcock Institute of Medical Research <b>Grant / Contract Description:</b> Funding for Phase 1b study of an orexin agonist <b>Additional Information:</b>		
<b>Recipient Type:</b> Institution <b>Grant / Contract Purpose:</b> Research		
Alkermes, Inc.	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Provide advice on narcolepsy <b>Additional Information:</b>		
EISAI INC.	Other	Self
<b>Category:</b> Other <b>Description:</b> Lecture fee <b>Additional Information:</b>		
Eli Lilly and Company	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Advisory Board for SURMOUNT OSA trial <b>Additional Information:</b> meetings every 2 months		
Vanda Pharmaceuticals Inc.	Grant / Contract	Self
<b>Recipient Name:</b> Woolcock Institute of Medical Research <b>Grant / Contract Description:</b> Set up Fee for proposed trial of tasimelteon in delayed sleep phase syndrome <b>Additional Information:</b>		
<b>Recipient Type:</b> Institution <b>Grant / Contract Purpose:</b> Research		

## Additional Questions

1. Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?

No.

2. What is the manuscript title?

Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity: The SURMOUNT-OSA Randomized Clinical Trials

3. Are you the corresponding author?

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.



**Discloser Identifier:** 793393

**Disclosure Purpose:** 24-04881

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
Apnimed	Grant / Contract	Other - Funds to Institution
<b>Recipient Name:</b> Brigham and Women's Hospital <b>Grant / Contract Description:</b> Study support <b>Additional Information:</b>		<b>Recipient Type:</b> Institution <b>Grant / Contract Purpose:</b> Research
Apnimed	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consulting <b>Additional Information:</b>		
Dynaflex	Grant / Contract	Other - Grant to Institution
<b>Recipient Name:</b> Brigham and Women's Hospital <b>Grant / Contract Description:</b> Research <b>Additional Information:</b>		<b>Recipient Type:</b> Institution <b>Grant / Contract Purpose:</b> Research
Forepont	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consulting, due diligence advice on scientific integrity of investment company <b>Additional Information:</b>		
Inspire Medical Systems, Inc.	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consultant <b>Additional Information:</b>		
LinguaFlex	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consultant, data analysis <b>Additional Information:</b>		
Nox Medical	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consultant, data analysis methods development <b>Additional Information:</b>		
Prosomnus	Grant / Contract	Other - Grant to Institution

Entity	Type	Interest Held By
<b>Recipient Name:</b> Brigham and Women's Hospital <b>Grant / Contract Description:</b> Research <b>Additional Information:</b>		
<b>Recipient Type:</b> Institution <b>Grant / Contract Purpose:</b> Research		
Respicardia, Inc.	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consulting, data analysis <b>Additional Information:</b>		

## Intellectual Property

Type	Is Licensed	Interest Held By						
Patent - Combination pharmacological interventions for mul	-	Self						
<b>Description:</b> Combination pharmacological interventions for multiple mechanisms of obstructive sleep apnea <b>Patent Number:</b> US20220362221A1 <b>Patent Status:</b> Pending <b>Filing Jurisdiction:</b> USA <b>Licensees:</b>								
<table border="1"> <thead> <tr> <th>Licensee</th> <th>Title</th> <th>Date</th> </tr> </thead> <tbody> <tr> <td>Apnimed</td> <td></td> <td></td> </tr> </tbody> </table>			Licensee	Title	Date	Apnimed		
Licensee	Title	Date						
Apnimed								
<b>Additional Information:</b>								
Patent - System and method for endo-phenotyping and risk s	-	Self						
<b>Description:</b> System and method for endo-phenotyping and risk stratfying obstructive sleep apnea <b>Patent Number:</b> US20230087299A1 <b>Patent Status:</b> Pending <b>Filing Jurisdiction:</b> USA <b>Licensees:</b> <b>Additional Information:</b> <a href="https://patents.google.com/patent/US20230087299A1/en">https://patents.google.com/patent/US20230087299A1/en</a>								

## Additional Questions

1. Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?

No.

2. What is the manuscript title?

Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity

3. Are you the corresponding author?

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.



Susan Redline

Discloser 778422  
Identifier:

Disclosure 24-04881  
Purpose:

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
<a href="#">Apnimed Inc</a>	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> consultant/member of Scientific Advisory board (unpaid for last year) <b>Additional Information:</b>		
<a href="#">Eli Lilly and Company</a>	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> provided input into a scientific protocol examining sleep apnea treatment <b>Additional Information:</b>		

## Additional Questions

1. **Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?**

No.

2. **What is the manuscript title?**

Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity: The SURMOUNT-OSA Randomized Clinical Trials

3. **Are you the corresponding author?**

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.



**Discloser Identifier:** 933190

**Disclosure Purpose:** 24-04881

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
Eli Lilly and Company	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> med ed <b>Additional Information:</b>		
Healcisio	Other Business Ownership	Self
<b>Form of Business Description:</b> cofounder <b>Additional Information:</b>		
LivaNova USA, Inc.	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> med ed <b>Additional Information:</b>		
sunrise	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> medical education <b>Additional Information:</b> advisor \$10K/yr		
University of California San Diego	Employment	Self
<b>Title:</b> prof <b>Additional Information:</b>		
<b>Position Description:</b> vice chair of medicin		
ZOLL Medical Corporation	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> med ed <b>Additional Information:</b>		

## Additional Questions

1. Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?

No.

2. What is the manuscript title?

Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity

**3. Are you the corresponding author?**

Yes.

**a. Please list the other authors' names here.**

Atul Malhotra, MD1, Ronald R. Grunstein, MD, PhD2, Ingo Fietze, MD3, Terri E. Weaver, PhD4, Susan Redline, MD, MPH5, Ali Azarbarzin, PhD5, Scott A. Sands, PhD5, Richard J. Schwab, MD6, Julia P. Dunn, MD, MS8, Sujatro Chakladar, PhD8, Mathijs C. Bunck, MD, PhD8, Josef Bednarik, MD, MBA8, for the SURMOUNT-OSA Investigators

## Certification

I certify that the information provided in this disclosure is complete and accurate.



**Discloser Identifier:** 1184234

**Disclosure Purpose:** 24-04881

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
CryOSA	Grant / Contract	Self
<b>Recipient Name:</b> <b>Grant / Contract Description:</b> Research grant <b>Additional Information:</b>		
Eli Lilly	Grant / Contract	Self
<b>Recipient Name:</b> <b>Grant / Contract Description:</b> Research Consultant <b>Additional Information:</b>		
eXciteOSA	Other	Self
<b>Category:</b> Other <b>Description:</b> Medical Advisory Board <b>Additional Information:</b>		
Inspire Medical Systems, Inc.	Grant / Contract	Self
<b>Recipient Name:</b> <b>Grant / Contract Description:</b> Research support <b>Additional Information:</b>		
Merck Manual	Other	Self
<b>Category:</b> Other <b>Description:</b> Royalties <b>Additional Information:</b>		
Resmed Corp	Grant / Contract	Self
<b>Recipient Name:</b> <b>Grant / Contract Description:</b> Grant/Research Support to study OSA/CSA and CPAP in hospitalized patients <b>Additional Information:</b>		
Sleep Evolution	Other	Self
<b>Category:</b> Other <b>Description:</b> Medical Advisory Board <b>Additional Information:</b>		
UpToDate	Other	Self

Entity	Type	Interest Held By
<b>Category:</b> Other <b>Description:</b> Royalties <b>Additional Information:</b>		

## Additional Questions

1. **Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?**

No.

2. **What is the manuscript title?**

Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity: The SURMOUNT-OSA Randomized Clinical Trials

3. **Are you the corresponding author?**

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.

**Discloser Identifier:** 1079632

**Disclosure Purpose:** 24-04881

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
Eli Lilly and Company	Employment	Self
<b>Title:</b> Associate Vice President <b>Additional Information:</b>		<b>Position Description:</b> I am the head of the tirzepatide obesity and NILEX medical development team.
Eli Lilly and Company	Stock	Self
<b>Additional Information:</b>		

## Additional Questions

1. **Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?**

No.

2. **What is the manuscript title?**

Tirzepatide for the treatment of obstructive sleep apnea and obesity

3. **Are you the corresponding author?**

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.

**Discloser Identifier:** 1184236

**Disclosure Purpose:** 24-04881

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
Eli Lilly and Company	Employment	Self
<b>Title:</b> Executive Director - Medical Tirzepatide-Obesity/NILEX <b>Additional Information:</b> full-time employee		<b>Position Description:</b> Clinical research physician in tirzepatide OSA development team

## Additional Questions

1. **Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?**

No.

2. **What is the manuscript title?**

Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity

3. **Are you the corresponding author?**

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.

**Discloser Identifier:** 793397

**Disclosure Purpose:** 24-04881

## Summary of Interests

### Company or Organization

Entity	Type	Interest Held By
<a href="#">Apnimed</a>	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consultant <b>Additional Information:</b>		
<a href="#">Cerebra</a>	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Working to include the hypoxic burden calculator into their PSG system. <b>Additional Information:</b>		
<a href="#">Eli Lilly and Company</a>	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consulting <b>Additional Information:</b>		
<a href="#">Inspire Medical Systems, Inc.</a>	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consulting <b>Additional Information:</b>		
<a href="#">LinguaFlex</a>	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consulting <b>Additional Information:</b>		
<a href="#">Somnifix</a>	Grant / Contract	Self
<b>Recipient Name:</b> Brigham and Women's Hospital <b>Grant / Contract Description:</b> Investigating effect of Somnifix mouth strips on adherence to CPAP <b>Additional Information:</b>		
<b>Recipient Type:</b> Institution <b>Grant / Contract Purpose:</b> Research		
<a href="#">Somnifix</a>	Consultant	Self
<b>Category:</b> Consultant <b>Description:</b> Consulting <b>Additional Information:</b>		
<a href="#">ZOLL Medical Corporation</a>	Consultant	Self

Entity	Type	Interest Held By
<b>Category:</b> Consultant <b>Description:</b> Consulting <b>Additional Information:</b>		

## Intellectual Property

Type	Is Licensed	Interest Held By
Patent - System and method for endo-phenotyping and risk s	-	Self
<b>Description:</b> System and method for endo-phenotyping and risk stratfying obstructive sleep apnea <b>Patent Status:</b> Pending <b>Filing Jurisdiction:</b> USA <b>Licensees:</b> <b>Additional Information:</b>		
<b>Patent Number:</b> US20230087299A1 <b>Patent Holder:</b> Current Institution		
Patent - Method, non-transitory computer readable medium a	-	Self
<b>Description:</b> Method, non-transitory computer readable medium and apparatus for arousal intensity scoring <b>Patent Status:</b> Issued <b>Filing Jurisdiction:</b> USA <b>Licensees:</b> <b>Additional Information:</b>		
<b>Patent Number:</b> 10349851 <b>Patent Holder:</b> Other Organization - Cerebra Health		

## Additional Questions

1. Are there other relationships or activities that readers could perceive to have influenced, or that give the appearance of potentially influencing, what you wrote in the submitted work?

No.

2. What is the manuscript title?

Tirzepatide for the Treatment of Obstructive Sleep Apnea and Obesity

3. Are you the corresponding author?

No.

## Certification

I certify that the information provided in this disclosure is complete and accurate.



# Protocol

Protocol for: Malhotra A, Grunstein RR, Fietze I, et al. Tirzepatide for the treatment of obstructive sleep apnea and obesity. *N Engl J Med*. DOI: 10.1056/NEJMoa2404881

This trial protocol has been provided by the authors to give readers additional information about the work.

# Protocol

This supplement contains the following items:

1. Master Protocol:
  - a. Original protocol, dated 27 January 2022
  - b. Final protocol with amendments, version (c), dated 02 June 2023
2. Study 1 (GPI1/ISA1)
  - a. Original protocol, dated 28 January 2022
  - b. Final protocol with amendments, version (b), dated 30 September 2022
3. Study 2 (GPI2/ISA2)
  - a. Original protocol, dated 28 January 2022
  - b. Final protocol with amendments, version (a), dated 12 February 2022
4. Statistical Analysis Plan:
  - a. Original statistical analysis plan, dated 12 January 2023
  - b. Final statistical analysis plan with revision history, dated 5 April 2024
  - c. Final integrated efficacy analysis plan with revision history, dated 27 February 2024

## Title Page

### Confidential Information

The information contained in this document is confidential and is intended for the use of clinical investigators. It is the property of Eli Lilly and Company or its subsidiaries and should not be copied by or distributed to persons not involved in the clinical investigation of tirzepatide (LY3298176), unless such persons are bound by a confidentiality agreement with Eli Lilly and Company or its subsidiaries.

**Note to Regulatory Authorities:** This document may contain protected personal data and/or commercially confidential information exempt from public disclosure. Eli Lilly and Company requests consultation regarding release/redaction prior to any public release.

In the United States, this document is subject to Freedom of Information Act (FOIA) Exemption 4 and may not be reproduced or otherwise disseminated without the written approval of Eli Lilly and Company or its subsidiaries.

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**Protocol Number:** I8F-MC-GPIF

**Amendment Number:** This is the initial protocol.

**List of Indication-Specific Appendices (ISAs):**

I8F-MC-GPI1: Participants with OSA unwilling or unable to use PAP therapy

I8F-MC-GPI2: Participants with OSA on PAP therapy

**Compound:** Tirzepatide (LY3298176)

**Brief Title:** A Master Protocol for Tirzepatide in Participants with Obstructive Sleep Apnea and Obesity

**Study Phase:** 3

**Acronym:** SURMOUNT-OSA

**Sponsor Name:** Eli Lilly and Company

**Legal Registered Address:** Indianapolis, Indiana, USA 46285

**Regulatory Agency Identifier Number:**

IND: 157090

**Approval Date:** Protocol Electronically Signed and Approved by Lilly on date provided below.

Approval Date: 27-Jan-2022 GMT

**Medical Monitor Name and Contact Information will be provided separately.**

## Table of Contents

<b>1.</b>	<b>Protocol Summary</b> .....	<b>7</b>
1.1.	Synopsis .....	7
1.2.	Schema .....	10
1.3.	Schedule of Activities (SoA) .....	12
<b>2.</b>	<b>Introduction</b> .....	<b>23</b>
2.1.	Study Rationale .....	23
2.2.	Background .....	23
2.3.	Benefit/Risk Assessment .....	24
<b>3.</b>	<b>Objectives, Endpoints, and Estimands</b> .....	<b>25</b>
<b>4.</b>	<b>Study Design</b> .....	<b>27</b>
4.1.	Overall Design .....	27
4.2.	Scientific Rationale for Study Design .....	27
4.3.	Justification for Dose .....	28
4.4.	End of Study Definition .....	28
<b>5.</b>	<b>Study Population</b> .....	<b>29</b>
5.1.	Inclusion Criteria .....	29
5.2.	Exclusion Criteria .....	30
5.3.	Lifestyle Considerations .....	34
5.3.1.	Meals and Dietary Restrictions .....	34
5.3.2.	Physical Activity .....	34
5.4.	Screen Failures .....	35
5.5.	Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant .....	35
<b>6.</b>	<b>Study Intervention(s) and Concomitant Therapy</b> .....	<b>36</b>
6.1.	Study Interventions Administered .....	36
6.1.1.	Medical Devices .....	37
6.2.	Preparation, Handling, Storage, and Accountability .....	37
6.3.	Measures to Minimize Bias: Randomization and Blinding .....	37
6.4.	Study Intervention Compliance .....	38
6.5.	Dose Modification .....	38
6.5.1.	Management of Participants with Gastrointestinal Symptoms .....	38
6.6.	Continued Access to Study Intervention after the End of the Study .....	39
6.7.	Treatment of Overdose .....	39
6.8.	Concomitant Therapy .....	39
6.8.1.	Management of Incident Diabetes .....	40
6.8.2.	Prohibited Concomitant Medications .....	40
<b>7.</b>	<b>Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal</b> .....	<b>42</b>
7.1.	Discontinuation of Study Intervention .....	42
7.1.1.	Liver Chemistry Stopping Criteria .....	44
7.1.2.	Temporary Discontinuation .....	44
7.2.	Participant Discontinuation/Withdrawal from the Study .....	45
7.3.	Lost to Follow up .....	46

<b>8.</b>	<b>Study Assessments and Procedures .....</b>	<b>47</b>
8.1.	Efficacy Assessments .....	47
8.1.1.	Primary Efficacy Assessment.....	47
8.1.2.	Secondary Efficacy Assessments .....	47
8.1.3.	Exploratory Efficacy Assessments .....	49
8.2.	Safety Assessments .....	51
8.2.1.	Physical Examinations.....	51
8.2.2.	Vital Signs.....	51
8.2.3.	Electrocardiograms (ECG).....	51
8.2.4.	Clinical Safety Laboratory Tests.....	51
8.2.5.	Pregnancy Testing .....	52
8.2.6.	Suicidal Ideation and Behavior Risk Monitoring.....	52
8.3.	Adverse Events, Serious Adverse Events, and Product Complaints .....	53
8.3.1.	Timing and Mechanism for Collecting Events .....	53
8.3.2.	Pregnancy.....	55
8.3.3.	Adverse Events of Special Interest.....	56
8.4.	Pharmacokinetics.....	57
8.4.1.	Bioanalytical Methods .....	57
8.5.	Pharmacodynamics.....	57
8.6.	Genetics .....	57
8.7.	Biomarkers .....	58
8.8.	Immunogenicity Assessments.....	58
8.9.	Health Economics OR Medical Resource Utilization and Health Economics.....	58
<b>9.</b>	<b>Statistical Considerations.....</b>	<b>59</b>
9.1.	Statistical Hypotheses.....	59
9.1.1.	Multiplicity Adjustment.....	59
9.2.	Analyses Sets .....	59
9.3.	Statistical Analyses.....	60
9.3.1.	General Considerations.....	60
9.3.2.	Primary Analysis .....	62
9.3.3.	Analysis of Key Secondary Endpoints .....	62
9.3.4.	Treatment Group Comparability .....	63
9.3.5.	Safety Analyses .....	63
9.3.6.	Evaluation of Immunogenicity.....	64
9.3.7.	Other Analyses .....	64
9.4.	Interim Analysis .....	65
9.5.	Sample Size Determination.....	65
<b>10.</b>	<b>Supporting Documentation and Operational Considerations .....</b>	<b>67</b>
10.1.	Appendix 1: Regulatory, Ethical, and Study Oversight Considerations.....	67
10.1.1.	Regulatory and Ethical Considerations.....	67
10.1.2.	Financial Disclosure .....	67
10.1.3.	Informed Consent Process .....	68
10.1.4.	Data Protection.....	68

10.1.5.	Committees Structure .....	68
10.1.6.	Dissemination of Clinical Study Data .....	69
10.1.7.	Data Quality Assurance .....	70
10.1.8.	Source Documents.....	71
10.1.9.	Study and Site Start and Closure.....	72
10.1.10.	Publication Policy.....	72
10.1.11.	Investigator Information .....	72
10.1.12.	Sample Retention.....	73
10.2.	Appendix 2: Clinical Laboratory Tests.....	74
10.2.1.	Laboratory Samples to be Obtained at the Time of a Systemic Hypersensitivity Event.....	76
10.3.	Appendix 3: Adverse Events and Serious Adverse Events: Definitions and Procedures for Recording, Evaluating, Follow-up, and Reporting .....	78
10.3.1.	Definition of AE .....	78
10.3.2.	Definition of SAE.....	79
10.3.3.	Definition of Product Complaints .....	80
10.3.4.	Recording and Follow-Up of AE and/or SAE and Product Complaints .....	81
10.3.5.	Reporting of SAEs.....	82
10.3.6.	Regulatory Reporting Requirements .....	83
10.3.7.	Special Safety Topics .....	83
10.4.	Appendix 4: Contraceptive and Barrier Guidance .....	86
10.4.1.	Definitions.....	86
10.4.2.	Contraception Guidance.....	87
10.5.	Appendix 5: Genetics .....	89
10.6.	Appendix 6: Liver Safety: Suggested Actions and Follow-up Assessments .....	90
10.6.1.	Hepatic Evaluation Testing.....	90
10.6.2.	Close Hepatic Monitoring.....	92
10.6.3.	Comprehensive Hepatic Evaluation .....	92
10.7.	Appendix 7: Medical Device Adverse Events (AEs), Adverse Device Effects (ADEs), Serious Adverse Events (SAEs) and Device Deficiencies: Definition and Procedures for Recording, Evaluating, Follow-up, and Reporting .....	94
10.8.	Appendix 8: Protocol GPIF Standardized Protocols for the Measurement of Height, Weight, Neck Circumference, Waist Circumference, Vital Signs, and Electrocardiogram .....	95
10.8.1.	Measuring Height .....	95
10.8.2.	Measuring Weight .....	95
10.8.3.	Measuring Hip and Waist Circumference.....	95
10.8.4.	Measuring Neck Circumference.....	96
10.8.5.	Vital Sign Measurements (Blood Pressure and Heart Rate).....	96
10.8.6.	Electrocardiogram .....	97
10.9.	Appendix 9: Provisions for Changes in Study Conduct Due to the COVID-19 Pandemic .....	98

10.10. Appendix 10: Abbreviations and Definitions ..... 102

10.11. Appendix 11: Country-Specific Requirements ..... 106

10.11.1. Germany..... 106

**11. References ..... 110**



## 1. Protocol Summary

### 1.1. Synopsis

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**Brief Title:** A Master Protocol for Tirzepatide in Participants with Obstructive Sleep Apnea and Obesity

#### **Rationale:**

Obstructive sleep apnea (OSA) is a breathing disorder associated with significant comorbidity and mortality. Currently available therapeutic approaches have shown moderate success in treating the clinical signs and symptoms of OSA (that is, snoring and excessive daytime sleepiness) but have failed to address the underlying pathophysiology of the disease and, more importantly, the cardiovascular (CV) morbidity and mortality associated with OSA.

Tirzepatide is a dual GIP/GLP-1R agonist that has demonstrated statistically significant and clinically relevant lowering in HbA1c and dose-dependent weight loss in 5 Phase 3 trials that enrolled patients with type 2 diabetes mellitus (T2DM) (Dahl et al. 2021; Frias et al. 2021; Lilly 2021; Ludvik et al. 2021; Rosenstock et al. 2021). Data from clinical studies show that a large proportion of participants have significant body weight reduction with tirzepatide treatment, suggesting that this agent may represent a promising pharmacologic treatment option that could reduce the frequency of apnea and hypopnea events as well as reduce weight, decrease blood pressure, and improve insulin resistance and dyslipidemia, which are all features associated with the increase in CV morbidity and mortality as seen in people living with OSA.

**Objectives, Endpoints, and Estimands:**

The following objectives and endpoints apply to both indication-specific appendices (ISAs).

Primary Objective	Endpoints
To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for mean percent decrease in AHI.	Percent change in AHI from baseline to Week 52.
Key Secondary Objectives (controlled for type I error)	Endpoints
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• A hierarchical assessment of PROs</li> <li>• Clinically meaningful change in AHI</li> <li>• Achieving OSA remission or mild nonsymptomatic OSA</li> <li>• Change in body weight</li> <li>• Change in SBP</li> <li>• Change in inflammatory status</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• A hierarchical combination of the following: <ul style="list-style-type: none"> <li>○ Change in FOSQ-10 score</li> <li>○ Change in FOSQ (30 items) Vigilance domain score</li> <li>○ Change in FOSQ (30 items) Activity Level domain score</li> </ul> </li> <li>• Percent of participants with <math>\geq 50\%</math> AHI reduction</li> <li>• Percent of participants with <ul style="list-style-type: none"> <li>○ AHI <math>&lt; 5</math> or</li> <li>○ AHI 5-14 with ESS <math>\leq 10</math></li> </ul> </li> <li>• Percent change in body weight</li> <li>• Change in SBP</li> <li>• Change in hsCRP concentration</li> </ul>

Abbreviations: AHI = Apnea-Hypopnea Index; ESS = Epworth Sleepiness Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; hsCRP = high-sensitivity C reactive protein; MTD = maximum-tolerated dose; OSA = obstructive sleep apnea; PROs = patient-reported outcomes; SBP = systolic blood pressure; QW = once weekly.

For estimands guiding statistical analyses, see Section 9.3.1.

**Overall Design:**

The overall study design consists of 2 components:

- Master Protocol: defines study elements common for both populations.
- ISAs: provide detailed population-specific information.

Study I8F-MC-GPIF (GPIF) is multicenter, randomized, parallel-arm, double-blind, placebo-controlled, Phase 3 study with 52-week treatment duration conducted under a basket-design, which will investigate the effects of treatment with weekly (QW) tirzepatide at the maximum tolerated dose (MTD) (10 mg or 15 mg), compared with placebo in participants who have moderate-to-severe OSA and obesity.

One master protocol will support 2 studies/ISAs.

- ISA 1 (GPI1) will include participants who are unwilling or unable to use PAP therapy.
- ISA 2 (GPI2) will include participants who are on PAP therapy for at least 3 months at time of screening and plan to continue PAP therapy during the study.

Participants will be assigned to the ISA which reflects their current PAP usage. The participant will then be randomly assigned 1:1 to treatment or placebo.

**Number of Participants:**

Approximately 412 participants will be randomly assigned to study intervention across the entire master protocol, with approximately 206 participants randomly assigned to study intervention in each ISA. See Section 9.5 for additional information.

An upper limit of approximately 70% enrollment of male participants will be used to ensure a sufficiently large sample of female participants.

**Intervention Groups and Duration:**

The study interventions are:

- tirzepatide at the MTD (10 mg or 15 mg) SC QW, or
- placebo.

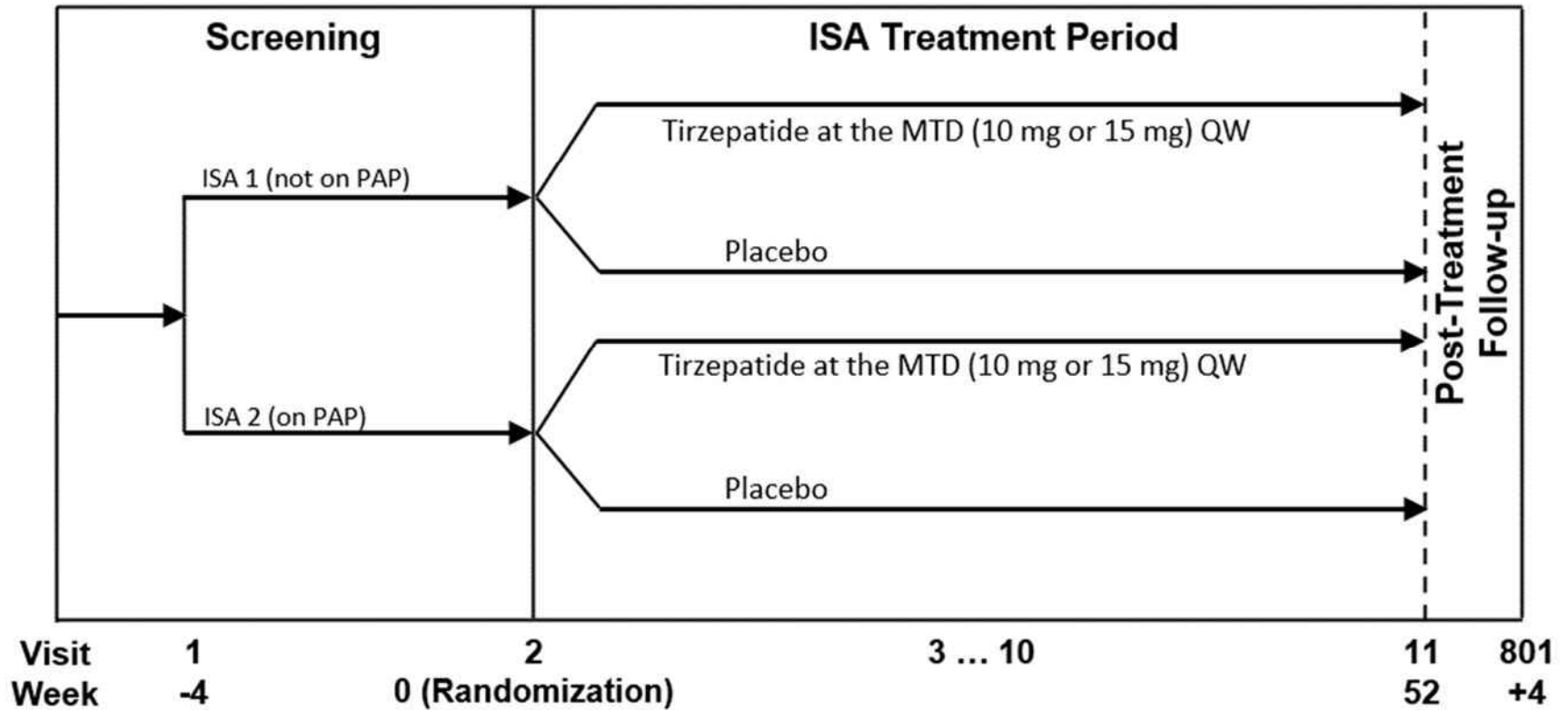
The expected total duration of study participation for each participant, including screening and the post-treatment follow-up periods, is 60 weeks across the following study periods:

- Screening: 4 weeks
- Treatment: 52 weeks
- Post-treatment follow-up: 4 weeks

The maximum duration of treatment is 52 weeks.

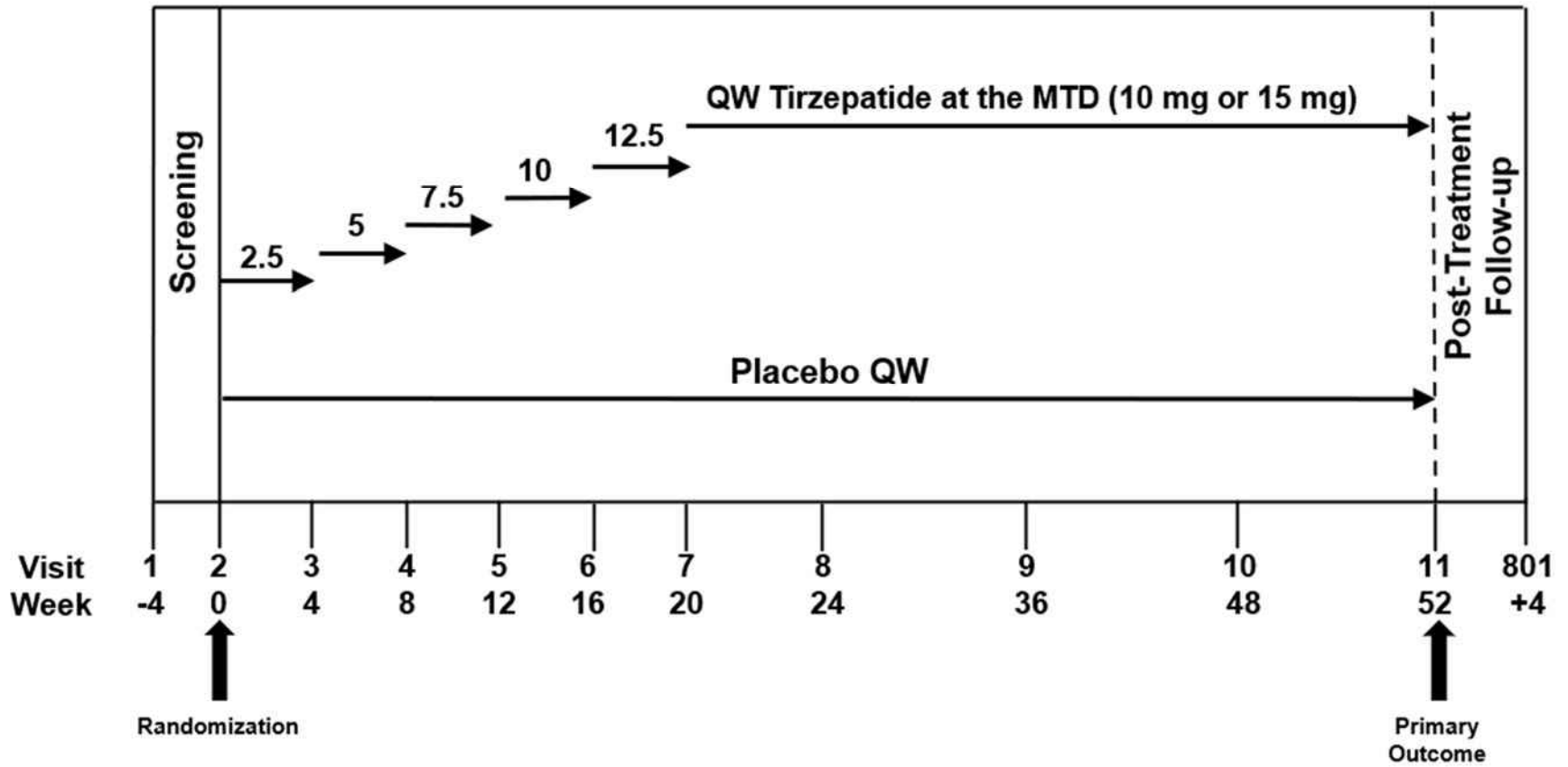
**Data Monitoring Committee: Yes**

**1.2. Schema**  
**Master Protocol and ISA Schema**



Abbreviations: ISA = intervention-specific appendix; MTD = maximum-tolerated dose; PAP = positive airway pressure.

**Dose Escalation and Visit Schema**



Abbreviations: MTD = maximum-tolerated dose; QW = once weekly.

### **1.3. Schedule of Activities (SoA)**

If Study Period I - Screening takes longer or shorter than 4 weeks to complete, it will not be considered a protocol deviation. Study Period I - Screening should not exceed 6 weeks unless approval from the sponsor is received.

Visit procedures may be conducted over more than 1 day as long as all activities are completed within the allowable visit tolerance period of each visit.

Study Period II - Treatment: For early discontinuations (ED) from study that occur before the last visit in treatment period, see the activities listed for ED in this table.

Note: Visit 10 is eligible to be conducted remotely (that is, by telephone, IT-assisted virtual visit, or in combination with on-site visit) at the direction of the sponsor, according to local laws and regulations.

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
		2	3	4	5	6	7	8	9	10	11	ED		
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b. If V10 is remote, then the visit will not be fasting.
Informed consent	X													The informed consent must be signed before any protocol-specific tests/procedures are performed.
Inclusion and exclusion criteria, review and confirm	X	X												Inclusion/Exclusion criteria should be confirmed prior to drug assignment and administration of first dose of study intervention.
Demographics	X													Includes ethnicity, year of birth, gender (sex), and race.
Preexisting conditions and medical history, including relevant surgical history	X													All conditions ongoing and relevant past surgical and medical history should be collected.
Prespecified medical history	X													Should include, but not be limited to, collecting diagnosis of OSA and obesity-related health problems.
Prior treatments for indication	X													Include OSA and obesity.
Substance use (alcohol, caffeine, tobacco use)	X													
Concomitant medications	X	X	X	X	X	X	X	X	X	X	X	X	X	

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b. If V10 is remote, then the visit will not be fasting.
AEs	X	X	X	X	X	X	X	X	X	X	X	X	X	Any events that occur after signing the informed consent are considered AEs as defined in Section 8.3.1. Additional data are collected for certain AEs.
<b>Physical Evaluation</b>														
Height	X													
Weight	X	X	X	X	X	X	X	X	X	X	X	X	X	Weight measurements should be obtained per the instructions in 10.8. If V10 is remote, then weight will not be obtained at the visit.
Waist circumference		X					X				X	X	X	Waist circumference should be obtained per the instructions in Section 10.8.
Hip circumference		X					X				X	X	X	Hip circumference should be obtained per the instructions in Section 10.8.
Neck circumference		X					X				X	X	X	Neck circumference should be obtained per the instructions in Section 10.8.
Vital signs	X	X	X	X	X	X	X	X	X	X	X	X	X	Includes PR and BP. Measured after participant has been sitting at least 5 minutes. Vital-sign measurements should be taken before obtaining an ECG tracing and before collection of blood samples for laboratory testing. See Sections 8.2.2 and 10.8. If V10 is remote, vital signs will not be obtained at the visit.



	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b. If V10 is remote, then the visit will not be fasting.
Complete physical examination	X													The complete physical examination is performed (excludes pelvic, rectal, and breast examinations unless clinically indicated).
Symptom-directed physical assessment		X	X	X	X	X	X	X	X	X	X	X	X	As indicated based on participant status and standard of care. May be performed by qualified personnel per local regulations. If V10 is remote, then symptom-directed physical assessment will not be done at the visit.
12-lead ECG (local)	X										X	X		ECG measurements should be obtained per the instructions in Section 8.2.3.
<b>Wearable Devices and PSG Assessments</b>														
Schedule Sleep Center Study for PSG	X					X				X		X		PSG results must be reviewed to confirm eligibility prior to randomization.
Sleep Center Study for PSG	X						X				X	X		PSG at V7 and 11 may be scheduled for any day +/- 14 days.
Participant wears the WatchPat300	X		X		X		X				X			Applicable only to participants in GPI1 (off PAP). Training documents will detail information on the dispensation, wearing, and return process.

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b. If V10 is remote, then the visit will not be fasting.
Participant wears actigraphy (AX6) device	X		X		X		X					X		Training documents will detail information on the dispensation, wearing, and return process.
<b>Participant Education and Supplies</b>														
eDiary education	X													Additional training can be repeated, as needed.
Train participant on study intervention administration		X												Re-training is available anytime.
Review Lifestyle Program instructions		X	X	X	X	X	X	X	X	X	X	X		
Review diet and exercise goals		X	X	X	X	X	X	X	X	X	X	X		All training should be repeated as needed to ensure participant compliance. Study personnel to provide reinforcement and encouragement for lifestyle modifications.
<b>Participant Diary (Electronic)</b>														
Participant diary dispensed	X													Includes the following: Sleep diary; Dosing diary; Hypoglycemic events (when applicable), Patient Diet and Exercise diary, PAP adherence (for GPI2 only).
Diary compliance check		X	X	X	X	X	X	X	X	X	X	X		
Diary return											X	X		

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b. If V10 is remote, then the visit will not be fasting.
<b>Patient-Reported Outcomes (Electronic)</b> <i>Complete prior to any clinical-administered assessments</i>														
ESS	X		X		X		X				X	X		When the PROs are scheduled for visits at which the PSG will be done, they should be completed in the following order (FOSQ, ESS, PROMIS Short Form v1.0 Sleep Disturbance 8b, PROMIS Short Form v1.0 Sleep-related Impairment 8a, PGIS, PGIC, SF-36v2 acute form, and EQ-5D-5L) before the PSG is conducted, and should be done at the same time of day for each of those visits.
EQ-5D-5L	X										X	X		
FOSQ	X				X		X				X	X		
PGIS (OSA Symptom Scales)	X		X		X		X				X	X		
PGIC (OSA Symptom Scales)			X		X		X				X	X		
PROMIS Short Form v1.0 Sleep Disturbance 8b	X		X		X		X				X	X		
PROMIS Short Form v1.0 Sleep-related Impairment 8a	X		X		X		X				X	X		
SF-36v2, acute	X						X				X	X		
PHQ-9	X	X					X				X	X		The PHQ-9 should be administered after assessment of AEs.

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b. If V10 is remote, then the visit will not be fasting.
<b>Clinician-Administered Assessments (Paper)</b>														
C-SSRS screening/baseline	X													The C-SSRS should be administered after assessment of AEs. The C-SSRS since last visit is adapted for the assessment of the ideation and behavior categories only. The Intensity of Ideation and Lethality of Behavior sections are removed.
C-SSRS (since last visit version)		X	X	X	X	X	X	X	X	X	X	X	X	
<b>Laboratory Tests and Sample Collections</b>														
Hematology	X				X			X	X		X	X	X	
HbA1c	X	X			X			X	X		X	X	X	
Clinical chemistry (includes glucose)	X	X			X			X	X		X	X	X	
Lipid panel		X						X			X	X	X	
hsCRP		X						X			X	X		
Serum pregnancy	X													Only for WOCBP and females with a history of tubal ligation. See Section 10.4, Appendix 4.

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b. If V10 is remote, then the visit will not be fasting.
Urine pregnancy (local)		X			X			X	X		X	X		A urine pregnancy test must be performed at V2 with the result available prior to first dose/injection of study intervention for WOCBP only. Additional pregnancy tests (beyond those required per the SoA) should be performed at any time during the trial if a menstrual period is missed, there is clinical suspicion of pregnancy, or as required by local law or regulation.
FSH	X													Optional; performed as needed to confirm postmenopausal status. See Section 10.4, Appendix 4.
Insulin		X						X			X	X	X	
C-Peptide		X						X			X	X	X	
Free fatty acids		X						X			X	X	X	
Cystatin-C	X							X			X	X	X	
Calcitonin	X										X	X	X	
Pancreatic amylase	X				X			X			X	X	X	
Lipase	X				X			X			X	X	X	
TSH	X													
eGFR	X	X			X			X	X		X	X	X	Calculated using CKD-EPI method.

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b. If V10 is remote, then the visit will not be fasting.
UACR	X	X			X			X	X		X	X	X	
PK samples		X	X		X			X			X	X	X	PK samples to be collected predose and close to ADA samples.
Immunogenicity (ADA) samples		X	X		X			X			X	X	X	In the event of systemic drug hypersensitivity reactions (immediate or nonimmediate), additional unscheduled samples should be collected as detailed in Section 10.3.7.2 (Hypersensitivity Reactions). Immunogenicity samples and PK samples for immunogenicity must be predose.
Stored Samples														
Genetics sample		X												Sample can be obtained at or after the specified visit.
Exploratory biomarker samples		X						X			X	X	X	
Randomization and Dosing														
Register visit with IWRS	X	X	X	X	X	X	X	X	X	X	X	X	X	X
ISA assignment via IWRS	X													
ISA treatment randomization via IWRS		X												

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b. If V10 is remote, then the visit will not be fasting.
Observe participant administer study intervention		X												Participants should administer their first dose of study intervention at the end of the V2, after other study procedures are completed.
Dispense study drug via IWRS		X	X	X	X	X	X	X	X	X				
Dispense study drug to participant (for at home dosing)		X	X	X	X	X	X	X	X	X				
Dispense ancillary supplies to participant		X												Dispensation of ancillary supplies may vary beyond V2 based on expiry dating of applicable supplies and/or participant needs.
Participant returns all unused study intervention			X	X	X	X	X	X	X	X	X	X		If V10 is remote, participant will return any unused study intervention dispensed at V9 during the V11 visit.
Assess study intervention compliance			X	X	X	X	X	X	X	X	X	X		If V10 is remote, the next study intervention compliance will be done at V11.

Abbreviations: ADA = antidrug antibody; AEs = adverse events; AX6 = Axivity 6; CKD EPI = Chronic Kidney Disease Epidemiology; hsCRP = high sensitivity C-reactive protein; C-SSRS = Columbia-Suicide Severity Rating Scale; ECG = electrocardiogram; ED = early discontinuation; eGFR = estimated glomerular filtration rate; EQ-5D-5L = EuroQol - 5 Dimension - 5 Level; ESS = Epworth Sleepiness Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; FSH = follicle-stimulating hormone; HbA1c = hemoglobin A1c; ISA = indication-specific appendix; IWRS = interactive web-response system; OSA = obstructive sleep apnea; PHQ = Patient Health Questionnaire-9; PGIC = Patient Global Impression of Change; PGIS = Patient Global Impression of Status; PK = pharmacokinetic; PSG = polysomnography; PR = pulse rate; PRO = patient reported outcome; PROMIS = Patient-Reported Outcomes Measurement Information System; SF-36v2 = Short-Form 36 version 2; SoA = schedule of activities; TSH = thyroid-stimulating hormone; UACR = urinary albumin/creatinine ratio; V = visit; WOCBP = woman of childbearing potential.

- a Post-treatment follow-up occurs approximately 4 weeks after the participant's final treatment period visit.
- b Fasting visit: On all office visits, study participants should be reminded to report to the site before taking study intervention in a fasting condition, after a period of approximately 8 hours without eating, drinking (except minimal amount of water, as needed), or any significant physical activity.



## **2. Introduction**

### **2.1. Study Rationale**

Obstructive sleep apnea is a breathing disorder associated with significant comorbidity and mortality. Currently available therapeutic approaches have shown moderate success in treating the clinical signs and symptoms of OSA (that is, snoring and excessive daytime sleepiness) but have failed to address the underlying pathophysiology of the disease and, more importantly, the CV morbidity and mortality associated with OSA.

Tirzepatide is a dual GIP/GLP-1R agonist that has demonstrated statistically significant and clinically relevant lowering in HbA1c and dose-dependent weight loss in 5 Phase 3 trials that enrolled patients with T2DM (Dahl et al. 2021; Frias et al. 2021; Lilly 2021; Ludvik et al. 2021; Rosenstock et al. 2021). Data from clinical studies show that a large proportion of participants have significant body weight reduction with tirzepatide treatment, suggesting that this agent may represent a promising pharmacologic treatment option that could reduce the frequency of apnea and hypopnea events as well as reduce weight, decrease BP, and improve insulin resistance and dyslipidemia, which are all features associated with the increase in CV morbidity and mortality as seen in people living with OSA.

### **2.2. Background**

Obstructive sleep apnea is a serious medical condition with a limited number of therapeutic options and high unmet need. OSA is a well-established risk factor for morbidity and mortality from CV and other metabolic diseases; further, it negatively affects daily life of patients (Tietjens et al. 2019; Gottlieb and Punjabi 2020).

Treatments for OSA include behavioral measures like weight loss programs, medical devices such as PAP, oral appliances or hypoglossal nerve stimulation, and surgery, including bariatric surgery in people living with obesity and OSA. Weight loss through caloric restriction and lifestyle interventions improves AHI, cardiometabolic comorbidities, and quality of life.

Although a limited number of studies have investigated the effect of pharmacological weight-loss therapy, their results support that adding weight-loss medication to lifestyle intervention reduces AHI and improves sleep quality and possibly other OSA-related outcomes. As such, standard of care recommends a prioritization of weight loss for patients with OSA and obesity or who are overweight, including consideration for FDA-approved antiobesity pharmacotherapy (Hudgel et al. 2018). Bariatric surgery is available for patients with severe obesity and can dramatically improve OSA (Currie et al. 2021). However, such surgical procedures are associated with greater risk for perioperative and postoperative complications (Opperer et al. 2016).

Positive airway pressure is the first-line treatment for moderate-to-severe and symptomatic mild OSA. Despite good clinical efficacy for the signs and symptoms of OSA, randomized controlled studies in PAP therapy have failed to show improvements in nonsleep-related outcomes linked to OSA such as stroke, heart attack, diabetes and depression (AHRQ 2021). It is thought inadequate PAP therapy adherence may explain why randomized clinical trials have failed to demonstrate benefit (Gottlieb and Punjabi 2020). Nonadherence to PAP therapy is a major clinical challenge

and studies have shown wide variability in adherence, ranging from 29% to 83% (Weaver and Grunstein 2008). Currently, there are only few pharmacologic alternatives indicated for symptoms in patients with OSA and none are disease modifying.

There is significant unmet need in treatment of patients with OSA. Tirzepatide, a GIP and GLP-1R dual agonist, has the potential to provide benefit to patients with OSA who have obesity by reducing weight, which may reduce the frequency of apnea and hypopnea events, decreasing BP, and improving insulin resistance and dyslipidemia.

### **2.3. Benefit/Risk Assessment**

More detailed information about the known and expected benefits and risks and reasonably expected AEs of tirzepatide may be found in the IB.

### 3. Objectives, Endpoints, and Estimands

The following objectives and endpoints apply to both ISAs.

Primary Objective	Endpoints
To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for mean percent decrease in AHI.	Percent change in AHI from baseline to Week 52.
Key Secondary Objectives (controlled for type I error)	Endpoints
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• A hierarchical assessment of PROs</li> <li>• Clinically meaningful change in AHI</li> <li>• Achieving OSA remission or mild nonsymptomatic OSA</li> <li>• Change in body weight</li> <li>• Change in SBP</li> <li>• Change in inflammatory status</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• A hierarchical combination of the following: <ul style="list-style-type: none"> <li>○ Change in FOSQ-10 score</li> <li>○ Change in FOSQ (30 items) Vigilance domain score</li> <li>○ Change in FOSQ (30 items) Activity Level domain score</li> </ul> </li> <li>• Percent of participants with <math>\geq 50\%</math> AHI reduction</li> <li>• Percent of participants with <ul style="list-style-type: none"> <li>○ AHI <math>&lt; 5</math> or</li> <li>○ AHI 5-14 with ESS <math>\leq 10</math></li> </ul> </li> <li>• Percent change in body weight</li> <li>• Change in SBP</li> <li>• Change in hsCRP concentration</li> </ul>
Other Secondary Objectives	Endpoints
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in excessive daytime sleepiness</li> <li>• Change in patient-reported functional status as assessed by FOSQ (30 items)</li> <li>• Change in body weight</li> <li>• Change in lipid parameters</li> <li>• Change in DBP</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in ESS score</li> <li>• Change in all other FOSQ domain scores</li> <li>• Percent of participants who achieve <ul style="list-style-type: none"> <li>○ <math>\geq 10\%</math> body weight reduction</li> <li>○ <math>\geq 15\%</math> body weight reduction</li> <li>○ <math>\geq 20\%</math> body weight reduction</li> </ul> </li> <li>• Change in <ul style="list-style-type: none"> <li>○ HDL-cholesterol</li> <li>○ non-HDL-cholesterol</li> <li>○ triglycerides</li> </ul> </li> <li>• Change in DBP</li> </ul>

<ul style="list-style-type: none"> <li>• Change in PROs</li>   <li>• Insulin</li> <li>• Hypoxic burden</li> </ul>	<ul style="list-style-type: none"> <li>• Change in:             <ul style="list-style-type: none"> <li>○ PROMIS Sleep-related impairment short form 8a score</li> <li>○ PROMIS Sleep disturbance short form 8b score</li> <li>○ SF-36v2 acute form domain scores</li> </ul> </li> <li>• Percent of participants with improved categorical shift in:             <ul style="list-style-type: none"> <li>○ PGIS-OSA Sleepiness</li> <li>○ PGIS-OSA Fatigue</li> <li>○ PGIS-OSA Snoring</li> </ul> </li> <li>• Change in fasting insulin</li> <li>• Change in SASHB (% min/hour)</li> </ul>
Exploratory Objectives	Endpoints
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in PROs</li>   <li>• To evaluate the effect of tirzepatide on sleep parameters as measured by Actigraphy (AX6)</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in             <ul style="list-style-type: none"> <li>○ EQ-5D-5L utility index</li> <li>○ EQ-VAS scores</li> </ul> </li> <li>• Percent of participants with improved categorical shift in:             <ul style="list-style-type: none"> <li>○ PGIC-OSA Sleepiness</li> <li>○ PGIC-OSA Fatigue</li> <li>○ PGIC-OSA Sleep quality</li> <li>○ PGIC-OSA Snoring</li> </ul> </li> <li>• Mean change from baseline to endpoint assessment in             <ul style="list-style-type: none"> <li>○ Daytime sleep duration</li> <li>○ Daily step counts</li> <li>○ Average acceleration</li> </ul> </li> </ul>

Abbreviations: AHI = Apnea-Hypopnea Index; AX6 = Axivity 6; DBP = diastolic blood pressure; ESS = Epworth Sleepiness Scale; EQ-5D-5L = EuroQol - 5 Dimension - 5 Level; EQ-VAS = EuroQol Visual Analogue Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; HDL = high-density lipoprotein; hsCRP = high-sensitivity C reactive protein; MTD = maximum-tolerated dose; OSA = obstructive sleep apnea; PGIC-OSA = Patient Global Impression of Change – Obstructive Sleep Apnea; PGIS-OSA = Patient Global Impression of Status – Obstructive Sleep Apnea; PROs = patient-reported outcomes; PROMIS = Patient-Reported Outcomes Measurement Information System; SBP = systolic blood pressure; SF-36v2 = Short-Form 36 version 2, SASHB = sleep apnea-specific hypoxic burden; QW = once weekly.

For estimands guiding statistical analyses, see Section 9.3.1.

## 4. Study Design

### 4.1. Overall Design

Study I8F-MC-GPIF (GPIF) is a multicenter, randomized, parallel-arm, double-blind, placebo-controlled Phase 3 study to evaluate the efficacy and safety of tirzepatide at the MTD (10 mg or 15 mg) QW versus placebo in participants who have obesity and moderate-to-severe OSA.

This basket-type master protocol will investigate 2 participant populations, described in 2 ISAs:

- GPI1 will include participants who are unwilling or are unable to use PAP therapy.
- GPI2 will include participants who are on PAP therapy for at least 3 months at time of screening and plan to continue PAP therapy during the study.

Participants to be assigned to whichever ISA they qualify for. Participants will then be randomly assigned to:

- tirzepatide at the MTD (10 mg or 15 mg) SC QW, or
- placebo.

The expected total duration of study participation for each participant, including screening and the post-treatment follow-up periods, is 60 weeks across the following study periods:

- Screening: 4 weeks
- Treatment: 52 weeks
- Post-treatment follow-up: 4 weeks

The maximum duration of treatment is 52 weeks. Procedures and assessments for each visit are presented in the SoA, Section 1.3.

### 4.2. Scientific Rationale for Study Design

The basket trial design employs a single overarching trial structure as a means to implement multiple investigations (Woodcock and LaVange 2017). Such approaches often allow for more streamlined and coordinated clinical trial logistics and consistency in data collection methods.

The 2 studies associated with the master protocol, also called “Intervention-Specific Appendices” (ISAs), will describe any objectives, endpoints and efficacy assessments specific to each study population. ISA refers to specific population and background intervention, as defined by inclusion criteria. Used together, the master protocol and the 2 ISAs will describe the investigations to be conducted.

ISA 1 and ISA 2 represent different populations with different treatment needs, and it is anticipated that tirzepatide may meet the needs of each participant group.

Inclusion of a placebo treatment arm is acceptable because there are no currently effective disease-modifying treatments for OSA; this approach is in agreement with the use of placebo described in the Declaration of Helsinki (WMA 2013). The use of a placebo comparator in Study GPIF is needed to determine the efficacy and safety of tirzepatide therapy.

### 4.3. Justification for Dose

Tirzepatide doses of up to 15 mg administered SC QW will be evaluated in this study. Participants may be treated with lower maintenance dose of 10 mg if they do not achieve full dose escalation to 15 mg and/or do not tolerate 15 mg.

These doses and associated escalation schemes were selected based on assessment of safety, efficacy (weight loss), and GI tolerability data in Phase 1, 2, and 3 studies in patients with T2DM, followed by exposure-response modeling of the data that predicted weight loss in patients with overweight or obesity.

Dosing algorithms starting at a low dose of 2.5 mg accompanied by dose escalation of 2.5 mg increments every 4 weeks should permit time for development of tolerance to GI events and are predicted to minimize GI tolerability concerns.

Similar to the GLP-1RA class, most of the tirzepatide AEs were dose-dependent and GI-related, consisting mainly of nausea, vomiting, and diarrhea. In general, these events were mild or moderate in severity, with few severe episodes, and transient.

Tirzepatide doses of 10 mg and 15 mg as MTD were selected based principally on the following criteria:

- each dose provides robust weight loss relative to placebo
- the percent of participants achieving  $\geq 10\%$  weight loss is higher with 15 mg than 10 mg, and
- safety and tolerability were supported by Phase 3 results in T2DM.

The proposed tirzepatide maintenance dose of up to 15 mg using a dose-escalation regimen is expected to safely maximize the potential for weight loss while minimizing GI tolerability concerns as has been achieved for approved GLP-1 RA drugs. The subsequent benefit of weight loss should result in clinically meaningful improvements in AHI.

### 4.4. End of Study Definition

A participant is considered to have completed the study if they have completed all required phases of the study including the last visit or the last scheduled procedure shown in the SoA. The end of the study for each ISA is defined as the date of the last visit of the last participant in the study or last scheduled procedure shown in the SoA for the last participant in the trial globally. The end of study timing for each ISA is independent of the other ISA end of study timing.

## 5. Study Population

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

### 5.1. Inclusion Criteria

Participants are eligible to be included in the study only if all of the following criteria apply:

#### Age

1. Participant must be at least 18 (or the legal age of consent in the jurisdiction in which the study is taking place) years of age, inclusive, at the time of signing the informed consent.

#### Type of Participant and Disease Characteristics

2. Previously diagnosed moderate-to-severe OSA with an AHI  $\geq 15$ , as diagnosed with PSG, home sleep apnea test (HSAT), or other method that meets local guidelines prior to Visit 1. See Section 10.10 for definitions of apnea and hypopnea.
3. AHI  $\geq 15$  on PSG as part of the trial at Visit 1.
4. In the investigator's opinion, are well-motivated, capable, and willing to
  - learn how to self-inject study intervention, as required for this protocol (visually impaired persons who are not able to perform the injections must have the assistance of a sighted individual trained to inject study intervention; persons with physical limitations who are not able to perform the injections must have the assistance of an individual trained to inject study intervention)
  - inject study intervention (or receive an injection from a trained individual if visually impaired or with physical limitations), and
  - follow study procedures for the duration of the study, including, but not limited to: follow lifestyle advice (for example, dietary restrictions and exercise plan), maintain a study diary, and complete required questionnaires.

#### Weight

5. BMI  $\geq 30$  kg/m<sup>2</sup>.
6. Have a history of at least 1 self-reported unsuccessful dietary effort to lose body weight.

#### Sex and Contraceptive/Barrier Requirements

7. Males and females may participate in this trial.  
Female participants must not be pregnant, intending to be pregnant, breastfeeding, or intending to breastfeed.  
Contraceptive use by participants should be consistent with local regulations regarding the methods of contraception for those participating in clinical studies. For definitions and the contraception requirements of this protocol, see Appendix 4 (Section 10.4) .

#### Informed Consent

8. Capable of giving signed informed consent as described in Appendix 1 (Section 10.1), which includes compliance with the requirements and restrictions listed in the ICF and in this protocol.

## 5.2. Exclusion Criteria

Participants are excluded from the study if any of the following criteria apply:

### Medical Conditions

#### *Diabetes-related*

9. Have T1DM or T2DM, history of ketoacidosis, or hyperosmolar state/coma.
10. HbA1c  $\geq$  6.5% ( $\geq$  48 mmol/mol) at Visit 1.

#### *OSA-related*

11. Any previous or planned surgery for sleep apnea or major ear, nose or throat surgery, including tonsillectomy and adenoidectomy that still may affect breathing at time of Visit 1. Inclusion of a participant with more minor ear, nose or throat surgery (for example, deviated septum) will be at the investigator's discretion.
12. Significant craniofacial abnormalities that may affect breathing at time of Visit 1.
13. Diagnosis of Central or Mixed Sleep Apnea with % of mixed or central apneas/hypopneas  $\geq$ 50%, or diagnosis of Cheyne Stokes Respiration.
14. Diagnosis of Obesity Hypoventilation Syndrome or daytime hypercapnia.
15. Active device treatment of OSA other than PAP therapy (for example, dental appliance), or other treatment, that in the opinion of the investigator, may interfere with study outcomes, unless willing to stop treatment at Visit 1 and throughout the study.
16. Respiratory and neuromuscular diseases that could interfere with the results of the trial in the opinion of the investigator.

#### *Obesity-related*

17. Have a self-reported change in body weight  $>$ 5 kg within 3 months prior to screening.
18. Have a prior or planned surgical treatment for obesity (excluding liposuction or abdominoplasty if performed more than 1 year prior to screening).
19. Have or plan to have endoscopic and/or device-based therapy for obesity or have had device removal within the last 6 months (for example, mucosal ablation, gastric artery embolization, intragastric balloon, and duodenal-jejunal bypass sleeve).

#### *Other medical*

20. History of clinically relevant medical, behavioral, or psychiatric disorder, other than OSA, that is associated with insomnia or excessive sleepiness.
21. Impaired renal function, defined as eGFR  $<$ 30 mL/min/1.73 m<sup>2</sup>.
22. Have a known clinically significant gastric emptying abnormality (for example, severe gastroparesis or gastric outlet obstruction) or chronically take drugs that directly affect GI motility.
23. History of chronic or acute pancreatitis.
24. Thyroid-stimulating hormone outside of the range of 0.4 to 6.0 mIU/L at the screening visit.

**Note:** Participants receiving treatment for hypothyroidism may be included, provided their thyroid hormone replacement dose has been stable for at least 3 months.

**Note:** TSH values above the normal range can, in some patients, suggest subclinical hypothyroidism. If, in the investigator's opinion, the participant has subclinical



hypothyroidism and may require initiation of thyroid hormone replacement during the course of the study, the participant should be excluded from the study.

25. Have obesity induced by other endocrinologic disorders (for example, Cushing Syndrome) or diagnosed monogenetic or syndromic forms of obesity (for example, Melanocortin 4 Receptor deficiency or Prader-Willi Syndrome).
  26. Are, in the judgment of the investigator, actively suicidal and therefore deemed to be at significant risk for suicide.
  27. Have answered “yes” to either Question 4 or Question 5 on the “Suicidal Ideation” portion of the C-SSRS **or** have answered “yes” to any of the suicide-related behaviors on the “suicidal behavior” portion of the C-SSRS, **and** the ideation or behavior occurred within the past month.
  28. PHQ-9 score of 15 or more at Visit 1 or 2, prior to randomization.
  29. Uncontrolled hypertension (SBP  $\geq$ 160 mmHg and/or DBP  $\geq$ 100 mmHg) at Visit 1.
  30. Any of the following CV conditions less than 3 months prior to randomization: acute MI, cerebrovascular accident (stroke), unstable angina, or hospitalization due to congestive heart failure.
  31. History of (less than 3 months prior to Visit 1) or planned CV procedure.
  32. Heart failure, including New York Heart Association Functional Classification Class IV
  33. Have acute or chronic hepatitis, signs and symptoms of any other liver disease other than nonalcoholic fatty liver disease, or any of the following, as determined by the central laboratory during screening:
    - ALT level  $>$ 3.0X the ULN for the reference range
    - ALP level  $>$ 1.5X the ULN for the reference range, or
    - TBL level  $>$ 1.2X the ULN for the reference range (except for cases of known Gilbert’s Syndrome).
- Note:* Participants with nonalcoholic fatty liver disease are eligible to participate in this trial if their ALT level is  $\leq$ 3.0X the ULN for the reference range.
34. Have a calcitonin level (at Visit 1) of:
    - a.  $\geq$ 20 ng/L at Visit 1, if eGFR  $\geq$ 60 mL/min/1.73 m<sup>2</sup>
    - b.  $\geq$ 35 ng/L at Visit 1, if eGFR  $<$ 60 mL/min/1.73 m<sup>2</sup>
  35. Have a family or personal history of medullary thyroid carcinoma or multiple endocrine neoplasia syndrome type 2.
  36. Have a history of an active or untreated malignancy or are in remission from a clinically significant malignancy (other than basal- or squamous-cell skin cancer, in situ carcinomas of the cervix, or in situ prostate cancer) for less than 5 years.
  37. Have any other condition not listed in this section (for example, hypersensitivity or intolerance) that is a contraindication to GLP-1R agonists.
  38. Have a history of any other condition (such as known drug or alcohol abuse, diagnosed eating disorder, or other psychiatric disorder) that, in the opinion of the investigator, may preclude the participant from following and completing the protocol.
  39. Have history of use of marijuana less than 3 months of V1 and unwillingness to abstain from marijuana use during the trial. Participants should also refrain from use of cannabidiol oil for the duration of the study.

40. Have had a transplanted organ (corneal transplants [keratoplasty] allowed) or awaiting an organ transplant.
41. Requires the use of supplemental oxygen.

**Prior/Concomitant Therapy**

42. Are receiving or have received within 3 months prior to screening
  - a. chronic (>2 weeks or 14 days) systemic glucocorticoid therapy (excluding topical, intraocular, intranasal, intra-articular, or inhaled preparations) or
  - b. have evidence of a significant, active autoimmune abnormality (for example, lupus or rheumatoid arthritis) that has required (within the last 3 months) or
  - c. is likely to require, in the opinion of the investigator, concurrent treatment with systemic glucocorticoids (excluding topical, intraocular, intranasal, intra-articular or inhaled preparations) in the next 12 months.
43. Have current or history of (less than 3 months prior to Visit 1) treatment with medications that may cause significant weight gain, including but not limited to: tricyclic antidepressants, atypical antipsychotic and mood stabilizers, for example:
  - imipramine
  - amitriptyline
  - mirtazapine
  - paroxetine
  - phenelzine
  - chlorpromazine
  - thioridazine
  - clozapine
  - olanzapine
  - valproic acid and its derivatives
  - lithium

**Note:** Selective serotonin reuptake inhibitors are permitted, except for paroxetine.

44. Have taken, less than 3 months prior to Visit 1, medications (prescribed or over-the-counter) or alternative remedies intended to promote weight loss. Examples include, but are not limited to:
  - Saxenda® (liraglutide 3.0 mg)
  - Xenical®/Alli® (orlistat)
  - Meridia® (sibutramine)
  - Acutrim® (phenylpropanolamine)
  - Sanorex® (mazindol)
  - Adipex® (phentermine)
  - BELVIQ® (lorcaserin)
  - Qsymia® (phentermine/topiramate combination)
  - Contrave® (naltrexone/bupropion)
  - Pramlintide

- Zonisamide
- Topiramate
- Wegovy®

**Note:** Use of metformin or any other glucose-lowering medication, whether prescribed for polycystic ovary syndrome or diabetes prevention is not permitted.

45. Use of stimulants less than 3 months prior to Visit 1 (for example, modafinil, armodafinil, solriamfetol, pitolisant, amphetamine, dextroamphetamine, dexmethylphenidate, methylphenidate, and lisdexamfetamine).
46. Use of hypnotics, mirtazapine, opioids, trazodone less than 3 months prior to Visit 1.
47. Use of GLP-1 RA less than 3 months prior to Visit 1.
48. Use of any over-the-counter or prescription medications that could affect the evaluation of excessive sleepiness, per investigator discretion.
49. Use of any over-the-counter or prescription medications that could affect the evaluation of excessive sleepiness, per investigator discretion.
50. Unwillingness to discontinue over-the-counter (herbal or supplemental) medication that, in the opinion of the investigator, can interfere with the study.

#### **Prior/Concurrent Clinical Study Experience**

51. Are currently enrolled in any other clinical study involving an investigational product or any other type of medical research judged not to be scientifically or medically compatible with this study.
52. Previously randomly assigned to study intervention in this study or any other study investigating tirzepatide.
53. Have participated, within the last 30 days (4 months for studies conducted in Japan; 3 months for studies conducted in the United Kingdom), in a clinical trial involving a study intervention. If the previous study intervention is scientifically or medically incompatible with this study and has a long half-life, 3 months or 5 half-lives (whichever is longer) should have passed prior to screening (participation in observational studies may be permitted upon review of the observational study protocol and approval by the sponsor).
54. Are Lilly employees or employees of third-party organizations involved with the study.
55. Are investigator site personnel directly affiliated with this study and/or their immediate families. Immediate family is defined as a spouse, parent, child, or sibling, whether biological or legally adopted.
56. Have any medical condition that, in the opinion of the investigator, would be a contraindication to participation in the trial

### 5.3. Lifestyle Considerations

Per the SoA (Section 1.3), participants will consult with study personnel experienced in diet and exercise counseling to receive lifestyle program instructions at timepoints indicated in the SoA.

Diet and exercise goals and the importance of adherence to the lifestyle program will be reinforced at each trial contact by study staff.

#### 5.3.1. Meals and Dietary Restrictions

At Visit 2 and subsequent visits, participants will consult with study personnel experienced in diet and exercise counseling to receive lifestyle program instructions at timepoints indicated in the SoA. Dietary counseling will consist of advice on healthy food choices and focus on calorie restriction using a hypocaloric diet with macronutrient composition of:

- maximum 30% of energy from fat
- approximately 20% of energy from protein
- approximately 50% of energy from carbohydrates
- an energy deficit of approximately 500 kcal/day compared to the participant's estimated TEE

To encourage adherence, it is recommended that a 3-day diet and exercise diary be completed prior to each counseling visit. During each visit, the participant's diet is reviewed and advice to maximize adherence is provided if needed.

The hypocaloric diet is continued after randomization and throughout the treatment period. If a BMI  $\leq 22$  kg/m<sup>2</sup> is reached, the recommended energy intake should be recalculated with no kcal deficit for the remainder of the trial.

Total energy expenditure is calculated by multiplying the estimated Basal Metabolic Rate (BMR) (see table below) with a Physical Activity Level value of 1.3 (FAO/WHO/UNU 2004), which reflects an inactive lifestyle. This calculation provides a conservative estimate of caloric requirements:

$$\text{TEE (kcal/day)} = \text{BMR} \times 1.3$$

#### Equations for estimating BMR in kcal/day<sup>a</sup>

Sex	Age	BMR (kcal/day)
Men	18-30 years	15.057 X actual weight in kg + 692.2
	31-60 years	11.472 X actual weight in kg + 873.1
	>60 years	11.711 X actual weight in kg + 587.7
Women	18-30 years	14.818 X actual weight in kg + 486.6
	31-60 years	8.126 X actual weight in kg + 845.6
	>60 years	9.082 X actual weight in kg + 658.5

Abbreviations: BMR = basal metabolic rate; WHO = World Health Organization.

<sup>a</sup>Revised WHO equations (Adapted from: FAO/WHO/UNU 2004).

#### 5.3.2. Physical Activity

At Visit 2 and all subsequent visits, participants will be advised to increase their physical activity to at least 150 minutes per week.

#### **5.4. Screen Failures**

A screen failure occurs when a participant who consents to participate in the clinical study is not subsequently randomly assigned to study intervention in the study. A minimal set of screen failure information is required to ensure transparent reporting of screen failure participants to meet the Consolidated Standards of Reporting Trials (CONSORT) publishing requirements and to respond to queries from regulatory authorities. Minimal information includes demography, screen failure details, eligibility criteria, and any SAE.

Individuals who have history of marijuana use less than 3 months of Visit 1 may be rescreened once provided the individual is willing to abstain from marijuana use during the trial. Rescreened participants should be assigned a new participant number for every screening/rescreening event.

See Section 5.4 of the respective ISA's for additional rescreening guidelines.

#### **5.5. Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant**

Not applicable.

## 6. Study Intervention(s) and Concomitant Therapy

Study intervention is defined as any investigational intervention(s), marketed product(s), placebo, or medical device(s) intended to be administered to/used by a study participant according to the study protocol.

### 6.1. Study Interventions Administered

<b>Intervention Name</b>	Tirzepatide (LY3298176)	Placebo
<b>Type</b>	Drug	Drug
<b>Dose Formulation</b>	Single-dose pen	Single-dose pen
<b>Dosage Level(s)</b>	10 mg QW, 15 mg QW	Not applicable
<b>Route of Administration</b>	SC	SC
<b>Use</b>	Experimental	Placebo
<b>IMP and NIMP</b>	IMP	IMP
<b>Sourcing</b>	Provided centrally by the sponsor and dispensed via IWRS	
<b>Packaging and Labeling</b>	Study intervention will be provided in autoinjectors (single-dose pens), packaged in cartons to be dispensed. Clinical study materials will be labeled according to country regulatory requirements	

Abbreviations: IMP = investigational medicinal product; IWRS = interactive web-response system; NIMP = non-investigational medicinal product; QW = once-weekly; SC = subcutaneous.

There are no restrictions on the time of day each weekly dose of study intervention is given, but it is advisable to administer the SC injections on the same day and same time each week. The actual date, time, and injection details of all dose administrations will be recorded in the diary by the participant. If a dose of study intervention is missed, the participant should take it as soon as possible unless it is within 72 hours of the next dose, in which case that dose should be skipped, and the next dose should be taken at the appropriate time. The day of weekly administration can be changed if necessary, as long as the last dose was administered 72 or more hours before.

All participants will inject study intervention subcutaneously in the abdomen or thigh using the injection supplies provided; a caregiver may administer the injection in the participant's upper arm. A new autoinjector will be used for each injection. If study intervention is to always be injected in the same body region, participants should be advised to alternate injection sites each week.

### **6.1.1. Medical Devices**

The combination products provided for use in the study are tirzepatide autoinjector (or matching placebo). Any medical-device incidents, including those resulting from malfunctions of the device, must be detected, documented, and reported by the investigator throughout the study (see Section 10.3).

## **6.2. Preparation, Handling, Storage, and Accountability**

- The investigator or designee must confirm appropriate storage conditions have been maintained during transit for all study intervention received and any discrepancies are reported and resolved before use of the study intervention.
- Only participants enrolled in the study may receive study intervention. Only authorized study personnel may supply study intervention. All study intervention must be stored in a secure, environmentally controlled, and monitored (manual or automated) area in accordance with the labeled storage conditions with access limited to the investigator and authorized study personnel.
- The investigator or authorized study personnel are responsible for study intervention accountability, reconciliation, and record maintenance (that is, receipt, reconciliation, and final disposition records).
- Further guidance and information for the final disposition of unused study interventions are in the provided instructions.

## **6.3. Measures to Minimize Bias: Randomization and Blinding**

All participants will be centrally randomly assigned to study intervention using an IWRS. Before the study is initiated, the log in information and directions for the IWRS will be provided to each site.

Study intervention will be dispensed at the study visits summarized in SoA.

Returned study intervention should not be re-dispensed to the participants.

This is a double-blind study in which participants and study personnel, are blinded to study intervention. The IWRS will be programmed with blind-breaking instructions. In case of an emergency, the investigator has the sole responsibility for determining if unblinding of a participant's intervention assignment is warranted. Participant safety must always be the first consideration in making such a determination. If the investigator decides that unblinding is warranted, the investigator should make every effort to contact the sponsor prior to unblinding a participant's intervention assignment unless this could delay emergency treatment for the participant. If a participant's intervention assignment is unblinded, the sponsor must be notified immediately within 24 hours of this occurrence. The date and reason that the blind was broken must be recorded.

If an investigator, site personnel performing assessments, or participant is unblinded, the participant must be discontinued from the study. In cases where there are ethical reasons to have the participant remain in the study, the investigator must obtain specific approval from a sponsor clinical research physician (CRP) for the participant to continue in the study.

## Stratification

Participants will be stratified at randomization per ISA by country/geographic region, baseline AHI (moderate/severe), and gender.

### 6.4. Study Intervention Compliance

A record of the number of single-dose pens dispensed to and taken by each participant must be maintained and reconciled with study intervention and compliance records. Intervention start and stop dates, including dates for intervention delays and/or dose reductions will also be recorded.

Treatment compliance for each visit interval is defined as taking at least 75% of the required doses of study intervention. Similarly, a participant will be considered significantly noncompliant if he or she is judged by the investigator to have intentionally or repeatedly taken more than the prescribed amount of medication (more than 125%).

Participants considered to be poorly compliant with their medication and/or the study procedures will receive additional training and instruction, as required, and will be reminded of the importance of complying with the protocol.

### 6.5. Dose Modification

Dose modification is permitted for management of intolerable GI symptoms (Section 6.5.1).

Participants who do not tolerate at least 10 mg even after 1 de-escalation and re-escalation attempt, will be discontinued from the study intervention but remain in the study for continued follow-up.

Interventions to optimize study intervention tolerance and adherence may be employed throughout the study and include, but are not limited to, brief temporary interruptions (Section 7.1.2) and use of additional medications to manage GI symptoms (for example, nausea, vomiting, and diarrhea).

#### 6.5.1. Management of Participants with Gastrointestinal Symptoms

In participants who experience intolerable GI symptoms (for example, nausea, vomiting, or diarrhea) at any time during the study, the following measures are recommended:

- counselling on dietary behaviors that may help mitigate nausea and vomiting, (for example, eating smaller meals, splitting 3 daily meals into 4 or more smaller ones, and stopping eating when they feel full).
- if symptoms persist despite #1, prescribing symptomatic medication (for example, antiemetic or antidiarrheal medication), at the investigator's discretion.
- if symptoms persist despite #1 and #2, interrupting study intervention for 1 dose, provided the participant has taken the last 3 weekly doses. Study treatment should be resumed at the assigned dose immediately, either alone or in combination with symptomatic medication, which can also be utilized to manage symptoms (Section 6.8).

During the first 24 weeks of the treatment period (20-week dose escalation plus 4 weeks), participants unable to tolerate 2.5 mg or 5 mg (despite the interventions mentioned above) will be discontinued from the study intervention but remain in the study. For participants unable to tolerate any dose escalation between 7.5 mg and 15 mg inclusive, a dose de-escalation step with



subsequent re-escalation by 2.5 mg every 4 weeks up to MTD will be allowed in a blinded fashion, to reach either the 10-mg or 15-mg dose as described below. Only 1 cycle of dose de-escalation and re-escalation is permitted during the first 24 weeks of the treatment period. MTD is either 10 mg or 15 mg.

Participants who tolerate

- 10 mg, but do not tolerate 12.5 mg or 15 mg even following 1 de-escalation and re-escalation attempt, will continue on 10 mg as their MTD dose.
- 12.5 mg, but do not tolerate 15 mg even after 1 de-escalation and re-escalation attempt, will continue on 10 mg as their MTD dose.
- 15 mg will continue on 15 mg as their MTD dose.

## **6.6. Continued Access to Study Intervention after the End of the Study**

Tirzepatide will not be made available to participants after conclusion of the study.

## **6.7. Treatment of Overdose**

Study intervention overdose (more than the specified number of injections in less than 72 hours) will be reported as an AE.

In the event of an overdose, the treating physician should:

- Contact the medical monitor immediately.
- Evaluate the participant to determine, in consultation with the medical monitor, whether study intervention should be interrupted or whether the dose should be reduced.
- Closely monitor the participant for any AE/SAE and laboratory abnormalities until study intervention no longer has a clinical effect.

## **6.8. Concomitant Therapy**

Participants will be permitted to use concomitant medications that they require during the study, except certain medications (Section 6.8.2) that may interfere with the assessment of efficacy and safety characteristics of the study treatments.

Investigative-site staff will inform participants that they must consult with the investigator or a designated site staff member upon being prescribed any new medications during the study. This may not be possible when initiated for treatment of medical emergencies, in which case, the participant will inform the investigator or a designated site staff member as soon as possible.

Nonstudy medications taken by participants who are screened but not randomly assigned to study intervention will not be reported to Lilly unless an SAE or AE occurs that the investigator believes may have been caused by a study procedure.

Any medication or vaccine (including over-the-counter or prescription medicines, vitamins, and/or herbal supplements, or other specific categories of interest) that the participant is receiving at the time of screening or receives during the study must be recorded along with:

- Reason for use
- Dates of administration including start and end dates

The medical monitor should be contacted if there are any questions regarding concomitant or prior therapy.

### **6.8.1. Management of Incident Diabetes**

Incident diabetes is defined when any 1 of the following occur after randomization (American Diabetes Association 2020):

- unequivocal hyperglycemia (random glucose  $\geq 200$  mg/dL) with signs or symptoms of hyperglycemia
- within a 4-week period, any 2 of the following criteria are observed or 1 abnormal value is observed and confirmed:
  - HbA1c  $\geq 6.5\%$  (48 mmol/mol)
  - FG or 0-hour serum glucose from 2-hour OGTT  $\geq 126$  mg/dL (7.0 mmol/L)
  - 2-hour glucose  $\geq 200$  mg/dL (11.1 mmol/L) by a 2-hour OGTT
  - initiation of any medication for the treatment of diabetes

Participants who develop type 2 diabetes during the study will be

- provided with and trained to use a glucometer
- educated on the signs and symptoms of hypoglycemia and its treatment (see Section [10.3.7.1](#)), and
- provided a diary to record hypoglycemic episodes per Section [10.3.7.1](#).

Participants will be referred to their usual care provider and provided with a letter showing the study results indicative of diabetes. The decision to further evaluate, to initiate antihyperglycemic therapy, and the choice of antihyperglycemic medication will be at the discretion of the participant's usual care provider, with the exception of use of DPP-4 inhibitors and open-label GLP-1R agonists, which are prohibited in the study (Section [6.8.2](#)). Monitoring for hypoglycemia includes capture of events as defined in Section [10.3](#) (Appendix 3). Date of diabetes diagnosis will be captured in the AE CRF.

Initiation of metformin for the treatment of diabetes is permitted, but metformin should not be initiated during the study for the treatment of other metabolic conditions (for example, polycystic ovary syndrome and diabetes prevention).

### **6.8.2. Prohibited Concomitant Medications**

The following medications are prohibited during the study:

- DPP-4 inhibitors
- Open-label GLP-1R agonists
- Stimulants (for example, modafinil, armodafinil, solriamfetol, pitolisant, amphetamine, dextroamphetamine, dexmethylphenidate, methylphenidate, and lisdexamfetamine)

- medications that may cause significant weight gain (such as, but not limited to, paroxetine, tricyclic antidepressants, atypical antipsychotic and mood stabilizers).
- Medications that may cause weight loss (such as, but not limited to, liraglutide, semaglutide, orlistat, sibutramine, phenylpropanolamine, mazindol, phentermine, lorcaserin, naltrexone/bupropion, phentermine/topiramate combination, pramlintide, zonisamide, and topiramate)
- OTC, herbal, or supplemental medications that, in the opinion of the investigator, may interfere with the study
- hypnotics, mirtazapine, opioids, trazodone, pramlintide, sibutramine, orlistat, and zonisamide
- Systemic glucocorticoid therapy, per discussion with sponsor
- Use of any over-the-counter or prescription medications that could affect the evaluation of excessive sleepiness, per investigator discretion (such as, but not limited to, CBD oil, THC, etc.)

In addition,

- Active device treatment of OSA other than PAP therapy (for example, dental appliance), or other treatment that, in the opinion of the investigator, may interfere with study outcomes.

Participants who initiate and will not or cannot discontinue a prohibited medication, device, or other treatment will be permanently discontinued from study intervention per Section 7.1.

## 7. Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal

Discontinuation of specific sites or of the study as a whole are handled as part of Section 10.1, Appendix 1.

### 7.1. Discontinuation of Study Intervention

When necessary, a participant may be permanently discontinued from study intervention. If so, the participant will remain in the study and follow procedures for remaining study visits, as shown in the SoA.

A participant who prematurely discontinues study intervention is strongly encouraged to remain in the study for safety and efficacy assessments through the treatment period and post-treatment follow-up.

If a participant who discontinues the double-blind study treatment prematurely declines to complete the remaining scheduled study visits, then the participant should complete the early ED procedures indicated in the SoA. Participants should be encouraged to come back for the last treatment visit (Visit 11) to complete the end-of-treatment phase procedures and return for the post-treatment follow-up Visit 801.

Possible reasons leading to permanent discontinuation of study intervention:

- **participant decision**
  - the participant requests to discontinue study intervention
- **clinical considerations**
  - initiation of a prohibited medication (Section 6.8.2), if participants will not or cannot discontinue them
  - inadvertent enrollment if continued treatment with study intervention would not be medically appropriate
  - BMI  $\leq 18.5$  kg/m<sup>2</sup> is reached at any time during the treatment period, study intervention discontinuation should be considered
  - TEAE
    - intolerable GI symptoms despite management (Section 6.5.1)  
*Note:* The investigator should contact the Sponsor CRP to discuss whether it is medically appropriate for the participant to continue study treatment.
    - significant elevation of calcitonin

- If the investigator determines that a systemic hypersensitivity reaction has occurred related to study intervention administration, the participant may be permanently discontinued from the study intervention, and the sponsor's designated medical monitor should be notified. If the investigator is uncertain about whether a systemic hypersensitivity reaction has occurred and whether discontinuation of study intervention is warranted, the investigator may consult the sponsor
- occurrence of any other TEAE, SAE, or clinically significant finding for which the investigator believes that permanent study intervention discontinuation is the appropriate measure to be taken
- Diagnosis of
  - T1DM
  - Thyroid C-cell hyperplasia, MTC or MEN Syndrome type 2 after randomization
  - acute or chronic pancreatitis
  - an active or untreated malignancy (other than basal or squamous cell skin cancer, in situ carcinomas of the cervix, or in situ prostate cancer) after randomization
- onset of pregnancy in a female participant (see Sections 7.2 and 8.3.2)
- Suicidal Ideation and Behavior
  - PHQ-9 score  $\geq 15$ 
    - Participants should be referred to a Mental Health Professional (MHP) to assist in deciding whether the participant should be discontinued from study intervention. If a participant's psychiatric disorder can be adequately treated with psycho- and/or pharmacotherapy, then the participant, at the discretion of the Investigator (in agreement with the MHP), may be continued in the trial on randomly assigned therapy.
  - in addition, study intervention may be discontinued if participants:
    - answered "yes" to Question 4 or Question 5 on the "Suicidal Ideation" portion of the C-SSRS, **or**
    - answered "yes" to any of the suicide-related behaviors on the Suicidal Behavior portion of the C-SSRS.

A psychiatrist or appropriately trained professional may assist in the decision to discontinue the participant.

### 7.1.1. Liver Chemistry Stopping Criteria

The study intervention should be interrupted or discontinued if one or more of these conditions occur:

Elevation	Exception
ALT or AST >8x ULN	
ALT or AST >5x ULN for more than 2 weeks	
ALT or AST >3x ULN and either TBL >2x ULN or INR >1.5	In participants with Gilbert's syndrome, doubling of direct bilirubin should be used for drug interruption/ discontinuation decisions rather than TBL>2x ULN.
ALT or AST >3x ULN with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%)	
ALP >3x ULN, when the source of increased ALP is the liver	
ALP >2.5x ULN and TBL > 2x ULN	In participants with Gilbert's syndrome, doubling of direct bilirubin should be used for drug interruption/ discontinuation decisions rather than TBL>2x ULN.
ALP >2.5x ULN with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%)	

Source: FDA 2009 and other consensus guidelines, with minor modifications

Abbreviations: ALT = alanine aminotransferase; AST = aspartate aminotransferase; ALP = alkaline phosphatase; INR = international normalized ratio; TBL = total bilirubin level; ULN = upper limit of normal.

Resumption of the study intervention can be considered only in consultation with the Lilly-designated medical monitor and only if the liver test results return to baseline and if a self-limited non-drug etiology is identified.

Participants who are discontinued from study intervention due to a hepatic event or liver test abnormality should have additional hepatic safety data collected as described in Section 10.6, Appendix 6.

### 7.1.2. Temporary Discontinuation

In certain situations, after randomization, the investigator may need to temporarily interrupt study intervention. Every effort should be made by the investigator to maintain participants on study intervention and to restart study intervention after any temporary interruption, as soon as it is safe to do so. Distribution of study intervention at the correct dose will be per IWRS instructions.

Investigators should inform the sponsor that study intervention has been temporarily interrupted. The data related to temporary interruption of study treatment will be documented in source documents and entered on the CRF.

If study intervention interruption is...	then...
2 consecutive doses or less	participant restarts study intervention at last administered dose, as per escalation schedule.
3 consecutive doses or more	participant restarts study intervention (at 5 mg, managed by IWRS) and repeats dose escalation scheme.
Due to an AE	the event is to be documented and followed according to the procedures in Section 8.3 of this protocol.
Due to intolerable persistent GI AE	participants should be treated as suggested in Section 6.5.1.

Abbreviations: AE = adverse event; GI = gastrointestinal; IWRS = interactive web-response system.

## 7.2. Participant Discontinuation/Withdrawal from the Study

Discontinuation is expected to be uncommon.

A participant may withdraw from the study:

- at any time at the participant's own request
- at the request of the participant's designee (for example, parents or legal guardian)
- at the discretion of the investigator for safety, behavioral, compliance, or administrative reasons
  - If the participant becomes pregnant during the study (see Section 8.3.2. for additional details)
- if enrolled in any other clinical study involving an investigational product, or enrolled in any other type of medical research judged not to be scientifically or medically compatible with this study
- if the participant, for any reason, requires treatment with a therapeutic agent that is prohibited by the protocol and has been demonstrated to be effective for treatment of the study indication. In this case, discontinuation from the study occurs prior to introduction of the new agent.

At the time of discontinuing from the study, if possible, the participant will complete procedures for an ED visit and post-treatment follow-up, if applicable, as shown in the SoA. If the participant has not already discontinued the study intervention, the participant will be permanently discontinued from the study intervention at the time of the decision to discontinue the study.

If the participant withdraws consent for disclosure of future information, the sponsor may retain and continue to use any data collected before such a withdrawal of consent. If a participant withdraws from the study, the participant may request destruction of any samples taken and not tested, and the investigator must document this in the site study records.

**7.3. Lost to Follow up**

A participant will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site. Site personnel or designee are expected to make diligent attempts to contact participants who fail to return for a scheduled visit or were otherwise unable to be followed up by the site.



## **8. Study Assessments and Procedures**

Study procedures and their timing are summarized in the SoA.

Immediate safety concerns should be discussed with the sponsor immediately upon occurrence or awareness to determine if the participant should continue or discontinue study intervention.

Adherence to the study design requirements, including those specified in the SoA, is essential and required for study conduct.

All screening evaluations must be completed and reviewed to confirm that potential participants meet all eligibility criteria. The investigator will maintain a screening log to record details of all participants screened and to confirm eligibility or record reasons for screening failure, as applicable.

### **8.1. Efficacy Assessments**

#### **8.1.1. Primary Efficacy Assessment**

##### **8.1.1.1. Polysomnography**

The primary efficacy assessment in this study is AHI. AHI measurements will be collected via polysomnography.

Polysomnography assessments (including AHI, blood oxygen saturation parameters, PR, sleep parameters) will be performed during 1-night, overnight clinic stays, per the SoA. Data from the PSGs will be read and scored centrally using the AASM 1B hypopnea scoring method (when there is  $\geq 4\%$  oxygen desaturation from pre-event baseline; see Section 10.10, Appendix 10 for definitions) (Hamilton et al. 2021).

#### **8.1.2. Secondary Efficacy Assessments**

##### **8.1.2.1. Patient-Reported Outcomes**

###### **8.1.2.1.1. Functional Outcomes of Sleep Questionnaire (FOSQ)**

The FOSQ will be included to assess change in FOSQ domains and total score from baseline to Week 52. The FOSQ is a 30-item sleep-specific, participant-completed questionnaire used to assess the effect of disorders associated with excessive daytime sleepiness (EDS) on daily functioning in adults. It assesses the following 5 domains of

- General productivity (8 items)
- Activity level (9 items)
- Vigilance (7 items)
- Social outcomes (2 items)
- Intimate and sexual relationships (4 items)

The FOSQ items assess participant's current status with each item rated on a scale of 1 (extreme difficulty) to 4 (no difficulty), with an additional not applicable (0 = "I don't do this activity for other reasons") also available. Individual domain scores are calculated by taking the mean of answered, non-zero items within each domain and a total score can be calculated by first

computing the mean score for each domain, then multiplying the mean of the domain scores by 5 (Weaver et al. 1997). The total score for the FOSQ 10-item short form (FOSQ-10) can also be calculated.

#### ***8.1.2.1.2. PROMIS Short Form v1.0 Sleep-related Impairment 8a***

The PROMIS Short Form v1.0 Sleep-related Impairment 8a assesses self-reported perceptions of alertness, sleepiness, and tiredness during usual waking hours, and the perceived functional impairments associated with sleep problems or impaired alertness. The PROMIS Short Form v1.0 Sleep-related Impairment 8a consists of 8 items each rated on a 5-point scale ranging from “not at all” to “very much.” Items have a recall period of “in the past 7 days.” Individual item scores are totaled to obtain a raw score, with higher scores indicating more sleep-related impairment. Raw scores can be converted to a T-score, which is standardized with a mean of 50 and a SD of 10. (Northwestern, 2016a)

#### ***8.1.2.1.3. PROMIS Short Form v1.0 Sleep Disturbance 8b***

The PROMIS Short Form v1.0 Sleep Disturbance 8b assesses self-reported perceptions of sleep quality, sleep depth, and restoration associated with sleep, including perceived difficulties and concerns with getting to sleep or staying asleep, as well as perceptions of the adequacy of and satisfaction with sleep. The PROMIS Short Form v1.0 Sleep Disturbance 8b consists of 8 items each rated on a 5-point scale ranging from “not at all” to “very much,” “never” to “always,” or “very poor” to “very good.” Items have a recall period of “in the past 7 days.” Individual item scores are totaled to obtain a raw score, with higher scores indicating more sleep disturbance. Raw scores can be converted to a T-score, which is standardized with a mean of 50 and a SD of 10. (Northwestern, 2016b)

#### ***8.1.2.1.4. Epworth Sleepiness Scale***

The ESS will be included to assess improvements in excessive daytime sleepiness from baseline to Week 52. The ESS is an 8-item participant-completed measure that asks the participant to rate on a scale of 0 (would never doze) to 3 (high chance of dozing), their usual chances of dozing in 8 different daytime situations, with a recall period of “in recent times.” The ESS total score is the sum of the 8-item scores and ranges from 0 to 24, with higher scores indicating greater daytime sleepiness (Johns 1991).

#### ***8.1.2.1.5. Short-Form 36 Version 2 Health Survey, Acute Form, 1-week Recall Version***

The SF-36v2 will be included to assess health-related quality of life from baseline to Week 52. The SF-36v2 acute form, 1-week recall version is a 36-item generic, participant-completed measure designed to assess the following 8 domains over “the past week.”

- Physical functioning
- Role-physical
- Bodily pain
- General health
- Vitality
- Social functioning
- Role-emotional, and
- Mental health

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

#### **8.1.2.1.6. *Patient Global Impression of Status – Obstructive Sleep Apnea (PGIS-OSA) Symptoms Scales***

Three patient global impression of status scales will be included to assess categorical shift in participant self-rated assessment of their OSA symptom severity from baseline to Week 52.

##### **PGIS-OSA Fatigue**

This is a single-item, participant self-rated assessment of their overall level of fatigue due to OSA, “over the past 7 days.” The item is rated on a 4-point scale ranging from “No fatigue” to “Severe fatigue.”

##### **PGIS-OSA Sleepiness**

This is a single-item, participant self-rated assessment of their overall level of sleepiness due to OSA during waking hours, “over the past 7 days.” The item is rated on a 4-point scale ranging from “Not at all sleepy” to “Very sleepy.”

##### **PGIS-OSA Snoring**

The PGIS-OSA Snoring scale consists of two items. The first item is a participant self-rated assessment of their overall perception of the severity of their snoring due to OSA, “over the past 7 days,” with respect to how much their snoring has affected their sleep. The item is rated on a 4-point scale ranging from “Not at all affected” to “Very affected.” For the second item, participants will be asked on a 3-point scale (“Not at all” to “All the time”) if they have ever been told by someone else that they snore in their sleep.

### **8.1.3. Exploratory Efficacy Assessments**

#### **8.1.3.1. Actigraphy**

The actigraphy device (AX6) will be utilized in the OSA trial to objectively evaluate the effect of tirzepatide on various sleep and physical activity parameters including but not limited to change of daytime sleep time, sleep efficiency, and change in daytime physical activity from baseline. The actigraphy device is a data logger capable of recording raw data from a suite of integrated sensors and can be configured to collect movement-relevant data in an uninterrupted fashion for up to 2 months, thus is ideal for collecting longitudinal movement data (for example, physical activity, sleep, etc.) in real-world health research and well-defined clinical trials. The actigraphy device meets the CE (Conformite Europeenne) mark requirements. Participants should wear the Actigraphy device for 7 consecutive days, per the SoA.

Lack of participation in actigraphy collections at any time is not considered a protocol deviation.

### **8.1.3.2. Patient Global Impression of Change – Obstructive Sleep Apnea (PGIC-OSA) Symptoms Scales**

Four patient global impression of change scales will be included to assess categorical shift in participant self-rated assessment of change in their OSA symptom severity from baseline to Week 52.

#### **PGIC-OSA Fatigue**

This is a single-item, participant self-rated assessment of the change in their overall level of fatigue due to OSA, “since you started taking the study medication.” The item is rated on a 5-point scale ranging from “Much worse” to “Much better.”

#### **PGIC-OSA Sleepiness**

This is a single-item, participant self-rated assessment of the change in their overall level of sleepiness due to OSA during waking hours, “since you started taking the study medication.” The item is rated on a 5-point scale ranging from “Much more sleepy” to “Much less sleepy.”

#### **PGIC-OSA Sleep Quality**

This is a single-item, participant self-rated assessment of the change in their overall sleep quality due to OSA, “since you started taking the study medication.” The item is rated on a 5-point scale ranging from “Much worse” to “Much better.”

#### **PGIC-OSA Snoring**

This is a single-item, participant self-rated assessment of the overall change in how their snoring has affected their sleep, “since you started taking the study medication.” The item is rated on a 5-point scale ranging from “My sleep is much more affected” to “My sleep is much less affected.”

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

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

- mobility
- selfcare
- usual activities
- pain/discomfort, and
- anxiety/depression.

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

- no problems
- slight problems
- moderate problems
- severe problems, and
- unable to perform/extreme problems.

A total of 3125 health states is possible. In addition to the health profile, a single health state index value can be derived based on a formula that attaches weights to each of the levels in each

dimension. This index value ranges between less than 0 (where 0 is a health state equivalent to death; negative values are valued as worse than dead) to 1 (perfect health). In addition, the EQ Visual Analog Scale records the respondent's self-rated health status on a vertical graduated (0 to 100) visual analog scale. The participant rates his/her perceived health from 0 (the worst imaginable health) to 100 (the best imaginable health). In conjunction with the health state data, it provides a composite picture of the respondent's health status.

## **8.2. Safety Assessments**

Planned time points for all safety assessments are provided in the SoA.

### **8.2.1. Physical Examinations**

A complete physical examination will include, at a minimum, assessments of the CV, respiratory, gastrointestinal and neurological systems as well as a thyroid examination. Height, weight, and waist circumference will also be measured and recorded.

A symptom directed physical assessment will be performed at visits indicated in the SoA, as clinically indicated.

Investigators should pay special attention to clinical signs related to previous serious illnesses.

### **8.2.2. Vital Signs**

For each participant, vital signs measurements should be conducted according to the SoA (Section 1.3).

### **8.2.3. Electrocardiograms (ECG)**

Single 12-lead ECG will be obtained as outlined in the SoA (see Section 1.3) using an ECG machine that automatically calculates the heart rate and measures PR, QRS, QT, and QTc intervals.

A local single 12-lead ECG will be collected at designated visits. ECGs should be collected at least 30 minutes prior to collection of blood samples for laboratory testing, including PK samples. Participants should be supine for approximately 5 to 10 minutes before ECG collections and remain supine but awake during the ECG collection. ECGs may be repeated at the investigator's discretion at any visit.

After enrollment, if a clinically significant increase in the QT/QTc interval from baseline or other clinically significant quantitative or qualitative change from baseline is identified, the participant will be assessed by the investigator for symptoms (for example, palpitations, near syncope, and syncope) and to determine whether the participant can continue in the study. The investigator or qualified designee is responsible for determining if any change in participant management is needed and must document his/her review of the ECG printed at the time of evaluation from at least 1 of the replicate ECGs from each time point.

### **8.2.4. Clinical Safety Laboratory Tests**

See Section 10.2, Appendix 2 for the list of clinical laboratory tests to be performed and the SoA for the timing and frequency.

The investigator must review the laboratory results, document this review, and report any clinically relevant changes occurring during the study as an AE. The laboratory results must be retained with source documents unless a Source Document Agreement or comparable document cites an electronic location that accommodates the expected retention duration. Clinically significant abnormal laboratory findings are those which are not associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant's condition.

All laboratory tests with values considered clinically significantly abnormal during participation in the study should be repeated until the values return to normal or baseline or are no longer considered clinically significant by the investigator or medical monitor.

- If such values do not return to normal/baseline within a period of time judged reasonable by the investigator, the etiology should be identified and the sponsor notified.
- All protocol-required laboratory assessments, as defined in Section 10.2, Appendix 2 must be conducted in accordance with the SoA, standard collection requirements, and laboratory manual.

If laboratory values from non-protocol specified laboratory assessments performed at an investigator-designated local laboratory require a change in participant management or are considered clinically significant by the investigator (for example, SAE or AE or dose modification), then report the information as an AE.

#### **8.2.5. Pregnancy Testing**

See the SoA (Section 1.3) for testing for pregnancy and timepoints.

See Appendix 4 (Section 10.4) for additional information for pregnancy.

#### **8.2.6. Suicidal Ideation and Behavior Risk Monitoring**

Patients with obesity may occasionally develop suicidal ideation or behavior.

Participants should be monitored appropriately and observed closely for SIB or any other unusual changes in behavior, especially at the beginning and end of the course of intervention, or at the time of dose changes, either increases or decreases. Participants who experience signs of SIB should undergo a risk assessment. All factors contributing to SIB should be evaluated and consideration should be given to discontinuation of the study intervention.

Baseline assessment of suicidal ideation and behavior/intervention emergent suicidal ideation and behavior will be monitored using the C-SSRS and PHQ-9.

##### **8.2.6.1. C-SSRS**

Columbia Suicide-Severity Rating Scale (C-SSRS) is a scale that captures the occurrence, severity, and frequency of suicidal ideation and behavior during the assessment period via a questionnaire. The scale was developed by the National Institute of Mental Health (NIMH) trial group (TASA) for the purpose of being counterpart to the Columbia Classification Algorithm of Suicide Assessment (C-CASA) categorization of suicidal events. (The Columbia Lighthouse Project, 2013)

For this study, the C-SSRS is adapted for the assessment of the ideation and behavior categories only. The Intensity of Ideation and Lethality of Behavior sections are removed.

### 8.2.6.2. PHQ-9

The PHQ-9 is a validated self-report screening tool that assesses the presence and intensity of depressive symptoms. The PHQ-9, which incorporates the 9 Diagnostic and Statistical Manual-IV depression criteria as “0” (not at all) to “3” (nearly every day), was developed for use in primary care settings (Kroenke et al. 2001).

## 8.3. Adverse Events, Serious Adverse Events, and Product Complaints

The definitions of the following events can be found in Section 10.3, Appendix 3:

- AEs
- SAEs
- PCs

These events will be reported by the participant (or, when appropriate, by a caregiver, surrogate, or the participant’s legally authorized representative).

The investigator and any qualified designees are responsible for detecting, documenting, and recording events that meet these definitions and remain responsible for following up events that are serious, considered related to the study intervention or study procedures, or that caused the participant to discontinue the study intervention before completing the study (see Section 7).

Care will be taken not to introduce bias when detecting events. Open-ended and non-leading verbal questioning of the participant is the preferred method to inquire about event occurrences.

After the initial report, the investigator is required to proactively follow each participant at subsequent visits/contacts. All SAEs and AEs of special interest (as defined in Section 8.3.3) will be followed until resolution, stabilization, the event is otherwise explained, or the participant is lost to follow-up (as defined in Section 7.3). For PCs, the investigator is responsible for ensuring that follow-up includes any supplemental investigations as indicated to elucidate the nature and/or causality. Further information on follow-up procedures is provided in Section 10.3, Appendix 3.

### 8.3.1. Timing and Mechanism for Collecting Events

This table describes the timing, deadlines, and mechanism for collecting events.

Event	Collection Start	Collection Stop	Timing for Reporting to Sponsor or Designee	Mechanism for Reporting	Back-Up Method of Reporting
<b>AE</b>					
AE	Signing of the ICF	participation in study has ended	As soon as possible upon site awareness	AE CRF	N/A

Event	Collection Start	Collection Stop	Timing for Reporting to Sponsor or Designee	Mechanism for Reporting	Back-Up Method of Reporting
<b>SAE</b>					
SAE and SAE updates – prior to start of study intervention <b>and</b> deemed reasonably possibly related to study procedures	Signing of the ICF	start of intervention	Within 24 hours of awareness	SAE CRF	SAE paper form
SAE and SAE updates – after start of study intervention	Start of intervention	participation in study has ended	Within 24 hours of awareness	SAE CRF	SAE paper form
SAE <sup>a</sup> – after participant’s study participation has ended <b>and</b> the investigator becomes aware	After participant’s study participation has ended	N/A	Promptly	SAE paper form	N/A
<b>Pregnancy</b>					
Pregnancy in female participants and female partners of male participants	After the start of study intervention	until 1 month after the last dose of study intervention	Within 24 hours (see Section 8.3.2)	Pregnancy paper form	Pregnancy paper form
<b>PCs</b>					
PC associated with an SAE or might have led to an SAE	Start of study intervention	End of study intervention	Within 24 hours of awareness	PC form	N/A



Event	Collection Start	Collection Stop	Timing for Reporting to Sponsor or Designee	Mechanism for Reporting	Back-Up Method of Reporting
PC not associated with an SAE	Start of study intervention	End of study intervention	Within 1 business day of awareness	PC form	N/A
Updated PC information	—	—	As soon as possible upon site awareness	Originally completed PC form with all changes signed and dated by the investigator	N/A
PC (if investigator becomes aware)	Participation in study has ended	N/A	Promptly	PC form	

Abbreviations: AE = adverse event; CRF = case report form; ICF = informed consent form; N/A = not applicable; PC = product complaint; SAE = serious adverse event.

<sup>a</sup> Serious adverse events should not be reported unless the investigator deems them to be possibly related to study treatment or study participation.

### 8.3.1.1. Adverse Event Monitoring with a Systematic Questionnaire

Nonleading AE collection should occur prior to the collection of the C-SSRS and PHQ-9.

If a suicide-related event is discovered *during the C-SSRS or PHQ-9* but was not captured during the nonleading AE collection, sites should not change the AE form.

If an AE is serious or leads to discontinuation, it needs to be included on the AE form and the process for reporting SAEs is followed.

### 8.3.2. Pregnancy

#### Collection of pregnancy information

##### *Male participants with partners who become pregnant*

- The investigator will attempt to collect pregnancy information on any male participant's female partner who becomes pregnant while the male participant is in this study. This applies only to male participants who receive study intervention.
- After learning of a pregnancy in the female partner of a study participant, the investigator will
  - obtain a consent to release information from the pregnant female partner directly, and
  - within 24 hours after obtaining this consent will record pregnancy information on the appropriate form and submit it to the sponsor.

- The female partner will also be followed to determine the outcome of the pregnancy. Information on the status of the mother and child will be forwarded to the sponsor. Generally, the follow-up will be no longer than 6 to 8 weeks following the estimated delivery date. Any termination of the pregnancy will be reported regardless of gestational age, fetal status (presence or absence of anomalies) or indication for the procedure.

#### *Female participants who become pregnant*

- The investigator will collect pregnancy information on any female participant who becomes pregnant while participating in this study. The initial information will be recorded on the appropriate form and submitted to the sponsor within 24 hours of learning of a participant's pregnancy.
- The participant will be followed to determine the outcome of the pregnancy. The investigator will collect follow-up information on the participant and the neonate and the information will be forwarded to the sponsor. Generally, follow-up will not be required for longer than 6 to 8 weeks beyond the estimated delivery date. Any termination of pregnancy will be reported, regardless of gestational age, fetal status (presence or absence of anomalies) or indication for the procedure.
- While pregnancy itself is not considered to be an AE or SAE, any pregnancy complication or elective termination of a pregnancy for medical reasons will be reported as an AE or SAE.
- A spontaneous abortion (occurring at <20 weeks gestational age) or still birth (occurring at ≥20 weeks gestational age) is always considered to be an SAE and will be reported as such.
- Any post-study pregnancy related SAE considered reasonably related to the study intervention by the investigator will be reported to the sponsor as described in protocol Section 8.3.1. While the investigator is not obligated to actively seek this information in former study participants, he or she may learn of an SAE through spontaneous reporting.
- Any female participant who becomes pregnant while participating in the study will discontinue study intervention and be withdrawn from the study. If the participant is discontinued from the study, follow the standard discontinuation process and continue directly to the study follow-up phase. The follow-up on the pregnancy outcome should continue independent of intervention or study discontinuation.

### **8.3.3. Adverse Events of Special Interest**

Adverse events of special interest are prospectively defined to include

- severe hypoglycemia
- MACE (adjudicated); includes, but not limited to CV death, nonfatal MI, nonfatal stroke, hospitalization for unstable angina, and hospitalization for heart failure
- treatment-emergent arrhythmias and cardiac conduction disorders
- hepatobiliary disorders; includes biliary colic, cholecystitis, and other gallbladder disease
- severe GI events
- acute renal events

- Major Depressive Disorder/suicidal behavior and ideation
- pancreatitis (adjudicated)
- c-cell hyperplasia and thyroid malignancies
- allergic/hypersensitivity reactions; includes ISRs and ADA formation

If these events are reported, investigators may be prompted to collect additional data about the event. Sections 10.1.5.1 and 10.1.5.2 (Appendix 1), outline additional information on CV and pancreatic adjudication. Section 10.3.7 (Appendix 3) outline additional information on hypoglycemia, hypersensitivity reactions, and ISRs.

## 8.4. Pharmacokinetics

Blood samples will be obtained from all participants enrolled in the 2 ISAs to enable the characterization of tirzepatide PK and exposure response relationships, as permissible. Samples will be collected with concurrent immunogenicity samples at timepoints indicated in the SoA (Section 1.3).

Blood samples collected from participants in the placebo arms will not be included in the bioanalyses of drug concentrations.

- Plasma samples will be collected for measurement of plasma concentrations of tirzepatide as specified in the SoA
- Instructions for the collection and handling of biological samples will be provided by the sponsor. The actual date and time (24-hour clock time) of each sample will be recorded.

Drug concentration information that may unblind the study will not be reported to investigative sites or blinded personnel.

### 8.4.1. Bioanalytical Methods

Samples will be analyzed at a laboratory approved by the sponsor, and stored at a facility designated by the sponsor. Concentrations of tirzepatide will be assayed using a validated liquid chromatography mass spectrometry method. Analyses of samples collected from placebo-treated participants are not planned. Bioanalytical samples collected to measure tirzepatide concentrations will be retained for a maximum of 1 year following last participant visit for the study. During this time, samples remaining after the bioanalyses may be used for exploratory analyses such as metabolism work, protein binding, and/or bioanalytical method cross-validation.

## 8.5. Pharmacodynamics

Samples to assess the PD properties of tirzepatide are included in the efficacy measures and not applicable in this section.

## 8.6. Genetics

A whole blood sample will be collected for pharmacogenetic analysis where local regulations allow.

## 8.7. Biomarkers

Blood samples will be collected to address questions of relevance to drug disposition, target engagement, PD, mechanism of action, variability of study participant response (including safety), and clinical outcome. Biomarkers will include measurement of biomolecules including proteins, lipids, and other cellular elements.

Samples will be collected according to the schedule described in the SoA and as detailed in the laboratory manual provided separately to sites.

All samples will be coded with the study participant number. These samples and any data generated can be linked back to the study participant only by the investigator site personnel.

Lilly may store samples after the end of the study to achieve study objectives. Additionally, with participant's consent, samples may be used for further research by Lilly or others such as universities or other companies to contribute to the understanding of OSA, obesity, or other diseases, the development of related or new treatments, or research methods.

See Section 10.1.12, Appendix 1 for details related to sample retention.

## 8.8. Immunogenicity Assessments

At the visits and times specified in the SoA (Section 1.3), venous blood samples will be collected for analysis to determine antibody production against tirzepatide. Antibodies may be further characterized for: cross-reactive binding to native GIP and GLP-1, neutralizing activity of tirzepatide on the GIP and GLP-1 receptors, and neutralizing activity to native GIP and/or GLP-1. To interpret the results of immunogenicity, a venous blood sample will be collected at the same time points to determine the serum concentrations of tirzepatide. All samples for immunogenicity should be taken predose when applicable and possible.

Treatment-emergent ADAs are defined in Section 10.3, Appendix 3.

Immunogenicity will be assessed by a validated assay designed to detect ADAs in the presence of tirzepatide at a laboratory approved by the sponsor. The purpose of retention, the maximum duration of retention, and facility for long-term storage of samples is described in Section 10.2, Appendix 2. Samples may also be used for development and control of an immunogenicity assay.

For details related to sample retention, see Section 10.1.12, Appendix 1.

## 8.9. Health Economics OR Medical Resource Utilization and Health Economics

Health economics or medical resource utilization and health economics parameters are not evaluated in this study.

## **9. Statistical Considerations**

This section is a summary of the planned statistical analyses of the most important endpoints, including primary and key secondary endpoints. The SAP will include more technical and detailed description of the statistical analyses described in this section.

Unblinding details will be specified in the unblinding plan section of the SAP or in a separate unblinding plan document.

### **9.1. Statistical Hypotheses**

For each ISA, the primary objective is to demonstrate that tirzepatide at the MTD (10 mg or 15 mg) is superior to Placebo in treating participants with OSA with respect to AHI endpoint. Thus the null and alternative hypotheses will be defined as below.

Null hypothesis: tirzepatide at the MTD (10 mg or 15 mg) is not different from the placebo with respect to the mean percent change from baseline in AHI at 52 weeks.

Alternative hypothesis: tirzepatide at the MTD (10 mg or 15 mg) is superior to the placebo with respect to the mean percent change from baseline in AHI at 52 weeks.

The treatment effect will be defined as the difference between the estimates of the mean percent change from baseline at 52 weeks for tirzepatide at the MTD (10 mg or 15 mg) and placebo.

#### **9.1.1. Multiplicity Adjustment**

The statistical comparisons for the primary efficacy endpoint and the key secondary endpoints will be carried out based on a graphical approach for multiple comparisons within each ISA (Bretz et al. 2011). The graphical approach is a closed testing procedure; hence, it strongly controls the familywise Type I error rate (2-sided alpha level of 0.05) across the primary and key secondary objectives (Alosh et al. 2014). Details about the graphical approach will be described in the SAP.

### **9.2. Analyses Sets**

This table describes the populations that will be used for statistical analyses within each ISA of the master protocol. Additional intervention-specific populations for analyses may be described in the respective ISA.

<b>Analysis Set or Population</b>	<b>Description</b>
<b>mITT population</b>	All randomly assigned participants who are exposed to at least 1 dose of study intervention.
<b>FAS</b>	Data obtained during treatment period from the mITT population, regardless of adherence to study intervention
<b>EAS</b>	Data obtained during treatment period from the mITT population, excluding data after discontinuation of study intervention
<b>SS</b>	Data obtained during treatment period plus safety follow-up period from mITT population, regardless of adherence to study intervention

Abbreviations: EAS = efficacy analysis set; FAS = full analysis set; mITT = modified intent-to-treat; SS = safety analysis set.

### **9.3. Statistical Analyses**

#### **9.3.1. General Considerations**

Statistical analysis will be the responsibility of the sponsor or its designee. Statistical analysis for each ISA will be conducted individually and a combined analysis with both ISAs is not planned.

The SAP will be finalized prior to the unblinding of the first ISA.

Changes to the data analysis methods will require an amendment only if a principal feature of the master protocol is changed. Any other change to the data analysis methods, and the justification for making the change, will be described in the SAP or the CSR for each respective ISA.

Additional exploratory analyses of the data will be conducted as deemed appropriate.

Efficacy analyses will be conducted on all participants randomly assigned to study intervention according to the treatment to which the participants are assigned. The primary and key secondary efficacy analysis will be guided by 2 estimands, the “hybrid” estimand and the “efficacy” estimand to support global regulatory submissions and publications. For the “hybrid” estimand, the analysis will be conducted using FAS. To minimize missing data, participants randomly assigned to study intervention who prematurely discontinue study treatment will be encouraged to remain in the study, however, some participants may choose to permanently discontinue from the study which will lead to missing endpoints. Details on handling missing values can be found in Missing Value Imputation section. For the “efficacy” estimand, the analysis will include data collected prior to permanent discontinuation of study intervention and will be conducted using the EAS.

Safety analysis will be conducted using the SS. Selected safety analyses may be conducted after excluding the data after permanent discontinuation of the study intervention.

### Missing Value Imputation

Missing values for the primary and multiplicity adjusted endpoints at Week 52 will be handled as follows. For efficacy analysis relative to the “hybrid” estimand, missing values will be imputed using multiple imputation based on the reason of intercurrent events. The statistical inference over multiple imputations will be guided by the method proposed by Rubin (1987). The intercurrent events and the resulting missing values will be handled as follows:

<b>Intercurrent events</b>	<b>Strategy to handle intercurrent events</b>	<b>Assumptions for missing values</b>	<b>Methods to handle missing values</b>
Treatment DC due to the COVID-19 pandemic (after other reasons for missing data are ruled out)	Hypothetical	MAR	Use observed data on treatment within each treatment to impute missing values
All other treatment discontinuations	Treatment policy	MNAR	Multiple imputation

Abbreviations: COVID-19 = coronavirus disease-2019; DC = discontinuation; MAR = missing at random; MNAR = missing not at random.

Unless otherwise noted, all tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, and the confidence interval will be calculated at 95%, 2-sided. In statistical summaries and analyses, participants will be analyzed as randomized. Countries in similar geographic regions with fewer than 10 participants, based on the all-randomized population, will be pooled to achieve a pooled country of at least 10 participants. All analyses using country in the model will use a pooled country, unless otherwise specified. The final pooling by country and geographic region will be finalized prior to data lock.

Baseline is defined as the last non missing measurement at or before the randomization visit (Visit 2) unless otherwise specified.

Analysis of covariance will be used to analyze continuous variables collected only at baseline and endpoint. The model will include treatment and strata (pooled country/geographic region, AHI stratum [moderate (AHI  $\geq 15$  and AHI  $< 30$ ), severe (AHI  $\geq 30$ )] and gender) as fixed effects and baseline as a covariate. The ANCOVA model for AHI analysis will include baseline AHI instead of the AHI stratum as a fixed covariate.

The MMRM analysis, a restricted-maximum-likelihood-based model, will be used to analyze continuous longitudinal variables. All the longitudinal observations at each scheduled postbaseline visit will be included in the analysis. The model will include the fixed class effects of treatment, strata (pooled country/geographic region and gender), visit, and treatment-by-visit interaction, as well as the continuous, fixed covariate of baseline value. For analyses of variables other than AHI, the AHI stratum will also be included in the model. Significance tests will be based on least-squares means and Type III tests.

For continuous measures, summary statistics may include sample size, mean, standard deviation, median, minimum, and maximum for both the actual and the change from baseline

measurements. Least-square means and standard errors derived from the analysis models will also be displayed for the change from baseline measurements. Treatment comparisons will be displayed showing the treatment difference least-square means and the 95% confidence intervals for the treatment differences, along with the p-values for the treatment comparisons.

For categorical measures, summary statistics may include sample size, frequency, and percentages. Fisher's exact test or Pearson's chi-square test will be used for treatment comparisons unless otherwise specified.

Handling of missing, unused, and spurious data are addressed prospectively in the overall statistical methods described in the protocol and in the SAP, where appropriate. Adjustments to the planned analyses are described in the final CSR.

### **9.3.2. Primary Analysis**

The primary objective of this study is to test the hypothesis that tirzepatide at the MTD (10 mg or 15 mg) is superior to placebo for participants with moderate to severe OSA on the mean percent AHI reduction from baseline to Week 52. There will be 2 primary analysis methods, each tested at the full significance level of 0.05.

The primary analysis guided by the "hybrid" estimand will use FAS. Missing values will be imputed following imputation methods described in Section 9.3.1, and more details about the imputation methods will be provided in the SAP. After the imputation, the primary efficacy comparison will be based on the contrast between tirzepatide at the MTD (10 mg or 15 mg) and placebo from the ANCOVA analysis of mean percent change from baseline to Week 52 in AHI using FAS as described in Section 9.3.1.

The primary analysis guided by the "efficacy" estimand will use EAS. The efficacy comparison will be based on the contrast between tirzepatide at the MTD (10 mg or 15 mg) and placebo at Week 52 (Visit 11) from the MMRM analysis of mean percent change from baseline in AHI using EAS as described in Section 9.3.1. More details about the model will be provided in the SAP.

### **9.3.3. Analysis of Key Secondary Endpoints**

The analyses for key secondary endpoints as in Section 3 will be performed for both the "hybrid" and "efficacy" estimand as described in Section 9.3.1 using the graphical testing scheme. The details of graphical testing scheme will be described in the SAP.

The analysis of the hierarchical combination PRO endpoint will be performed with the Finkelstein-Schoenfeld method, and the win ratio (Pocock et al. 2012) will be reported as the measure of treatment effect.

The last measurement prior to randomization for all three PRO endpoints will be used as baseline. For the "hybrid" estimand, missing values at Week 52 will be imputed through multiple imputations based on the reason of missingness as described in Section 9.3.1.

Analysis of percent change from baseline in body weight, change in SBP, and CRP at the 52-week visit will be conducted in a manner similar to the primary efficacy analyses with baseline AHI stratum added in the model, and baseline of the corresponding variable as a covariate.



Comparisons at the 52-week visit between the treatments relative to the proportion of participants achieving  $\geq 50\%$  AHI reduction and  $AHI < 5$  or ( $AHI \geq 5$  and  $AHI \leq 14$  and  $ESS \leq 10$ ) will be conducted using logistic regression analysis including terms for treatment, pooled country, gender, and baseline AHI as a covariate.

### **9.3.4. Treatment Group Comparability**

#### **9.3.4.1. Participant Disposition**

Participants who discontinue from the study will be identified, and the extent of their participation in the study will be reported for each ISA. A detailed description of participant disposition and the reasons for discontinuation will be summarized by treatment group for each ISA at the end of the study. Intervention-specific analyses will be detailed in each respective ISA.

#### **9.3.4.2. Participant Characteristics**

Participant characteristics and baseline clinical measures will be summarized for each treatment. For all participant characteristics, the summaries will include descriptive statistics for continuous measures (for example, sample size, mean, standard deviation, median, minimum, and maximum) and for categorical measures (for example, sample size, counts and percentages). Additional intervention-specific analyses will be specified in each respective ISA.

#### **9.3.4.3. Concomitant Therapy**

Concomitant medications used during the study will be summarized for each ISA.

### **9.3.5. Safety Analyses**

Unless specified otherwise, safety assessments will compare safety of tirzepatide at the MTD (10 mg or 15 mg) with placebo irrespective of adherence to study intervention. Thus, safety analyses will be conducted using the SS.

#### **9.3.5.1. Adverse Events**

Adverse events will be classified by system organ class and preferred term as defined by the Medical Dictionary for Regulatory Activities.

All conditions existing prior to randomization at Visit 2 will be used as baseline. The postbaseline visits during the placebo-controlled phase will be included as the postbaseline period for analysis.

For events that are gender specific, the denominator and computation of the percentage will only include participants of the given gender.

#### **9.3.5.2. Hypoglycemic Events**

Incidence of documented symptomatic hypoglycemia events and severe hypoglycemia will be summarized and compared between tirzepatide at the MTD (10 mg or 15 mg) and placebo. Rate of hypoglycemic episodes will also be analyzed. Some analyses may be conducted excluding data after introducing another antihyperglycemic therapy.

### **9.3.5.3. Gastrointestinal Events**

Summaries and analyses for incidence and severity of nausea, vomiting, and diarrhea will be provided by each treatment.

### **9.3.5.4. Adjudicated Cardiovascular Events**

Listings of deaths, myocardial infarctions, strokes, and hospitalizations for unstable angina or heart failure confirmed by an independent Clinical Endpoint Committee (CEC) will be provided.

### **9.3.5.5. Central Laboratory Measures and Vital Signs**

Values and change from baseline to postbaseline values of central laboratory measures and vital signs will be summarized and compared between tirzepatide at the MTD (10 mg or 15 mg) and placebo at each scheduled visit.

### **9.3.5.6. Analysis of C-SSRS Data**

Suicide-related thoughts and behaviors occurring during treatment will be summarized based on responses to the C-SSRS consistent with the C-SSRS Scoring and Data Analysis Guide (The Columbia Lighthouse Project, 2013).

## **9.3.6. Evaluation of Immunogenicity**

The frequency and percentage of participants with preexisting ADA and with TE ADA+ to tirzepatide will be tabulated. Treatment-emergent ADAs are defined as those with a titer 2-fold (1 dilution) greater than the minimum required dilution (1:10) of the ADA assay if no ADAs were detected at baseline (treatment-induced ADA), or those with a 4-fold (2 dilutions) increase in titer compared with baseline if ADAs were detected at baseline (treatment-boosted ADA). For the TE ADA+ participants, the distribution of maximum titers will be described. The frequency of neutralizing antibodies may also be tabulated in TE ADA+ participants. The relationship between the presence of antibodies and the PK parameters and PD response including safety and efficacy to tirzepatide may be assessed.

## **9.3.7. Other Analyses**

### **9.3.7.1. Health Economics**

Analyses of actual and change from baseline in PRO scores will be conducted using linear models with baseline PRO scores, treatment, stratification factors and other factors that may be considered relevant. These variables will be specified in the SAP.

### **9.3.7.2. Subgroup Analyses**

The following subgroups will be analyzed using the “efficacy” estimand on percent change in AHI values from baseline to 52 week visit if there are sufficient numbers of participants in each treatment by subgroup (for example, 10%): Age (<50 years, ≥50 years),

- Baseline OSA severity (Moderate, Severe),
- Race,
- Ethnicity,
- Country,

- Gender (Male or Female),
- Baseline BMI ( $<30$ ,  $\geq 30$  and  $<35$ ,  $\geq 35$  and  $<40$ ,  $\geq 40$  kg/m<sup>2</sup>), and
- Baseline ESS (ESS $\leq 10$ , ESS $>10$ ).

Analyses for percent change from baseline in AHI will be performed using an MMRM model that includes the same fixed effects given for the primary analysis model plus factors of subgroup, 2-way interaction of subgroup and treatment, 2-way interaction of subgroup and visit, and 3-way interaction of treatment, visit and subgroup. The interaction of subgroup and treatment at the primary endpoint (Week 52) will be evaluated to assess the treatment by subgroup interaction. When analyzing OSA severity (Moderate, Severe) as a subgroup, the baseline AHI will be not be included as a covariate to avoid confounding. Additional subgroup analyses may also be performed.

More details on other analyses will be described in the SAP.

#### **9.4. Interim Analysis**

Based on the projected enrollment, approximately 3 interim analyses of safety will be conducted. The first interim analysis is planned to occur when approximately 20% of the anticipated number of participants are randomly assigned to study intervention or 6 months after the first participant is randomly assigned to study intervention, whichever occurs later, from one of the ISAs. Subsequent reviews will follow approximately every 6 months throughout the study.

Only the DMC is authorized to evaluate unblinded interim analyses. Study sites will receive information about interim results only if they need to know for the safety of their participants.

Unblinding details are specified in SAP and/or a separate unblinding plan document.

The DMC charter will describe the specific parameters of the planned interim analysis.

See Section 10.1.5, Appendix 1 And the DMC Charter for details related to the DMC.

The timing of dissemination of data summaries based on interim analyses is addressed in Section 10.1.6, Appendix 1.

#### **9.5. Sample Size Determination**

Approximately 206 participants per ISA will be randomly assigned to either tirzepatide or placebo in a 1:1 ratio (approximately 103 participants per treatment arm), and the statistical power is evaluated for the primary efficacy endpoint and key secondary combination PRO endpoint at a 2-sided significance level of 0.05. This sample size will provide

- At least 90% power to demonstrate superiority of tirzepatide at the MTD (10 mg or 15 mg) to placebo for the mean percent change from baseline in AHI, assuming 50% improvement, with a common standard deviation of 50%, and a dropout rate of 25%
- at least 90% power to demonstrate superiority of tirzepatide at the MTD (10 mg or 15 mg) to placebo for the hierarchical combination PRO endpoint using Finkelstein-Schoenfeld method (Finkelstein and Schoenfeld 1999) with a dropout rate of 25%

An upper limit of approximately 70% enrollment of male participants will be used to ensure a sufficiently large sample of female participants.

The sample size and power for the key secondary hierarchical combination PRO endpoint is estimated through simulations under the following assumptions (Blackman et al. 2016, Weaver et al. 2021):

<b>Variable</b>	<b>Placebo</b>	<b>Tirzepatide at the MTD (10 mg or 15 mg)</b>	<b>Definition of Win</b>
FOSQ-10 – change from baseline (mean $\pm$ SD)	1.24 $\pm$ 1.84	1.95 $\pm$ 1.84	If the change from baseline FOSQ-10 for TZP at MTD (10 mg or 15 mg) minus change from baseline FOSQ-10 for placebo $\geq$ 2.2
FOSQ Vigilance – change from baseline (mean $\pm$ SD)	0.3 $\pm$ 0.5	0.5 $\pm$ 0.5	If the change from baseline FOSQ Vigilance for TZP at MTD (10 mg or 15 mg) minus change from baseline FOSQ Vigilance for placebo $\geq$ 0.44
FOSQ Activity Level – change from baseline (mean $\pm$ SD)	0.3 $\pm$ 0.5	0.5 $\pm$ 0.5	If the change from baseline FOSQ Activity Level for TZP at MTD (10 mg or 15 mg) minus change from baseline FOSQ Activity Level for placebo $\geq$ 0.44

Abbreviations: FOSQ = Functional Outcomes of Sleep Questionnaire; FOSQ-10 = FOSQ 10-item short form; MTD = maximum tolerated dose; SD = standard deviation; TZP = tirzepatide.

## **10. Supporting Documentation and Operational Considerations**

### **10.1. Appendix 1: Regulatory, Ethical, and Study Oversight Considerations**

#### **10.1.1. Regulatory and Ethical Considerations**

This study will be conducted in accordance with the protocol and with the following:

- Consensus ethical principles derived from international guidelines including the Declaration of Helsinki and Council for International Organizations of Medical Sciences International Ethical Guidelines
- Applicable ICH GCP Guidelines
- International Organization for Standardization (ISO) 14155
- Applicable laws and regulations

The protocol, protocol amendments, ICFs, IB, and other relevant documents (for example, advertisements) must be submitted to an IRB/IEC by the investigator and reviewed and approved by the IRB/IEC before the study is initiated.

Any amendments to the protocol will require IRB/IEC approval before implementation of changes made to the study design, except for changes necessary to eliminate an immediate hazard to study participants.

Protocols and any substantial amendments to the protocol will require health authority approval prior to initiation except for changes necessary to eliminate an immediate hazard to study participants.

The investigator will be responsible for the following:

- Providing written summaries of the status of the study to the IRB/IEC annually or more frequently in accordance with the requirements, policies, and procedures established by the IRB/IEC
- Notifying the IRB/IEC of SAEs or other significant safety findings as required by IRB/IEC procedures
- Providing oversight of study conduct for participants under their responsibility and adherence to requirements of 21 Code of Federal Regulations, ICH guidelines, the IRB/IEC, European regulation 536/2014 for clinical studies (if applicable), and all other applicable local regulations

Investigator sites are compensated for participation in the study as detailed in the clinical trial agreement.

#### **10.1.2. Financial Disclosure**

Investigators and sub-investigators will provide the sponsor with sufficient, accurate financial information as requested to allow the sponsor to submit complete and accurate financial certification or disclosure statements to the appropriate regulatory authorities. Investigators are responsible for providing information on financial interests during the course of the study and for 1 year after completion of the study.

### **10.1.3. Informed Consent Process**

The investigator or the investigator's representative will explain the nature of the study, including the risks and benefits, to the participant or the participant's legally authorized representative and answer all questions regarding the study.

Participants must be informed that their participation is voluntary. Participants or their legally authorized representatives will be required to sign a statement of informed consent that meets the requirements of 21 Code of Federal Regulations 50, local regulations, ICH guidelines, privacy and data protection requirements, where applicable, and the IRB/IEC or study center.

The medical record must include a statement that written informed consent was obtained before the participant was entered in the study and the date the written consent was obtained. The authorized person obtaining the informed consent must also sign the ICF.

Participants must be re-consented to the most current version of the ICF(s) during their participation in the study.

A copy of the ICF(s) must be provided to the participant or the participant's legally authorized representative and is kept on file.

Participants who are rescreened are required to sign a new ICF.

### **10.1.4. Data Protection**

Participants will be assigned a unique identifier by the sponsor. Any participant records, datasets or tissue samples that are transferred to the sponsor will contain the identifier only; participant names or any information which would make the participant identifiable will not be transferred.

The participant must be informed that the participant's personal study-related data will be used by the sponsor in accordance with local data protection law. The level of disclosure must also be explained to the participant who will be required to give consent for their data to be used as described in the informed consent.

The participant must be informed that their medical records may be examined by Clinical Quality Assurance auditors or other authorized personnel appointed by the sponsor, by appropriate IRB/IEC members, and by inspectors from regulatory authorities.

The sponsor has processes in place to ensure data protection, information security and data integrity. These processes include appropriate contingency plan(s) for appropriate and timely response in the event of a data security breach.

### **10.1.5. Committees Structure**

Prospective adjudication of major adverse CV events and pancreatic Aes will be performed by the independent adjudication committees for this study.

Sections 10.1.5.1 and 10.1.5.2, Appendix 1 outline additional information on pancreatic and CV adjudication.

An independent DMC for the interim analysis (Section 9.4) will include members with no conflict of financial interest. An independent DMC with members all external to the sponsor will be used to monitor participant safety in an unblinded fashion. For details on the DMC, refer to the DMC charter.

#### **10.1.5.1. Cardiovascular Adjudicated Events**

Deaths and nonfatal CV Aes will be adjudicated by a committee blinded to treatment assignment. The nonfatal CV Aes to be adjudicated include:

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

#### **10.1.5.2. Pancreatitis Adjudicated Event**

Acute pancreatitis is defined as an AE of special interest in this trial (Section 8.3.3).

The diagnosis of acute pancreatitis requires 2 of the following 3 features:

- abdominal pain, characteristic of acute pancreatitis (generally located in the epigastrium and radiates to the back in approximately half the cases) (Banks and Freeman 2006; Koizumi et al. 2006); the pain is often associated with nausea and vomiting
- serum amylase (total and/or pancreatic) and/or lipase  $\geq 3X$  ULN, and
- characteristic findings of acute pancreatitis on CT scan or MRI.

All suspected cases of acute or chronic pancreatitis will be adjudicated by a committee blinded to treatment assignment. In addition, Aes of severe or serious abdominal pain of unknown etiology will also be submitted to the adjudication committee to assess for possible pancreatitis or other pancreatic disease. Relevant data from participants with acute or chronic pancreatitis and those with severe or serious abdominal pain will be entered into a specifically designed CRF page. The adjudication committee representative will enter the results of adjudication in a corresponding CRF page.

#### **10.1.6. Dissemination of Clinical Study Data**

##### **Reports**

The sponsor will disclose a summary of study information, including tabular study results, on publicly available websites where required by local law or regulation.

The summary of results will be posted within the time frame specified by local law or regulation. If the study remains ongoing in some countries and a statistical analysis of an incomplete data set would result in analyses lacking scientific rigor (for example, underpowered) or compromise the integrity of the overall analyses (for example, trial not yet unblinded), the summary of results will be submitted within 1 year after the end of the study globally or as soon as available, whichever is earlier.

The publication policy for Study GPIF is outlined in Section 10.1.10, Appendix 1 and further described in the Clinical Trial Agreement.

**Data**

The sponsor provides access to all individual participant data collected during the trial, after anonymization, with the exception of PK or genetic data.

Data are available to request 6 months after the indication studied has been approved in the US and EU and after primary publication acceptance, whichever is later. No expiration date of data requests is currently set once data are made available.

Access is provided after a proposal has been approved by an independent review committee identified for this purpose and after receipt of a signed data sharing agreement.

Data and documents, including the study protocol, SAP, CSR, and blank or annotated CRFs, will be provided in a secure data sharing environment for up to 2 years per proposal.

For details on submitting a request, see the instructions provided at [www.vivli.org](http://www.vivli.org).

**10.1.7. Data Quality Assurance**

All participant data relating to the study will be recorded on printed or electronic CRFs unless transmitted to the sponsor or designee electronically (for example, laboratory data). The investigator is responsible for verifying that data entries are accurate and correct by physically or electronically signing the CRF.

The investigator must maintain accurate documentation (source data) that supports the information entered in the CRF.

The investigator must permit study-related monitoring, audits, IRB/IEC review, and regulatory agency inspections and provide direct access to source data documents.

Quality tolerance limits (QTLs) will be pre-defined to identify systematic issues that can impact participant safety and/or reliability of study results. These pre-defined parameters will be monitored during the study and important excursions from the QTLs and remedial actions taken will be summarized in the CSR.

Monitoring details describing strategy (for example, risk-based initiatives in operations and quality such as risk management and mitigation strategies and analytical risk-based monitoring), methods, responsibilities and requirements, including handling of noncompliance issues and monitoring techniques are provided in the Monitoring Plan.

The sponsor or designee is responsible for the data management of this study including quality checking of the data.

The sponsor assumes accountability for actions delegated to other individuals (for example, contract research organizations).

Study monitors will perform ongoing source data verification to confirm that data transcribed into the CRF by authorized site personnel are accurate, complete, and verifiable from source documents; that the safety and rights of participants are being protected; and that the study is being conducted in accordance with the currently approved protocol and any other study agreements, ICH GCP, and all applicable regulatory requirements.

Records and documents, including signed ICFs, pertaining to the conduct of this study must be retained by the investigator for the time period outlined in the Clinical Trial Agreement unless



local regulations or institutional policies require a longer retention period. No records may be destroyed during the retention period without the written approval of the sponsor. No records may be transferred to another location or party without written notification to the sponsor.

In addition, sponsor or its representatives will periodically check a sample of the participant data recorded against source documents at the study site. The study may be audited by sponsor or its representatives, and/or regulatory agencies at any time. Investigators will be given notice before an audit occurs.

### **Data Capture System**

The investigator is responsible for ensuring the accuracy, completeness, legibility, and timeliness of the data reported to the sponsor.

An EDC system will be used in this study for the collection of CRF data. The investigator maintains a separate source for the data entered by the investigator or designee into the sponsor-provided EDC system. The investigator is responsible for the identification of any data to be considered source and for the confirmation that data reported are accurate and complete by signing the CRF.

Additionally, COA data (clinician-reported outcome instrument) for suicidality assessments will be collected by the authorized study personnel, via a paper source document and will be transcribed by the authorized study personnel into the EDC system.

Additionally, electronic COA data (participant-focused outcome instrument) will be directly recorded by the participant, into an instrument (for example, handheld smart phone or tablet). The electronic COA data will serve as the source documentation and the investigator does not maintain a separate, written or electronic record of these data.

Data collected via the sponsor-provided data capture system(s) will be stored at third parties. The investigator will have continuous access to the data during the study and until decommissioning of the data capture system(s). Prior to decommissioning, the investigator will receive or access an archival copy of pertinent data for retention.

Data managed by a central vendor, such as laboratory test data, will be stored electronically in the central vendor's database system and reports/electronic transfers will be provided to the investigator for review and retention. Data will subsequently be transferred from the central vendor to the sponsor data warehouse.

Data collected on the actigraphy device will be transferred electronically to the sponsor data warehouse, via a third party.

Data from complaint forms submitted to the sponsor will be encoded and stored in the global PC management system.

#### **10.1.8. Source Documents**

Source documents provide evidence for the existence of the participant and substantiate the integrity of the data collected. Source documents are filed at the investigator's site.

Data reported on or entered in the CRF and are transcribed from source documents must be consistent with the source documents or the discrepancies must be explained. The investigator

may need to request previous medical records or transfer records, depending on the study. Also, current medical records must be available.

Definition of what constitutes source data can be found in Section 10.1.7, Appendix 1.

### **10.1.9. Study and Site Start and Closure**

#### **First Act of Recruitment**

The study start date is the date on which the clinical study will be open for recruitment of participants. The first act of recruitment is the first site open and will be the study start date.

#### **Study or Site Termination**

The sponsor or sponsor's designee reserves the right to close the study site or terminate the study at any time for any reason at the sole discretion of the sponsor. Study sites will be closed upon study completion. A study site is considered closed when all required documents and study supplies have been collected and a study-site closure visit has been performed.

The investigator may initiate study-site closure at any time, provided there is reasonable cause and sufficient notice is given in advance of the intended termination.

Reasons for the early closure of a study site by the sponsor or investigator may include but are not limited to:

- For study termination:
  - Discontinuation of further study intervention development
- For site termination:
  - Failure of the investigator to comply with the protocol, the requirements of the IRB/IEC or local health authorities, the sponsor's procedures, or GCP guidelines
  - Inadequate recruitment (evaluated after a reasonable amount of time) of participants by the investigator
  - Total number of participants included earlier than expected.

If the study is prematurely terminated or suspended, the sponsor shall promptly inform the investigators, the IECs/IRBs, the regulatory authorities, and any contract research organization(s) used in the study of the reason for termination or suspension, as specified by the applicable regulatory requirements. The investigator shall promptly inform the participant and should assure appropriate participant therapy and/or follow-up.

### **10.1.10. Publication Policy**

In accordance with the sponsor's publication policy, the results of this study will be submitted for publication by a peer-reviewed journal.

### **10.1.11. Investigator Information**

Researchers with appropriate education, training, and experience, as determined by the sponsor, will participate as investigators in this clinical trial.

**10.1.12. Sample Retention**

Sample retention enables use of new technologies, response to regulatory questions, and investigation of variable response that may not be observed until later in the development of tirzepatide or after tirzepatide becomes commercially available.

<b>Sample Type</b>	<b>Custodian</b>	<b>Retention Period After Last Participant Visit<sup>a</sup></b>
Exploratory Biomarker Samples	Sponsor or Designee	15 years
PK Samples	Sponsor or Designee	2 years
Genetics	Sponsor or Designee	15 years
Immunogenicity (ADA) Samples	Sponsor or Designee	15 years

Abbreviations: ADA = antidrug antibody; PK = pharmacokinetic.

<sup>a</sup>Retention periods may differ locally.

## 10.2. Appendix 2: Clinical Laboratory Tests

The tests detailed in the table below will be performed by the central laboratory.

Local laboratory results are only required in the event that the central laboratory results are not available in time for either study intervention administration and/or response evaluation. If a local sample is required, it is important that the sample for central analysis is obtained at the same time.

In circumstances where the sponsor approves local laboratory testing in lieu of central laboratory testing (in the table below), the local laboratory must be qualified in accordance with applicable local regulations.

Protocol-specific requirements for inclusion or exclusion of participants are detailed in Section 5 of the protocol.

Additional tests may be performed at any time during the study as determined necessary by the investigator or required by local regulations.

Investigators must document their review of the laboratory safety results.

Laboratory results that could unblind the study will not be reported to investigative sites or other blinded personnel.

Clinical Laboratory Tests	Comments
<b>Hematology</b>	Assayed by Lilly-designated laboratory
Hemoglobin	
Hematocrit	
Erythrocyte count (RBCs)	
Mean cell volume	
Mean cell hemoglobin	
Mean cell hemoglobin concentration	
Leukocytes (WBCs)	
Differential	
Percent and/or Absolutes Count of:	
Neutrophils, segmented	
Bands	
Lymphocytes	
Monocytes	
Eosinophils	
Basophils	
Platelets	
Cell morphology (RBCs and WBCs)	
<b>Clinical Chemistry</b>	Assayed by Lilly-designated laboratory
Sodium	
Potassium	
Chloride	
Bicarbonate	

<b>Clinical Laboratory Tests</b>	<b>Comments</b>
Total bilirubin	
Direct bilirubin	
ALP	
ALT	
AST	
BUN	
Creatinine	
CK	
Uric acid	
Albumin	
Calcium	
Glucose	
<b>Lipid Panel</b>	Assayed by Lilly designated laboratory.
Cholesterol	
Triglycerides	
HDL-C	.
LDL-C	Generated by Lilly-designated laboratory. If Triglycerides are >400; direct LDL will be measured.
VLDL-C	Generated by Lilly-designated laboratory.
<b>Hormones (female)</b>	
Serum Pregnancy	Assayed by Lilly-designated laboratory.
Urine Pregnancy	Assayed and Evaluated locally
FSH	Assayed by Lilly-designated laboratory.
<b>Urine Chemistry</b>	Assayed by Lilly-designated laboratory.
Albumin	
Creatinine	
<b>Calculations</b>	Generated by Lilly-designated laboratory.
eGFR (CKD-EPI)	
UACR	
<b>PK Samples – Tirzepatide</b>	Assayed by Lilly-designated laboratory. Results will not be provided to the investigative sites.
<b>Immunogenicity (ADA) Samples</b>	Assayed by Lilly-designated laboratory. Results will not be provided to the investigative sites.
Anti-tirzepatide antibodies OR Tirzepatide antibodies	
Anti-tirzepatide antibodies neutralization OR Tirzepatide antibodies neutralization	
<b>Additional Testing</b>	Assayed by Lilly-designated laboratory.
HbA1c	
Calcitonin	
Pancreatic Amylase	
Lipase	

Clinical Laboratory Tests	Comments
Insulin	Results will not be provided to the investigative sites.
C-Peptide	Results will not be provided to the investigative sites.
Free Fatty Acids	Results will not be provided to the investigative sites.
hsCRP	Results will not be provided to the investigative sites.
Cystatin-C	
TSH	
<b>Stored Samples</b>	Assayed by Lilly-designated laboratory. Results will not be provided to the investigative sites.
Genetics Sample	
Exploratory Biomarker Samples	
Serum	
Plasma (EDTA)	
Plasma (P800)	

Abbreviations: ADA = antidrug antibody; ALP = alkaline phosphatase; ALT = alanine aminotransferase;; AST = aspartate aminotransferase; BUN = blood urea nitrogen; CK = creatinine kinase, CKD-EPI = Chronic Kidney Disease-Epidemiology; EDTA = ethylenediaminetetraacetic acid; eGFR = estimated glomerular filtration rate; FSH = follicle-stimulating hormone; HbA1c = hemoglobin A1c; HDL-C = high-density lipoprotein cholesterol; hsCRP = C-Reactive Protein, high-sensitivity; IWRS = interactive web-response system; LDL-C = low-density lipoprotein cholesterol; PK = pharmacokinetics; RBC = red blood cells; TSH = thyroid-stimulating hormone; UACR = urine albumin/creatinine ratio; VLDL-C = very-low-density lipoprotein cholesterol; WBC = white blood cells.

### 10.2.1. Laboratory Samples to be Obtained at the Time of a Systemic Hypersensitivity Event

#### Purpose of collecting samples after a systemic hypersensitivity event

The samples listed in this appendix are not collected for acute study participant management. The sponsor will use the laboratory tests results from these samples to characterize hypersensitivity events across the clinical development program.

#### When to collect samples after a systemic hypersensitivity event occurs

Collect the samples listed below if a systemic hypersensitivity event is suspected. The timing should be as designated in the table, assuming the participant has been stabilized.

Obtain follow-up pre-dose samples at the next regularly scheduled laboratory sample collection (ideally prior to the next dose after the event) to assess post-event return-to-baseline values.

Timing	Sample Type	Laboratory Test <sup>a</sup>
Collect from 30 minutes to 4 hours after the start of the event.  ● Note: The optimal collection time is from 1 to 2 hours after the start of event.	Serum	total tryptase
	Serum	complements (C3, C3a, and C5a)
	Serum	cytokine panel (IL-6, IL-1 $\beta$ , IL-10 or any cytokine panel that includes these 3 cytokines)
Collect only if not already collected on the same day as the event.  ● Note: If collecting, collect up to 12 hours after the start of the event.	Serum	Tirzepatide ADA
	Serum/plasma	Tirzepatide concentration

Abbreviations: ADA = anti-drug antibodies; IL = interleukin.

A All samples for hypersensitivity testing will be assayed by Lilly-designated laboratory. Results will not be provided to the study site. If samples are not collected or are collected outside the specified time period, this will not be considered a protocol deviation.

### What information to record

Record the date and time when the samples are collected.

### Allowed additional testing for participant management

The investigator may perform additional tests locally, if clinically indicated, for acute study participant management.

**10.3. Appendix 3: Adverse Events and Serious Adverse Events: Definitions and Procedures for Recording, Evaluating, Follow-up, and Reporting**

- The definitions and procedures detailed in this appendix are in accordance with International Organization for Standardization (ISO) 14155.
- Both the investigator and the sponsor will comply with all local medical device reporting requirements.
- The detection and documentation procedures described in this protocol apply to all sponsor medical devices provided for use in the study. See Section 6.1.1 for the list of sponsor medical devices).

**10.3.1. Definition of AE**

<b>AE Definition</b>
<ul style="list-style-type: none"> <li>• An AE is any untoward medical occurrence in a participant administered a pharmaceutical product and which does not necessarily have a causal relationship with the study intervention. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product.</li> <li>• An AE is any untoward medical occurrence, unintended disease or injury, or untoward clinical signs (including abnormal laboratory finding) in study participants, users, or other persons, whether or not related to the investigational medical device. This definition includes events related to the investigational medical device or comparator and events related to the procedures involved except for events in users or other persons, which only include events related to investigational devices.</li> </ul>

<b>Events Meeting the AE Definition</b>
<ul style="list-style-type: none"> <li>• Any abnormal laboratory test results (hematology, clinical chemistry, or urinalysis) or other safety assessments (for example, ECG, radiological scans, vital signs measurements), including those that worsen from baseline, considered clinically significant in the medical and scientific judgment of the investigator (that is, not related to progression of underlying disease).</li> <li>• Exacerbation of a chronic or intermittent pre-existing condition including either an increase in frequency and/or intensity of the condition.</li> <li>• New condition detected or diagnosed after study intervention administration even though it may have been present before the start of the study.</li> <li>• Signs, symptoms, or the clinical sequelae of a suspected drug-drug interaction.</li> <li>• Medication error, misuse, or abuse of study intervention, including signs, symptoms, or clinical sequelae.</li> <li>• Lack of efficacy or failure of expected pharmacological action per se will not be reported as an AE or SAE. Such instances will be captured in the efficacy assessments.</li> </ul>



However, the signs, symptoms, and/or clinical sequelae resulting from lack of efficacy will be reported as AE or SAE if they fulfill the definition of an AE or SAE.

#### Events NOT Meeting the AE Definition

- Any clinically significant abnormal laboratory findings or other abnormal safety assessments that are associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant's condition.
- The disease/disorder being studied or expected progression, signs, or symptoms of the disease/disorder being studied, unless more severe than expected for the participant's condition.
- Medical or surgical procedure (for example, endoscopy, appendectomy): the condition that leads to the procedure is the AE.
- Situations in which an untoward medical occurrence did not occur (social and/or convenience admission to a hospital).
- Anticipated day-to-day fluctuations of pre-existing disease(s) or condition(s) present or detected at the start of the study that do not worsen.

#### 10.3.2. Definition of SAE

**An SAE is defined as any untoward medical occurrence that, at any dose, meets one or more of the criteria listed:**

##### a. Results in death

##### b. Is life-threatening

The term *life-threatening* in the definition of *serious* refers to an event in which the participant was at risk of death at the time of the event. It does not refer to an event, which hypothetically might have caused death, if it were more severe.

##### c. Requires inpatient hospitalization or prolongation of existing hospitalization

- In general, hospitalization signifies that the participant has been admitted to hospital or emergency ward (usually involving at least an overnight stay) for observation and/or treatment that would not have been appropriate in the physician's office or outpatient setting. Complications that occur during hospitalization are AEs. If a complication prolongs hospitalization or fulfills any other serious criteria, the event is serious. When in doubt as to whether hospitalization occurred or was necessary, the AE should be considered serious.
- Hospitalization for elective treatment of a pre-existing condition that did not worsen from baseline is not considered an AE.

##### d. Results in persistent disability/incapacity

- The term disability means a substantial disruption of a person's ability to conduct normal life functions.
- This definition is not intended to include experiences of relatively minor medical significance such as uncomplicated headache, nausea, vomiting, diarrhea, influenza,

and accidental trauma (for example, sprained ankle) which may interfere with or prevent everyday life functions but do not constitute a substantial disruption.
<p><b>e. Is a congenital anomaly/birth defect</b></p> <ul style="list-style-type: none"> <li>Abnormal pregnancy outcomes (for example, spontaneous abortion, fetal death, stillbirth, congenital anomalies, ectopic pregnancy) are considered SAEs.</li> </ul>
<p><b>f. Other situations:</b></p> <ul style="list-style-type: none"> <li>Medical or scientific judgment should be exercised by the investigator in deciding whether SAE reporting is appropriate in other situations such as important medical events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the participant or may require medical or surgical intervention to prevent one of the other outcomes listed in the above definition. These events should usually be considered serious.</li> <li>Examples of such events include invasive or malignant cancers, intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias or convulsions that do not result in hospitalization, or development of drug dependency or drug abuse.</li> </ul>
<p><b>g.</b> Resulted in medical or surgical intervention to prevent life-threatening illness or injury or permanent impairment to a body structure or a body function.</p>

### 10.3.3. Definition of Product Complaints

Product Complaint
<ul style="list-style-type: none"> <li>A PC is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, durability, reliability, safety, effectiveness or performance of a study intervention. When the ability to use the study intervention safely is impacted, the following are also PCs: <ul style="list-style-type: none"> <li>Deficiencies in labeling information, and</li> <li>Use errors for device or drug-device combination products due to ergonomic design elements of the product.</li> </ul> </li> <li>Product complaints related to study interventions used in clinical trials are collected in order to ensure the safety of participants, monitor quality, and to facilitate process and product improvements.</li> <li>Investigators will instruct participants to contact the site as soon as possible if he or she has a PC or problem with the study intervention so that the situation can be assessed.</li> <li>An event may meet the definition of both a PC and an AE/SAE. In such cases, it should be reported as both a PC and as an AE/SAE.</li> </ul>

**10.3.4. Recording and Follow-Up of AE and/or SAE and Product Complaints**

<b>AE, SAE, and PC Recording</b>
<ul style="list-style-type: none"> <li>• When an AE/SAE/PC occurs, it is the responsibility of the investigator to review all documentation (for example, hospital progress notes, laboratory reports, and diagnostics reports) related to the event.</li> <li>• The investigator will then record all relevant AE/SAE/PC information in the participant's medical records, in accordance with the investigator's normal clinical practice. AE/SAE information is reported on the appropriate CRF page and PC information is reported on the PC Form.</li> </ul> <p>Note: An event may meet the definition of both a PC and an AE/SAE. In such cases, it should be reported as both a PC and as an AE/SAE.</p> <ul style="list-style-type: none"> <li>• It is <b>not</b> acceptable for the investigator to send photocopies of the participant's medical records to sponsor or designee in lieu of completion of the CRF page for AE/SAE and the PC Form for product complaints.</li> <li>• There may be instances when copies of medical records for certain cases are requested by sponsor or designee. In this case, all participant identifiers, with the exception of the participant number, will be redacted on the copies of the medical records before submission to sponsor or designee.</li> <li>• The investigator will attempt to establish a diagnosis of the event based on signs, symptoms, and/or other clinical information. Whenever possible, the diagnosis (not the individual signs/symptoms) will be documented as the AE/SAE.</li> </ul>
<b>Assessment of Intensity</b>
<p>The investigator will make an assessment of intensity for each AE and SAE reported during the study and assign it to one of the following categories:</p> <ul style="list-style-type: none"> <li>• <b>Mild:</b> A type of AE that is usually transient and may require only minimal treatment or therapeutic intervention. The event does not generally interfere with usual activities of daily living.</li> <li>• <b>Moderate:</b> A type of AE that is usually alleviated with additional specific therapeutic intervention. The event interferes with usual activities of daily living, causing discomfort but poses no significant or permanent risk of harm to the research participant.</li> <li>• <b>Severe:</b> A type of AE that interrupts usual activities of daily living, or significantly affects clinical status, or may require intensive therapeutic intervention. An AE that is assessed as severe should not be confused with an SAE. Severe is a category utilized for rating the intensity of an event; and both AEs and SAEs can be assessed as severe.</li> </ul> <p>An event is defined as "serious" when it meets at least one of the predefined outcomes as described in the definition of an SAE, NOT when it is rated as severe.</p>

### Assessment of Causality

- The investigator is obligated to assess the relationship between study intervention and each occurrence of each AE/SAE. The investigator will use clinical judgment to determine the relationship.
- A “reasonable possibility” of a relationship conveys that there are facts, evidence, and/or arguments to suggest a causal relationship, rather than a relationship cannot be ruled out.
- Alternative causes, such as underlying disease(s), concomitant therapy, and other risk factors, as well as the temporal relationship of the event to study intervention administration will be considered and investigated.
- The investigator will also consult the IB in their assessment.
- For each AE/SAE, the investigator **must** document in the medical notes that he/she has reviewed the AE/SAE and has provided an assessment of causality.
- There may be situations in which an SAE has occurred and the investigator has minimal information to include in the initial report to sponsor or designee. However, it is very important that the investigator always make an assessment of causality for every event before the initial transmission of the SAE data to sponsor or designee.
- The investigator may change their opinion of causality in light of follow-up information and send an SAE follow-up report with the updated causality assessment.
- The causality assessment is one of the criteria used when determining regulatory reporting requirements.

### Follow-Up of Aes and SAEs

- The investigator is obligated to perform or arrange for the conduct of supplemental measurements and/or evaluations as medically indicated or as requested by sponsor or designee to elucidate the nature and/or causality of the AE or SAE as fully as possible. This may include additional laboratory tests or investigations, histopathological examinations, or consultation with other health care professionals.
- If a participant dies during participation in the study or during a recognized follow-up period, the investigator will provide sponsor or designee with a copy of any post-mortem findings including histopathology.

### 10.3.5. Reporting of SAEs

#### SAE Reporting via an Electronic Data Collection Tool

- The primary mechanism for reporting an SAE will be the electronic data collection tool.
- If the electronic system is unavailable, then the site will use the SAE paper form (see next section) in order to report the event within 24 hours.
- The site will enter the SAE data into the electronic system as soon as it becomes available.

- After the study is completed at a given site, the electronic data collection tool will be taken off-line to prevent the entry of new data or changes to existing data.
- If a site receives a report of a new SAE from a study participant or receives updated data on a previously reported SAE after the electronic data collection tool has been taken off-line, then the site can report this information on an SAE paper form (see next section) or to the sponsor or designee by telephone.
- Contacts for SAE reporting can be found on the SAE form.

#### **SAE Reporting via Paper Form**

- Facsimile transmission of the SAE paper form is the preferred method to transmit this information to the sponsor or designee.
- Initial notification via telephone does not replace the need for the investigator to complete and sign the SAE CRF pages within the designated reporting time frames.
- Contacts for SAE reporting can be found on the SAE form.

### **10.3.6. Regulatory Reporting Requirements**

#### **SAE Regulatory Reporting**

- Prompt notification by the investigator to the sponsor of an SAE is essential so that legal obligations and ethical responsibilities toward the safety of participants and the safety of a study intervention under clinical investigation are met.
- The sponsor has a legal responsibility to notify both the local regulatory authority and other regulatory agencies about the safety of a study intervention under clinical investigation. The sponsor will comply with country-specific regulatory requirements relating to safety reporting to the regulatory authority, IRB/IEC, and investigators.
- An investigator who receives an investigator safety report describing an SAE or other specific safety information (for example, summary or listing of SAEs) from the sponsor will review and then file it along with the IB and will notify the IRB/IEC, if appropriate according to local requirements.

### **10.3.7. Special Safety Topics**

#### **10.3.7.1. Hypoglycemia**

Participants will be trained by authorized study personnel about signs and symptoms of hypoglycemia and how to treat hypoglycemia, and how to collect appropriate information for each episode of hypoglycemia.

Hypoglycemia may be identified by spontaneous reporting of symptoms from participants (whether confirmed or unconfirmed by simultaneous glucose values) or by blood glucose samples collected during study visits.

All participants who develop diabetes during the study will be provided with glucometers. Participants without diabetes may, at the investigator's discretion, be given glucometers to assist in the evaluation of reported symptoms consistent with hypoglycemia. Participants receiving glucometers will be provided a diary to record relevant information (for example, glucose values, symptoms).

Hypoglycemic episodes will be recorded on a specific CRF and should not be recorded as AEs unless the event meets serious criteria. If a hypoglycemic event meets severe criteria (see definition below), it should be recorded as serious on the AE and SAE CRFs, and reported to Lilly as an SAE.

Investigators should use the following classification of hypoglycemia (ADA 2020):

### **Level 1 hypoglycemia:**

**Glucose <70 mg/dL (3.9 mmol/L) and  $\geq$  54 mg/dL (3.0 mmol/L):** Level 1 hypoglycemia can alert a person to take action such as treatment with fast-acting carbohydrates. Providers should continue to counsel participants to treat hypoglycemia at this glucose alert value.

### **Level 2 hypoglycemia:**

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

### **Level 3 hypoglycemia:**

**Severe hypoglycemia (in adults):** A severe event characterized by altered mental and/or physical status requiring assistance for treatment of hypoglycemia. For example, participants had altered mental status, and could not assist in their own care, or were semiconscious or unconscious, or experienced coma with or without seizures, and the assistance of another person was needed to actively administer carbohydrate, glucagon, or other resuscitative actions. Glucose measurements may not be available during such an event, but neurological recovery attributable to the restoration of glucose concentration to normal is considered sufficient evidence that the event was induced by a low glucose concentration.

- The determination of a hypoglycemic event as an episode of severe hypoglycemia, as defined above, is made by the investigator based on the medical need of the participant to have required assistance and is not predicated on the report of a participant simply having received assistance.
- If a hypoglycemic event meets the criteria of severe hypoglycemia, the investigator must record the event as serious on the AE CRF and report it to Lilly as an SAE.

### **Nocturnal hypoglycemia:**

Nocturnal hypoglycemia is a hypoglycemia event (including severe hypoglycemia) that **occurs at night** and presumably during sleep.

To avoid duplicate reporting, all consecutive BG values <70 mg/dL (<3.9 mmol/L) occurring within a 1-hour period may be considered to be a single hypoglycemic event (Weinberg et al. 2010; Danne et al. 2013).

#### **10.3.7.2. Hypersensitivity Reactions**

Many drugs, including oral agents and biologic agents, carry the risk of systemic hypersensitivity reactions. If such a reaction occurs, additional data should be provided to the sponsor in the designated CRFs.

Sites should have appropriately trained medical staff and appropriate medical equipment available when study participants are receiving study intervention. It is recommended that participants who experience a systemic hypersensitivity reaction be treated per national and international guidelines.

In the case of a suspected systemic hypersensitivity event, additional blood samples should be collected as described in Section 10.2.1. Laboratory results are provided to the sponsor via the central laboratory.

#### **10.3.7.3. Injection-Site Reactions**

Symptoms and signs of a local ISR may include erythema, induration, pain, pruritus, and edema.

If an ISR is reported by a participant or parent or guardian or site staff, the ISR CRF will be used to capture additional information about this reaction, for example, injection site pain, degree and area of erythema, induration, pruritis and edema.

At the time of AE occurrence in the tirzepatide group, samples will be collected for measurement of tirzepatide ADAs and tirzepatide concentration.

## 10.4. Appendix 4: Contraceptive and Barrier Guidance

### 10.4.1. Definitions

Word/Phrase	Definition
Women of child bearing potential	<p>Females are considered a WOCBP if</p> <ul style="list-style-type: none"> <li>• they have had at least 1 cycle of menses, or</li> <li>• they have Tanner 4 breast development.</li> </ul> <p>Any amount of spotting should be considered menarche. If Tanner Staging of breasts is performed as part of study procedures, please refer to the Reproductive, Pregnancy and Pediatrics Safety Committee Safety Guidance for Children in Clinical Trial regarding Tanner staging.</p>
Women not of child bearing potential	<p>Females are considered women not of child bearing potential if</p> <ul style="list-style-type: none"> <li>• they have a congenital anomaly such as Mullerian agenesis</li> <li>• they are infertile due to surgical sterilization, or</li> <li>• they are post-menopausal.</li> </ul> <p>Examples of surgical sterilization include: hysterectomy, bilateral oophorectomy, or tubal ligation.</p>
Post-menopausal state	<p>The post-menopausal state should be defined as:</p> <ol style="list-style-type: none"> <li>1. A woman at any age at least 6 weeks post-surgical bilateral oophorectomy with or without hysterectomy, confirmed by operative note; or</li> <li>2. A woman at least 40 years of age and up to 55 years old with an intact uterus, not on hormone therapy, who has had cessation of menses for at least 12 consecutive months without an alternative medical cause, AND with a follicle-stimulating hormone &gt;40 mIU/mL; or</li> <li>3. A woman 55 or older not on hormone therapy, who has had at least 12 months of spontaneous amenorrhea; or</li> <li>4. A woman at least 55 years of age with a diagnosis of menopause prior to starting hormone replacement therapy</li> </ol> <p>* Women should not be taking medications during amenorrhea such as oral contraceptives, hormones, gonadotropin-releasing hormone, anti-estrogens, SERMs, or chemotherapy that could induce transient amenorrhea.</p>

Abbreviations: SERM = selective estrogen receptor modulators; WOCBP = woman of child bearing potential.



**10.4.2. Contraception Guidance**

**10.4.2.1. Females**

**WOCBP who are completely abstinent as their preferred and usual lifestyle, or in a same sex relationship, as part of their preferred and usual lifestyle**

Must...	Must not...
<ul style="list-style-type: none"> <li>• agree to either remain abstinent, or</li> <li>• stay in a same sex relationship without sexual relationships with males</li> </ul>	<ul style="list-style-type: none"> <li>• use periodic abstinence methods                             <ul style="list-style-type: none"> <li>○ calendar</li> <li>○ ovulation</li> <li>○ symptothermal, or</li> <li>○ post-ovulation</li> </ul> </li> <li>• declare abstinence just for the duration of a trial, or</li> <li>• use the withdrawal method</li> </ul>

**WOCBP who are NOT completely abstinent as their preferred and usual lifestyle, or in a same sex relationship, as part of their preferred and usual lifestyle**

Topic	Condition
Pregnancy testing	Negative serum result at screening, followed by a negative urine result within 24 hours prior to treatment exposure.
Contraception	Agree to use 2 forms of effective contraception, where at least one form must be highly effective (less than 1% failure rate) for the duration of the trial and for 30 days thereafter.

**Examples of different forms of contraception:**

Methods	Examples
Highly effective contraception	<ul style="list-style-type: none"> <li>• combination oral contraceptive pill and mini-pill</li> <li>• implanted contraceptives</li> <li>• injectable contraceptives</li> <li>• contraceptive patch (only women &lt;198 pounds or 90 kg)</li> <li>• total abstinence</li> <li>• vasectomy (if only sexual partner)</li> <li>• fallopian tube implants (if confirmed by hysterosalpingogram)</li> <li>• combined contraceptive vaginal ring, or</li> <li>• intrauterine devices</li> </ul>
Effective contraception	<ul style="list-style-type: none"> <li>• male or female condoms with spermicide</li> <li>• diaphragms with spermicide or cervical sponges</li> <li>• barrier method with use of a spermicide                             <ul style="list-style-type: none"> <li>○ condom with spermicide</li> </ul> </li> </ul>

	<ul style="list-style-type: none"> <li>○ diaphragm with spermicide, or</li> <li>○ female condom with spermicide</li> </ul> <p>Note: The barrier method must include use of a spermicide (that is, condom with spermicide, diaphragm with spermicide, female condom with spermicide) to be considered effective.</p>
Ineffective forms of contraception	<ul style="list-style-type: none"> <li>● spermicide alone</li> <li>● immunocontraceptives</li> <li>● periodic abstinence</li> <li>● fertility awareness (calendar method, temperature method, combination of above 2, cervical mucus, symptothermal)</li> <li>● withdrawal</li> <li>● post coital douche</li> <li>● lactational amenorrhea</li> </ul>

**10.4.2.2. Males**

The table below describes contraception guidance for all men.

Topic	Guidance
For all men	should refrain from sperm donation for the duration of the study and for 120 days (4 months) from the last dose of study intervention received
Contraception for men with partners of childbearing potential	<ul style="list-style-type: none"> <li>● either remain abstinent (if this is their preferred and usual lifestyle), or</li> <li>● must use condoms during intercourse for the duration of the study, <b>and</b></li> <li>● for 120 days (4 months) from the last dose of study intervention received</li> </ul>
Contraception for men in exclusively same sex relationships, as their preferred and usual lifestyle	Are not required to use contraception

Examples of highly effective, effective and unacceptable methods of contraception are listed in Section 10.4.2.1, Appendix 4.

## 10.5. Appendix 5: Genetics

### Use/Analysis of DNA

Genetic variation may impact a participant's response to study intervention, susceptibility to, and severity and progression of disease. Variable response to study intervention may be due to genetic determinants that impact drug absorption, distribution, metabolism, and excretion; mechanism of action of the drug; disease etiology; and/or molecular subtype of the disease being treated. Therefore, where local regulations and IRB/IEC allow, a blood sample will be collected for DNA analysis from consenting participants.

DNA samples will be used for research related to tirzepatide, OSA, obesity and related diseases. They may also be used to develop tests/assays including diagnostic tests related to tirzepatide, OSA, and obesity. Genetic research may consist of the analysis of 1 or more candidate genes or the analysis of genetic markers throughout the genome or analysis of the entire genome (as appropriate).

The samples may be analyzed as part of a multi-study assessment of genetic factors involved in the response to tirzepatide or study interventions of this class to understand study disease or related conditions.

The results of genetic analyses may be reported in the CSR or in a separate study summary.

The sponsor will store the DNA samples in a secure storage space with adequate measures to protect confidentiality.

The samples will be retained while research on tirzepatide, OSA, and obesity continues but no longer than the sample retention limits described in Section 10.1.12, Appendix 1.

## 10.6. Appendix 6: Liver Safety: Suggested Actions and Follow-up Assessments

### 10.6.1. Hepatic Evaluation Testing

See Sections 10.6.2 and 10.6.3, Appendix 6 for guidance on appropriate test selection.

The Lilly-designated central laboratory must complete the analysis of all selected testing except for microbiology testing.

Local testing may be performed *in addition to central testing* when necessary for immediate participant management.

Results will be reported if a validated test or calculation is available.

<b>Hematology</b>	<b>Clinical Chemistry</b>
Hemoglobin	Total bilirubin
Hematocrit	Direct bilirubin
Erythrocytes (RBCs – red blood cells)	Alkaline phosphatase (ALP)
Leukocytes (WBCs – white blood cells)	Alanine aminotransferase (ALT)
Differential:	Aspartate aminotransferase (AST)
Neutrophils, segmented	Gamma-glutamyl transferase (GGT)
Lymphocytes	Creatine kinase (CK)
Monocytes	<b>Other Chemistry</b>
Basophils	Acetaminophen
Eosinophils	Acetaminophen protein adducts
Platelets	Alkaline phosphatase isoenzymes
Cell morphology (RBC and WBC)	Ceruloplasmin
	Copper
<b>Coagulation</b>	Ethyl alcohol (EtOH) (quantitative)
Prothrombin time, INR (PT-INR)	Haptoglobin
<b>Serology</b>	Immunoglobulin A (IgA) (quantitative)
Hepatitis A virus (HAV) testing:	Immunoglobulin G (IgG) (quantitative)
HAV total antibody	Immunoglobulin M (IgM) (quantitative)
HAV IgM antibody	Phosphatidylethanol (Peth)
Hepatitis B virus (HBV) testing:	<b>Urine Chemistry</b>
Hepatitis B surface antigen (HbsAg)	Drug screen
Hepatitis B surface antibody (anti-HBs)	Ethyl glucuronide (EtG)
Hepatitis B core total antibody (anti-HBc)	<b>Other Serology</b>
Hepatitis B core IgM antibody	Anti-nuclear antibody (ANA)

Hepatitis B core IgG antibody	Anti-smooth muscle antibody (ASMA) <sup>a</sup>
HBV DNA <sup>b</sup>	Anti-actin antibody <sup>c</sup>
Hepatitis C virus (HCV) testing:	Epstein-Barr virus (EBV) testing:
HCV antibody	EBV antibody
HCV RNA <sup>b</sup>	EBV DNA <sup>b</sup>
Hepatitis D virus (HDV) testing:	Cytomegalovirus (CMV) testing:
HDV antibody	CMV antibody
Hepatitis E virus (HEV) testing:	CMV DNA <sup>b</sup>
HEV IgG antibody	Herpes simplex virus (HSV) testing:
HEV IgM antibody	HSV (Type 1 and 2) antibody
HEV RNA <sup>b</sup>	HSV (Type 1 and 2) DNA <sup>b</sup>
<b>Microbiology <sup>d</sup></b>	Liver kidney microsomal type 1 (LKM-1) antibody
Culture:	
Blood	
Urine	

<sup>a</sup> Not required if anti-actin antibody is tested.

<sup>b</sup> Reflex/confirmation dependent on regulatory requirements, testing availability, or both.

<sup>c</sup> Not required if anti-smooth muscle antibody (ASMA) is tested.

<sup>d</sup> Assayed ONLY by investigator-designated local laboratory; no central testing available.

### 10.6.2. Close Hepatic Monitoring

Laboratory tests (Section 10.2, Appendix 2), including ALT, AST, ALP, TBL, direct bilirubin, gamma-glutamyl transferase, and creatine kinase, should be repeated within 48 to 72 hours to confirm the abnormality and to determine if it is increasing or decreasing, if one or more of these conditions occur:

If a participant with baseline results of...	develops the following elevations:
ALT or AST <1.5x ULN	ALT or AST ≥3x ULN
ALP <1.5x ULN	ALP ≥2x ULN
TBL <1.5x ULN	TBL ≥2x ULN (except for participants with Gilbert's syndrome)
ALT or AST ≥1.5x ULN	ALT or AST ≥2x baseline
ALP ≥1.5x ULN	ALP ≥2x baseline
TBL ≥1.5x ULN	TBL ≥1.5x baseline (except for participants with Gilbert's syndrome)

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; TBL = total bilirubin level; ULN = upper limit of normal.

If the abnormality persists or worsens, clinical and laboratory monitoring, and evaluation for possible causes of abnormal liver tests should be initiated by the investigator in consultation with the Lilly-designated medical monitor. At a minimum, this evaluation should include physical examination and a thorough medical history, including symptoms, recent illnesses (for example, heart failure, systemic infection, hypotension, or seizures), recent travel, history of concomitant medications (including over-the-counter), herbal and dietary supplements, history of alcohol drinking and other substance abuse.

Initially, monitoring of symptoms and hepatic biochemical tests should be done at a frequency of 1 to 3 times weekly, based on the participant's clinical condition and hepatic biochemical tests. Subsequently, the frequency of monitoring may be lowered to once every 1 to 2 weeks, if the participant's clinical condition and lab results stabilize. Monitoring of ALT, AST, ALP, and total bilirubin level should continue until levels normalize or return to approximate baseline levels.

### 10.6.3. Comprehensive Hepatic Evaluation

A comprehensive evaluation should be performed to search for possible causes of liver injury if one or more of these conditions occur:

If a participant with baseline results of..	develops the following elevations:
ALT or AST <1.5x ULN	ALT or AST $\geq 3x$ ULN with hepatic signs/symptoms <sup>a</sup> , or ALT or AST $\geq 5x$ ULN
ALP <1.5x ULN	ALP $\geq 3x$ ULN
TBL <1.5x ULN	TBL $\geq 2x$ ULN (except for participants with Gilbert's syndrome)
ALT or AST $\geq 1.5x$ ULN	ALT or AST $\geq 2x$ baseline with hepatic signs/symptoms <sup>a</sup> , or ALT or AST $\geq 3x$ baseline
ALP $\geq 1.5x$ ULN	ALP $\geq 2x$ baseline
TBL $\geq 1.5x$ ULN	TBL $\geq 2x$ baseline (except for participants with Gilbert's syndrome)

Abbreviations: ALP = alkaline phosphatase; ALT= alanine aminotransaminase; AST = aspartate aminotransferase; TBL = total bilirubin level; ULN = upper limit of normal.

a Hepatic signs/symptoms are severe fatigue, nausea, vomiting, right upper quadrant abdominal pain, fever, rash, and/or eosinophilia >5%.

At a minimum, this evaluation should include physical examination and a thorough medical history, as outlined above, as well as tests for PT-INR; tests for viral hepatitis A, B, C, or E; tests for autoimmune hepatitis; and an abdominal imaging study (for example, ultrasound or CT scan).

Based on the participant's history and initial results, further testing should be considered in consultation with the Lilly-designated medical monitor, including tests for hepatitis D virus, cytomegalovirus, Epstein-Barr virus, acetaminophen levels, acetaminophen protein adducts, urine toxicology screen, Wilson's disease, blood alcohol levels, urinary ethyl glucuronide, and blood phosphatidylethanol. Based on the circumstances and the investigator's assessment of the participant's clinical condition, the investigator should consider referring the participant for a hepatologist or gastroenterologist consultation, magnetic resonance cholangiopancreatography, endoscopic retrograde cholangiopancreatography, cardiac echocardiogram, or a liver biopsy.

#### **Additional hepatic data collection (hepatic safety CRF) in study participants who have abnormal liver tests during the study**

Additional hepatic safety data collection in hepatic safety CRFs should be performed in study participants who meet 1 or more of the following 5 conditions:

1. Elevation of serum ALT to  $\geq 5x$  ULN on 2 or more consecutive blood tests (if baseline ALT <1.5x ULN)
  - In participants with baseline ALT  $\geq 1.5x$  ULN, the threshold is ALT  $\geq 3x$  baseline on 2 or more consecutive tests
2. Elevated TBL to  $\geq 2x$  ULN (if baseline TBL <1.5x ULN) (except for cases of known Gilbert's syndrome)
  - In participants with baseline TBL  $\geq 1.5x$  ULN, the threshold should be TBL  $\geq 2x$  baseline
3. Elevation of serum ALP to  $\geq 2x$  ULN on 2 or more consecutive blood tests (if baseline ALP <1.5x ULN)
  - In participants with baseline ALP  $\geq 1.5x$  ULN, the threshold is ALP  $\geq 2x$  baseline on 2 or more consecutive blood tests
4. Hepatic event considered to be an SAE
5. Discontinuation of study intervention due to a hepatic event

**Note:** the interval between the 2 consecutive blood tests should be at least 2 days.

**10.7. Appendix 7: Medical Device Adverse Events (AEs), Adverse Device Effects (ADEs), Serious Adverse Events (SAEs) and Device Deficiencies: Definition and Procedures for Recording, Evaluating, Follow-up, and Reporting**

Refer to Appendix 3 for definitions and procedures for recording, evaluating, follow-up, and reporting of all events.



## **10.8. Appendix 8: Protocol GPIF Standardized Protocols for the Measurement of Height, Weight, Neck Circumference, Waist Circumference, Vital Signs, and Electrocardiogram**

The following information has been adapted from standardized physical measurement protocols for the World Health Organization's STEPwise approach to Surveillance (STEPS) (WHO 2017).

### **10.8.1. Measuring Height**

**Step 1.** Ask the participant to remove their footwear and any headgear (light headgear worn for religious reasons can remain, but this should be worn by the participant at every clinic visit when their height is measured).

**Step 2.** Ask the participant to stand on the calibrated height measuring board (stadiometer) or against a wall with their feet together and their knees straight with their heels against the backboard, the stadiometer, or the wall.

**Step 3.** Ask the participant to look straight ahead without tilting their head up.

**Step 4.** Ask the participant to breathe in and stand tall. Measure and record the participant's height in centimeters to 1 decimal place.

### **10.8.2. Measuring Weight**

- Body weight measurements should be done in a consistent manner using a calibrated electronic scale capable of measuring weight in kilograms to 1 decimal place.
- All weights for a given participant should be measured using the same scale, whenever possible, at approximately the same time in the morning after evacuation of bladder contents.
- Body weight will be measured in fasting state. If the participant is not fasting, the participant should be called in for a new visit within the visit window to have the fasting body weight measured.

**Step 1.** Ask the participant to empty their pockets, remove their footwear, outerwear (coat, jacket, etc.), and any headgear (light headgear worn for religious reasons can remain, but this should be worn by the participant at every clinic visit when weight is measured).

**Step 2.** Make sure the scale is placed on a firm, flat, even surface (not on carpet, on a sloping surface, or a rough, uneven surface).

**Step 3.** Ask the participant to step onto the scale with 1 foot on each side of the scale.

**Step 4.** Ask the participant to stand still with arms by sides and then record weight in kilograms to the nearest one-tenth kilogram.

### **10.8.3. Measuring Hip and Waist Circumference**

- Hip circumference measurements should be obtained with the participant in the standing position. The hip circumference should be measured at the maximal circumference of the buttocks.

- Waist circumference should be measured in the horizontal plane and at the midpoint between the lower margin of the last palpable rib and the top of the iliac crest.
- Measurements should be taken at the end of a normal expiration using a nonstretchable measuring tape. The tape should lie flat against the skin without compressing the soft tissue.
- The waist circumference should be measured twice, rounded to the nearest 0.5 cm. The measuring tape should be removed between the 2 measurements. Both measurements will be recorded in the CRF. If the difference between the 2 measurements exceeds 1 cm, this set of measurements should be discarded and the 2 measurements repeated.

**Step 1.** Ask the participant to wear light clothing (if available, patient gowns could also be used).

**Step 2.** Ask the participant to stand with their feet close together, arms at their side, body weight evenly distributed.

**Step 3.** Ask the participant to relax and measure the participant's waist circumference.

#### **10.8.4. Measuring Neck Circumference**

- Participants should look straight ahead during the measurement, with shoulders down (not hunched).
- Measure the neck circumference at a point just below the larynx (Adam's Apple) and perpendicular to the long axis of the neck.
- Do not place the tape measure over the Adam's Apple.
- The tape will be as close to horizontal as anatomically feasible (the tape line in the front of the neck should be at the same height as the tape line in the back of the neck).
- Care should be taken so as not to involve the shoulder/neck muscles (trapezius) in the measurement.
- Round neck measurement up to the nearest half centimeter.

#### **10.8.5. Vital Sign Measurements (Blood Pressure and Heart Rate)**

- Vital sign measurements (BP and heart rate, measured by pulse) should be taken before obtaining an ECG tracing and before collection of blood samples for laboratory testing
- The participant should sit quietly for 5 minutes before vital signs measurements are taken
- For each parameter, 3 measurements will be taken using the same arm, preferably the nondominant arm
- The recordings should be taken at least 1 minute apart. Each measurement of sitting pulse and BP needs to be recorded in the CRF
- Blood pressure must be taken with an automated BP instrument

- If BP and pulse measurements are taken separately, pulse should be taken prior to BP.

*Note:* In the event pulse measurement cannot be taken via an automated BP instrument, the preferred location for measurement of pulse is the radial artery.

#### **10.8.6. Electrocardiogram**

- All digital ECGs will be obtained using local ECG machines.
- 12-lead ECGs should be obtained after the participant has rested in a supine position for at least 10 minutes.
- Electrocardiograms should be collected at least 30 minutes prior to collection of blood samples for laboratory testing, including PK samples.
- Electrocardiograms should be obtained approximately 1 minute apart, with all 3 tracings to be obtained within approximately 5 minutes. Measurements that deviate substantially from previous readings should be repeated immediately.

## **10.9. Appendix 9: Provisions for Changes in Study Conduct Due to the COVID-19 Pandemic**

### **Implementation of this appendix**

The changes to procedures described in this appendix are temporary measures intended to be used only during specific time periods as directed by the sponsor in partnership with the investigator.

### **Study disruptions due to the COVID-19 pandemic**

Individual, site, or regional restrictions due to the COVID-19 pandemic may cause disruptions to the conduct of the study. These disruptions may limit the ability of the investigators or participants, or both, to attend on-site visits or to conduct planned study procedures.

### **Implementing changes due to the COVID-19 pandemic**

After receiving the sponsor's written approval, sites may implement changes if permitted by local regulations.

After approval by local ERBs/IRBs, regulatory bodies and any other relevant local authorities, implementation of these changes will not typically require additional notification to these groups, unless they have specific requirements in which notification is required (for example, upon implementation and suspension of changes). All approvals and notifications must be retained in the study records.

If the sponsor grants written approval for changes in study conduct, the sponsor will also provide additional written guidance, if needed.

### **Considerations for making a change**

The prevailing consideration for making a change is ensuring the safety of study participants. Additional important considerations for making a change are compliance with GCP, enabling participants to continue safely in the study and maintaining the integrity of the study.

### **Informed consent**

Additional consent from the participant will be obtained, if required, for:

- participation in remote visits, as defined in Section "Remote Visits"
- dispensation of additional study intervention during an extended treatment period
- alternate delivery of study intervention and ancillary supplies, and
- provision of their personal or medical information required prior to implementation of these activities.

### **Changes in study conduct due to the COVID-19 pandemic**

Changes in study conduct not described in this appendix, or not consistent with applicable local regulations, are not allowed.

The following changes in study conduct will not be considered protocol deviations.

***Remote visits****Types of remote visits***Telephone/Telemedicine**

Telephone or technology-assisted virtual visits, or both, are acceptable to complete appropriate assessments. Assessments to be completed in this manner include, but are not limited to, collection of AEs and PCs, concomitant medications review, review study participant diary (including study intervention compliance), review diet and exercise goals, C-SSRS (Since Last Visit Assessed) and PHQ-9. PROs will be completed by the participant on the provisioned device per the SoA.

**Mobile healthcare visit**

Healthcare visits may be performed by a mobile healthcare provider at locations other than the study site when participants cannot travel to the site due to the COVID pandemic, if written approval is provided by the sponsor. Procedures performed at such visits include, but are not limited to, weight measurement, blood sample collection, vital signs (temperature, PR, BP), concomitant medication review, conducting physical assessments, collection of AEs and PCs, and collecting health information. PROs will be completed by the participant on the provisioned device per the SoA.

Every effort should be made for the participant to return to on-site visits as soon as reasonably possible, while ensuring the safety of the participant and investigational site staff.

*Data capture*

In source documents and the CRF, the study site should capture the visit method, with a specific explanation for any data missing because of missed in-person site visits.

*Safety reporting*

Regardless of the type of remote visits implemented, the protocol requirements regarding the reporting of AEs, SAEs, and PCs remain unchanged.

*Return to on-site visits*

Every effort should be made to enable participants to return to on-site visits as soon as reasonably possible, while ensuring the safety of both the participants and the site staff.

***Local laboratory testing option***

Local laboratory testing may be conducted in lieu of central laboratory testing. However, central laboratory testing must be retained for PK, immunogenicity, hsCRP, insulin and lipid samples. The local laboratory must be qualified in accordance with applicable local regulations.

***Study intervention and supplies***

When a participant is unable to go to the site to receive study supplies during normal on-site visits, the site should work with the sponsor to determine appropriate actions. These actions may include:

- asking the participant to go to the site and receive study supplies from site staff without completion of a full study visit

- asking the participant's designee to go to the site and receive study supplies on a participant's behalf, and
- arranging delivery of study supplies.

These requirements must be met before action is taken:

- Alternate delivery of study intervention should be performed in a manner that does not compromise treatment blinding and ensures product integrity. The existing protocol requirements for product accountability remain unchanged, including verification of participant's receipt of study supplies.
- When delivering supplies to a location other than the study site (for example, participant's home), the investigator, sponsor, or both should ensure oversight of the shipping process to ensure accountability and product quality (that is, storage conditions maintained and intact packaging upon receipt).
- Instructions may be provided to the participant or designee on the final disposition of any unused or completed study supplies.

### ***Screening period guidance***

To ensure safety of study participants, laboratory values and other eligibility assessments taken at Visit 1 are valid for a maximum of 90 days. The following rules will be applied for active participants not randomly assigned to study intervention whose participation in the study must be paused due to the COVID-19 pandemic:

- If screening is paused for less than 90 days from Visit 1 to Visit 2: the participant will proceed to the next study visit per the usual SoA, provided that Visit 2 must be conducted within 90 days from Visit 1.
  - The site should conduct the next visit if the participant's eligibility criteria are confirmed, and the site should document the reason for delay.
  - Due to the pause in screening, sites should also reconfirm the impacted participant's consent and document this confirmation in the source documentation.
- If screening is paused for more than 90 days from Visit 1 to Visit 2: the participant must be discontinued because of screening interruption due to the COVID-19 pandemic. This is documented as a screen failure in the CRF. The participant can re-consent and be rescreened as a new participant. This rescreen is in addition to the one allowed by the main protocol. The screening procedures per the usual SoA should be followed, starting at Visit 1 to ensure participant eligibility by Visit 2.

### ***Adjustments to visit windows***

Whenever possible and safe to do so, as determined by the investigator's discretion, participants should complete the usual SoA. To maximize the possibility that these visits can be conducted as on-site visits, the windows for visits may be adjusted, upon further guidance from the sponsor. This minimizes missing data and preserves the intended conduct of the study.

This table describes the allowed adjustments to visit windows.

<b>Visit Number</b>	<b>Tolerance</b>
Visit 1 (Screening)	No change
Visit 2 (Randomization)	Within 90 days after Visit 1.
Visits 3 through 6	Within 7 days before or after the intended date.
Visits 7 through 10	Within 14 days before or after the intended date.
Visit 11	Within 14 days before the intended date, or up to 28 days after the intended date.
Visit 801	Up to 28 days after the intended date

For participants whose visits have extended windows, additional study intervention may need to be provided to avoid interruption and maintain overall integrity of the study.

### **Documentation**

*Changes to study conduct will be documented*

Sites will identify and document the details of how participants, visits types, and conducted activities were affected due to the COVID-19 pandemic. Dispensing/shipment records of study intervention and relevant communications, including delegation, should be filed with site study records.

*Source documents at alternate locations*

Source documents generated at a location other than the study site should be part of the investigator's source documentation and should be transferred to the site in a secure and timely manner.

**10.10. Appendix 10: Abbreviations and Definitions**

<b>Term</b>	<b>Definition</b>
<b>AASM</b>	American Academy of Sleep Medicine
<b>Abuse</b>	Use of a study intervention for recreational purposes or to maintain an addiction or dependence
<b>ADA</b>	antidrug antibody
<b>AE</b>	adverse event
<b>AHI</b>	Apnea-Hypopnea Index
<b>ALP</b>	alkaline phosphatase
<b>ALT</b>	alanine aminotransferase
<b>ANCOVA</b>	analysis of covariance
<b>Apnea</b>	decrease in airflow $\geq 90\%$ from baseline for $\geq 10$ seconds
<b>AST</b>	aspartate aminotransferase
<b>AX6</b>	Axivity 6
<b>blinding/masking</b>	A single-blind study is one in which the investigator and/or the investigator's staff are aware of the treatment but the participant is not, or vice versa, or when the sponsor is aware of the treatment but the investigator and/the investigator's staff and the participant are not. A double-blind study is one in which neither the participant nor any of the investigator or sponsor staff who are involved in the treatment or clinical evaluation of the participants are aware of the treatment received.
<b>BMI</b>	body mass index
<b>BP</b>	blood pressure
<b>CBD</b>	cannabidiol
<b>COA</b>	clinical outcome assessment
<b>complaint</b>	A complaint is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, purity, durability, reliability, safety or effectiveness, or performance of a drug or drug delivery system.
<b>compliance</b>	Adherence to all study-related, good clinical practice (GCP), and applicable regulatory requirements.
<b>CRF</b>	case report form; a printed, optical, or electronic document designed to record all of the protocol-required information to be reported to the sponsor for each trial participant.
<b>CSR</b>	clinical study report
<b>C-SSRS</b>	Columbia-Suicide Severity Rating Scale
<b>CV</b>	cardiovascular
<b>Device deficiencies</b>	Equivalent to product complaint
<b>DMC</b>	data monitoring committee. A data monitoring committee, or data monitoring board (DMB) is a group of independent scientists who are appointed to monitor the safety and scientific integrity of a human research intervention, and to make recommendations to the sponsor regarding the stopping of a study for efficacy, or for harms, or for futility. The composition of the committee is dependent upon the scientific skills and knowledge required for monitoring the particular study.
<b>DPP-4</b>	dipeptidyl-peptidase-4
<b>EAS</b>	efficacy analysis set
<b>ECG</b>	electrocardiogram
<b>ED</b>	early discontinuation
<b>EDC</b>	electronic data capture
<b>eGFR</b>	estimated glomerular filtration rate
<b>enroll</b>	The act of assigning a participant to a treatment. Participants who are enrolled in the study are those who have been assigned to a treatment.



<b>enter</b>	Participants entered into a study are those who sign the informed consent form directly or through their legally acceptable representatives.
<b>ERB</b>	ethical review board
<b>EQ-5D-5L</b>	EuroQol - 5 Dimension - 5 Level
<b>ESS</b>	Epworth Sleepiness Scale
<b>FAS</b>	full analysis set
<b>FG</b>	fasting glucose
<b>FOSQ</b>	Functional Outcomes of Sleep Questionnaire
<b>GCP</b>	good clinical practice
<b>GI</b>	gastrointestinal
<b>GIP</b>	glucose-dependent insulinotropic polypeptide
<b>GLP-1</b>	glucagon-like peptide-1
<b>GLP-1R</b>	glucagon-like peptide-1 receptor
<b>GLP-1RA</b>	glucagon like peptide 1 receptor
<b>HbA1c</b>	hemoglobin A1c
<b>HDL</b>	high-density lipoprotein
<b>HSAT</b>	home sleep apnea test
<b>hsCRP</b>	high-sensitivity C-reactive protein
<b>Hypopnea</b>	an abnormal respiratory event lasting $\geq 10$ seconds with $\geq 30\%$ reduction in thoracoabdominal movement or airflow as compared to baseline, and with $\geq 4\%$ oxygen desaturation.
<b>IB</b>	Investigator's Brochure
<b>ICF</b>	informed consent form
<b>ICH</b>	International Council for Harmonisation
<b>IEC</b>	Independent Ethics Committees
<b>IMP</b>	Investigational Medicinal Product (see also "investigational product") A medicinal product which is being tested or used as a reference, including as a placebo, in a clinical trial.
<b>informed consent</b>	A process by which a participant voluntarily confirms their willingness to participate in a particular study, after having been informed of all aspects of the study that are relevant to the participant's decision to participate. Informed consent is documented by means of a written, signed and dated informed consent form.
<b>INR</b>	International normalized ratio
<b>interim analysis</b>	An interim analysis is an analysis of clinical study data, separated into treatment groups, that is conducted before the final reporting database is created/locked.
<b>investigational product</b>	A pharmaceutical form of an active ingredient or placebo being tested or used as a reference in a clinical trial, including products already on the market when used or assembled (formulated or packaged) in a way different from the authorized form, or marketed products used for an unauthorized indication, or marketed products used to gain further information about the authorized form.
<b>IRB</b>	institutional review board
<b>ISA</b>	Indication-Specific Appendix
<b>ISR</b>	injection site reaction
<b>ITT</b>	intention to treat: The principle that asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a participant (that is, the planned treatment regimen) rather than the actual treatment given. It has the consequence that participant allocated to a treatment group should be followed up, assessed, and analyzed as members of that group irrespective of their compliance to the planned course of treatment.
<b>IWRS</b>	interactive web-response system
<b>MACE</b>	major adverse cardiovascular events

<b>medication error</b>	Errors in the prescribing, dispensing, or administration of a study intervention, regardless of whether or not the medication is administered to the participant or the error leads to an AE. Medication error generally involve a failure to uphold one or more of the five “rights” of medication use: the right participant, the right drug, the right dose, right route, at the right time. In addition to the core five rights, the following may also represent medication errors: <ul style="list-style-type: none"> <li>• dose omission associated with an AE or a product complaint</li> <li>• dispensing or use of expired medication</li> <li>• use of medication past the recommended in-use date</li> <li>• dispensing or use of an improperly stored medication</li> <li>• use of an adulterated dosage form or administration technique inconsistent with the medication's labeling (for example, Summary of Product Characteristics, IB, local label, protocol), or</li> <li>• shared use of cartridges, prefilled pens, or both.</li> </ul>
<b>MEN</b>	multiple endocrine neoplasia
<b>MI</b>	myocardial infarction
<b>misuse</b>	Use of a study intervention for self-treatment that either is inconsistent with the prescribed dosing regimen, indication, or both, or is obtained without a prescription.
<b>MMRM</b>	Mixed model repeated measures
<b>MTC</b>	medullary thyroid cancer
<b>MTD</b>	maximum tolerated dose
<b>OGTT</b>	oral glucose tolerance test
<b>OSA</b>	obstructive sleep apnea
<b>OTC</b>	Over the counter
<b>PAP</b>	positive airway pressure
<b>participant</b>	Equivalent to CDISC term “subject”: an individual who participates in a clinical trial, either as recipient of an investigational medicinal product or as a control.
<b>PC</b>	product complaint
<b>PGIC</b>	Patient Global Impression of Change
<b>PGIS</b>	Patient Global Impression of Status
<b>PHQ-9</b>	Patient Health Questionnaire-9
<b>PK/PD</b>	pharmacokinetics/pharmacodynamics
<b>PR</b>	pulse rate
<b>PRO/ePRO</b>	patient-reported outcomes/electronic patient-reported outcomes
<b>PROMIS</b>	Patient-Reported Outcomes Measurement Information System
<b>PSG</b>	polysomnography
<b>QTc</b>	corrected QT interval
<b>QW</b>	weekly
<b>SAE</b>	serious adverse event
<b>SAP</b>	statistical analysis plan
<b>SASHB</b>	sleep apnea specific hypoxic burden
<b>SBP</b>	systolic blood pressure
<b>SC</b>	subcutaneous
<b>SD</b>	standard deviation
<b>screen</b>	The act of determining if an individual meets minimum requirements to become part of a pool of potential candidates for participation in a clinical study.
<b>SF-36v2</b>	Short-Form 36 version 2
<b>SIB</b>	suicidal ideation and behavior
<b>SoA</b>	Schedule of Activities
<b>SS</b>	safety analysis set
<b>T1DM</b>	Type 1 diabetes mellites

<b>T2DM</b>	Type 2 diabetes mellites
<b>TBL</b>	total bilirubin level
<b>TEE</b>	total energy expenditure
<b>TEAE</b>	treatment-emergent adverse event: An untoward medical occurrence that emerges during a defined treatment period, having been absent pretreatment, or worsens relative to the pretreatment state, and does not necessarily have to have a causal relationship with this treatment.
<b>THC</b>	Tetrahydrocannabinol
<b>TSH</b>	thyroid-stimulating hormone
<b>UACR</b>	urinary albumin/creatinine ratio
<b>ULN</b>	upper limit of normal
<b>WOCBP</b>	woman of childbearing potential

---

## 10.11. Appendix 11: Country-Specific Requirements

### 10.11.1. Germany

This section describes protocol changes applicable for adult participants in study sites in Germany.

This table describes the changes and provides a rationale for the changes.

Protocol Section Number and Name	Description of the Change	Brief Rationale
7.2. Participant Discontinuation/Withdrawal from the Study		
8.3. Adverse Events, Serious Adverse Events, and Product Complaints	Deleted references to “legally authorized representative,” “legal guardian,” “parents”	The German Drug Law (Arzneimittelgesetz – AMG) requires per Paragraph 40 (1-3) and Paragraph 41 (3) that adult participants act on their own behalf and provide their own written informed consent. If written consent is not possible, verbal consent with a witness is acceptable. No legal representative consent is accepted.
10.1.3 Informed Consent Process		
10.10. Abbreviations and Definitions		

The revised text in the following subsections show the changes applicable for adult participants to study sites in Germany. Additions are identified by underline. Deletions are identified by ~~strikethrough format~~.

### 7.2. Participant Discontinuation/Withdrawal from the Study

Discontinuation is expected to be uncommon.

A participant may withdraw from the study:

- at any time at the participant’s own request
- ~~at the request of the participant’s designee (for example, parents or legal guardian)~~
- at the discretion of the investigator for safety, behavioral, compliance, or administrative reasons
  - If the participant becomes pregnant during the study (see Section 8.3.2 for additional details)
- if enrolled in any other clinical study involving an investigational product, or enrolled in any other type of medical research judged not to be scientifically or medically compatible with this study
- if the participant, for any reason, requires treatment with a therapeutic agent that is prohibited by the protocol and has been demonstrated to be effective for treatment of the study indication. In this case, discontinuation from the study occurs prior to introduction of the new agent.

**8.3. Adverse Events, Serious Adverse Events, and Product Complaints**

The definitions of the following events can be found in Section 10.3, Appendix 3:

- AEs
- SAEs
- PCs

These events will be reported by the participant (or, when appropriate, by a caregiver, or surrogate, or the participant's legally authorized representative).

### 10.1.3 Informed Consent Process

The investigator or the investigator's representative will explain the nature of the study, including the risks and benefits, to the participant ~~or the participant's legally authorized representative~~ and answer all questions regarding the study.

Participants must be informed that their participation is voluntary. Participants ~~or their legally authorized representatives~~ will be required to sign a statement of informed consent that meets the requirements of 21 Code of Federal Regulations 50, local regulations, ICH guidelines, privacy and data protection requirements, where applicable, and the IRB/IEC or study center.

The medical record must include a statement that written informed consent was obtained before the participant was entered in the study and the date the written consent was obtained. The authorized person obtaining the informed consent must also sign the ICF.

Participants must be re-consented to the most current version of the ICF(s) during their participation in the study.

A copy of the ICF(s) must be provided to the participant ~~or the participant's legally authorized representative~~ and is kept on file.

Participants who are rescreened are required to sign a new ICF.

**10.10. Appendix 10: Abbreviations and Definitions**

**enter**

Participants entered into a study are those who sign the informed consent form directly ~~or through their legally acceptable representatives.~~

## 11. References

- [ADA] American Diabetes Association Glycemic targets: Standards of medical care in diabetes-2020. *Diabetes Care*. 2020;43(Suppl 1):S66–S76. <https://doi.org/10.2337/dc20-S006>
- [AHRQ] Agency for Healthcare Research and Quality. Continuous positive airway pressure treatment for obstructive sleep apnea. Accessed July 20, 2021. <https://www.ahrq.gov/sites/default/files/wysiwyg/research/findings/ta/drafts-for-review/sleep-apnea-draftreport.pdf>
- Alosh M, Bretz F, Mohammad H. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693-713. <https://doi.org/10.1002/sim.5974>
- Banks PA, Freeman ML; Practice Parameters Committee of the American College of Gastroenterology. Practice guidelines in acute pancreatitis. *Am J Gastroenterol*. 2006;101(10):2379-2400.
- Blackman A, Foster GD, Zammit G, et al. Effect of liraglutide 3.0 mg in individuals with obesity and moderate or severe obstructive sleep apnea: the SCALE Sleep Apnea randomized clinical trial. *Int J Obes (Lond)*. 2016;40(8):1310-1319.
- Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913. <https://doi.org/10.1002/bimj.201000239>
- Currie AC, Kaur V, Carey I, et al. Obstructive sleep apnea remission following bariatric surgery: a national registry cohort study. *Surgery Obes Relat Dis*. 2021;17(9):1576-1582. <https://doi.org/10.1016/j.soard.2021.05.021>
- Dahl D, Onishi Y, Norwood P, et al. Tirzepatide, a dual GIP/GLP-1 receptor agonist, is effective and safe when added to basal insulin for treatment of type 2 diabetes (SURPASS-5). *Diabetes*. 2021;70(suppl\_1). <https://doi.org/10.2337/db21-80-LB>
- Danne T, Philotheou A, Goldman D, et al. A randomized trial comparing the rate of hypoglycemia – assessed using continuous glucose monitoring – in 125 preschool children with type 1 diabetes treated with insulin glargine or NPH insulin (the PRESCHOOL study). *Pediatr Diabetes*. 2013;14(8):593-601. <https://doi.org/10.1111/pedi.12051>
- EuroQol Research Foundation. EQ-5D-5L user guide, version 3.0. Updated September 2019. <https://euroqol.org/publications/user-guides>
- FAO/WHO/UNU. *Human energy requirements. Report of a joint FAO/WHO/UNU Expert Consultation*. Rome. 2004.
- [FDA] Food and Drug Administration. Guidance for Industry. Drug-Induced Liver Injury: Premarketing Clinical Evaluation. July 2009. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/drug-induced-liver-injury-premarketing-clinical-evaluation>
- Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. *Stat Med*. 1999;18(11):1341-1354. [https://doi.org/10.1002/\(sici\)1097-0258\(19990615\)18:11<1341::aid-sim129>3.0.co;2-7](https://doi.org/10.1002/(sici)1097-0258(19990615)18:11<1341::aid-sim129>3.0.co;2-7)



- Frias JP, Davies MJ, Rosenstock J, et al. Tirzepatide versus semaglutide once weekly in patients with type 2 diabetes. *N Engl J Med*. 2021;385(6):503-515.  
<https://doi.org/10.1056/nejmoa2107519>
- Gottlieb DJ, Punjabi NM. Diagnosis and management of obstructive sleep apnea: a review. *JAMA*. 2020;323(14):1389-1400. <https://doi.org/10.1001/jama.2020.3514>
- Hamilton GS, Gupta R, Vizcarra D, et al; Guidelines Committee Members and Governing Council of the World Sleep Society. Endorsement of: “clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American academy of sleep medicine clinical practice guideline” by the World Sleep Society. *Sleep Med*. 2021;79:152-154.  
<https://doi.org/10.1016/j.sleep.2020.12.044>.
- Hudgel DW, Patel SR, Ahasic AM, et al. The role of weight management in the treatment of adult obstructive sleep apnea. An Official American Thoracic Society Clinical Practice Guideline. *Am J Respir Crit Care Med*. 2018;198(6):e70-e87.  
<https://doi.org/10.1164/rccm.201807-1326ST>
- Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540-545. <https://doi.org/10.1093/sleep/14.6.540>
- Koizumi M, Takada T, Kawarada Y, et al. JPN guidelines for the management of acute pancreatitis: diagnostic criteria for acute pancreatitis. *J Hepatobiliary Pancreat Surg*. 2006;13(1):25-32. <https://doi.org/10.1007/s00534-005-1048-2>
- Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613.
- [Lilly]. Eli Lilly and Company. Lilly’s tirzepatide achieves all primary and key secondary study outcomes against insulin glargine in adults with type 2 diabetes and increased cardiovascular risk in SURPASS-4 trial. Published May 20, 2021. Accessed July 15, 2021.  
<http://lilly.mediaroom.com/2021-05-20-Lillys-tirzepatide-achieves-all-primary-and-key-secondary-study-outcomes-against-insulin-glargine-in-adults-with-type-2-diabetes-and-increased-cardiovascular-risk-in-SURPASS-4-trial>
- Ludvik B, Giorgino F, Jodar E, et al. Efficacy and safety of tirzepatide, a dual GIP/GLP-1 receptor agonist, compared with insulin degludec in patients with type 2 diabetes (SURPASS 3). *Diabetes*. 2021;70(suppl\_1). <https://doi.org/10.2337/db21-78-LB>
- Maruish ME, editor. User’s manual for the SF-36v2 Health Survey. 3rd ed. Lincoln, RI: Quality Metric Incorporated; 2011.
- [Northwestern]. Northwestern University. PROMIS Short Form v1.0 – Sleep-Related Impairment 8a. Published May 09, 2016a. Accessed Dec 01, 2021.  
[https://www.healthmeasures.net/index.php?option=com\\_instruments&view=measure&id=185&Itemid=992](https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=185&Itemid=992)
- [Northwestern]. Northwestern University. PROMIS Short Form v1.0 – Sleep Disturbance 8b. Published May 03, 2016b. Accessed Dec 01, 2021.  
[https://www.healthmeasures.net/index.php?option=com\\_instruments&view=measure&id=184&Itemid=992](https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=184&Itemid=992)

- Opperer M, Cozowicz C, Bugada D, et al. Does obstructive sleep apnea influence perioperative outcome? A qualitative systematic review for the society of anesthesia and sleep medicine task force on preoperative preparation of patients with sleep-disordered breathing. *Anesth Analg*. 2016;122(5):1321-1334. <https://doi.org/10.1213/ane.0000000000001178>
- Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. 2012;33(2):176-182. <https://doi.org/10.1093/eurheartj/her352>
- Rosenstock J, Wysham C, Frias JP, et al. Efficacy and safety of a novel dual GIP and GLP-1 receptor agonist tirzepatide in patients with type 2 diabetes (SURPASS-1): a double-blind, randomised, phase 3 trial. *Lancet*. 2021;398(10295):143-155. [https://doi.org/10.1016/s0140-6736\(21\)01324-6](https://doi.org/10.1016/s0140-6736(21)01324-6)
- Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons Inc.; 1987.
- The Columbia Lighthouse Project. The Columbia Protocol for Research. Columbia-Suicide Severity Rating Scale Scoring and Data Analysis Guide. Version 2.0. Published Feb 2013. Accessed August 03, 2021. <http://cssrs.columbia.edu/wp-content/uploads/ScoringandDataAnalysisGuide-for-Clinical-Trials-1.pdf>
- Tietjens JR, Claman D, Kezirian EJ, et al. Obstructive sleep apnea in cardiovascular disease: a review of the literature and proposed multidisciplinary clinical management strategy. *J Am Heart Assoc*. 2019;8(1):e010440. <https://doi.org/10.1161/JAHA.118.010440>
- Weaver TE, Grunstein RR. Adherence to continuous positive airway pressure therapy: the challenge to effective treatment. *Proc Am Thorac Soc*. 2008;5(2):173-178. <https://doi.org/10.1513/pats.200708-119MG>
- Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*. 1997;20(10):835-843.
- Weaver T, Menno D, Bron M, et al. Determination of thresholds for minimally important difference and clinically important response on the functional outcomes of sleep questionnaire short version in adults with narcolepsy or obstructive sleep apnea. *Sleep Breath*. 2021;25(3):1701-1715.
- Weinberg ME, Bacchetti P, Rushakoff RJ. Frequently repeated glucose measurements overestimate the incidence of inpatient hypoglycemia and severe hyperglycemia. *J Diabetes Sci Technol*. 2010;4(3):577-582. <https://doi.org/10.1177/193229681000400311>
- [WMA] World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191-2194. <https://doi.org/10.1001/jama.2013.281053>
- [WHO] World Health Organization. WHO STEPS surveillance manual: the WHO STEPwise approach to noncommunicable disease risk factor surveillance. Updated January 26, 2017. Accessed November 04, 2021. [https://www.who.int/ncds/surveillance/steps/STEPS\\_Manual.pdf](https://www.who.int/ncds/surveillance/steps/STEPS_Manual.pdf)
- Woodcock J, LaVange LM. Master protocols to study multiple therapies, multiple diseases, or both. *N Engl J Med*. 2017;377(1):62-70. <https://doi.org/10.1056/NEJMra1510062>

PPD

## Title Page

### Confidential Information

The information contained in this document is confidential and is intended for the use of clinical investigators. It is the property of Eli Lilly and Company or its subsidiaries and should not be copied by or distributed to persons not involved in the clinical investigation of tirzepatide (LY3298176), unless such persons are bound by a confidentiality agreement with Eli Lilly and Company or its subsidiaries.

**Note to Regulatory Authorities:** This document may contain protected personal data and/or commercially confidential information exempt from public disclosure. Eli Lilly and Company requests consultation regarding release/redaction prior to any public release.

In the United States, this document is subject to Freedom of Information Act (FOIA) Exemption 4 and may not be reproduced or otherwise disseminated without the written approval of Eli Lilly and Company or its subsidiaries.

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**Protocol Number:** I8F-MC-GPIF

**Amendment Number:** c

**List of Intervention-Specific Appendices (ISAs):**

I8F-MC-GPI1: Participants with OSA unwilling or unable to use PAP therapy

I8F-MC-GPI2: Participants with OSA on PAP therapy

**Compound:** Tirzepatide (LY3298176)

**Brief Title:** A Master Protocol for Tirzepatide in Participants with Obstructive Sleep Apnea and Obesity

**Study Phase:** 3

**Acronym:** SURMOUNT-OSA

**Sponsor Name:** Eli Lilly and Company

**Legal Registered Address:** Indianapolis, Indiana, USA 46285

**Regulatory Agency Identifier Number:**

IND: 157090

**Approval Date:** Protocol Amendment (c) Electronically Signed and Approved by Lilly on date provided below.

**Document ID:** VV-CLIN-116270

**Medical Monitor Name and Contact Information will be provided separately.**

## Protocol Amendment Summary of Changes Table

DOCUMENT HISTORY	
Document	Date
<i>Amendment b</i>	<i>30-Sep-2022</i>
<i>Amendment a</i>	<i>10-Feb-2022</i>
<i>Original Protocol</i>	<i>27-Jan-2022</i>

### Amendment [c]

This amendment is considered to be nonsubstantial.

#### Overall Rationale for the Amendment:

This amendment includes changes made to the primary and key secondary endpoints for change in AHI and clarification around timing of the PHQ-9 assessment. Revisions were made to clarify the role of central reading of PSG scores.

Protocol changes have been made as outlined in the following table.

Section # and Name	Description of Change	Brief Rationale
1.1. Synopsis and 3. Objectives, Endpoints, and Estimands	Primary Objective: Removed “percent”	To align with change made to primary endpoint. Note, the current sample size is deemed adequate for the assessment of the updated primary endpoint and all key secondary endpoints
	Primary Endpoints: Updated primary endpoint from “Percent change in AHI from baseline to Week 52” to “Change in AHI from baseline to Week 52 (events per hour)”	In response to regulatory recommendation
	Key Secondary Endpoints: Updated first endpoint from “Change in AHI” to “Percent change in AHI”	Changed to reflect revisions in endpoint hierarchy
1.3. Schedule of Activities (SoA)	“PHQ-9” row: Added additional information to Comments: “PHQ-9 at V7 and V11 may be scheduled for any day +/- 14 days.”	Clarification
5.2 Exclusion Criteria	EC#44: Added glucose-lowering medication, including metformin	Clarification

<b>Section # and Name</b>	<b>Description of Change</b>	<b>Brief Rationale</b>
	to the list of medications prohibited within 3 months of Visit 1	
6.8.2 Prohibited Concomitant Medications	Added glucose-lowering medication, including metformin to the list of prohibited concomitant medications	Clarification
8.1.1.1 Polysomnography	Added text, “The eligibility criteria and AHI related endpoints of the study will be assessed based on central reading of the PSG.”	Clarification that PSG based endpoints are based on result determined from central reading of PSG
9.1 Statistical Hypotheses	“Percent” has been removed from the description of the treatment effect definition	To align with changes made in Section 1.1 and Section 3
9.3.1 General Considerations	A statement was added to clarify that efficacy analyses will be conducted on all participants meeting study eligibility criteria	Clarifications added to ensure consistency of text in Section 9.3.1 with the table in Section 9.2
	The first intercurrent event has been updated to include incorrect study procedures leading to invalid measurements	Clarifying how invalid measurements will be handled in the analysis.
9.3.2 Primary Analysis	“Percent” has been removed when describing the primary endpoint	To align with changes made in Section 1.1 and Section 3
9.3.3 Analysis of Key Secondary Endpoints	“Percent” has been added when describing the analysis	To align with changes made in Section 1.1 and Section 3
Throughout	Editorial corrections	Minor, therefore not described

## Table of Contents

<b>1.</b>	<b>Protocol Summary .....</b>	<b>9</b>
1.1.	Synopsis .....	9
1.2.	Schema.....	12
1.3.	Schedule of Activities (SoA).....	14
<b>2.</b>	<b>Introduction.....</b>	<b>24</b>
2.1.	Study Rationale.....	24
2.2.	Background.....	24
2.3.	Benefit/Risk Assessment .....	25
<b>3.</b>	<b>Objectives, Endpoints, and Estimands .....</b>	<b>26</b>
<b>4.</b>	<b>Study Design.....</b>	<b>29</b>
4.1.	Overall Design .....	29
4.2.	Scientific Rationale for Study Design .....	29
4.3.	Justification for Dose .....	30
4.4.	End of Study Definition.....	30
<b>5.</b>	<b>Study Population.....</b>	<b>31</b>
5.1.	Inclusion Criteria .....	31
5.2.	Exclusion Criteria .....	32
5.3.	Lifestyle Considerations .....	36
5.3.1.	Meals and Dietary Restrictions.....	36
5.3.2.	Physical Activity.....	36
5.4.	Screen Failures.....	37
5.5.	Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant.....	37
<b>6.</b>	<b>Study Intervention(s) and Concomitant Therapy .....</b>	<b>38</b>
6.1.	Study Interventions Administered .....	38
6.1.1.	Medical Devices.....	39
6.2.	Preparation, Handling, Storage, and Accountability .....	39
6.3.	Measures to Minimize Bias: Randomization and Blinding.....	39
6.4.	Study Intervention Compliance .....	40
6.5.	Dose Modification .....	40
6.5.1.	Management of Participants with Gastrointestinal Symptoms.....	40
6.6.	Continued Access to Study Intervention after the End of the Study.....	41
6.7.	Treatment of Overdose .....	41
6.8.	Concomitant Therapy .....	41
6.8.1.	Management of Incident Diabetes .....	42
6.8.2.	Prohibited Concomitant Medications .....	43
<b>7.</b>	<b>Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal.....</b>	<b>44</b>
7.1.	Discontinuation of Study Intervention.....	44
7.1.1.	Liver Chemistry Stopping Criteria.....	46
7.1.2.	Temporary Discontinuation .....	46
7.2.	Participant Discontinuation/Withdrawal from the Study.....	47
7.3.	Lost to Follow up.....	48



<b>8.</b>	<b>Study Assessments and Procedures.....</b>	<b>49</b>
8.1.	Efficacy Assessments .....	49
8.1.1.	Primary Efficacy Assessment .....	49
8.1.2.	Secondary Efficacy Assessments.....	49
8.1.3.	Exploratory Efficacy Assessments .....	51
8.2.	Safety Assessments.....	53
8.2.1.	Physical Examinations.....	53
8.2.2.	Vital Signs.....	53
8.2.3.	Electrocardiograms (ECG).....	53
8.2.4.	Clinical Safety Laboratory Tests .....	53
8.2.5.	Pregnancy Testing.....	54
8.2.6.	Suicidal Ideation and Behavior Risk Monitoring .....	54
8.3.	Adverse Events, Serious Adverse Events, and Product Complaints.....	55
8.3.1.	Timing and Mechanism for Collecting Events .....	55
8.3.2.	Pregnancy.....	57
8.3.3.	Adverse Events of Special Interest .....	58
8.4.	Pharmacokinetics .....	59
8.4.1.	Bioanalytical Methods .....	59
8.5.	Pharmacodynamics .....	59
8.6.	Genetics .....	59
8.7.	Biomarkers.....	60
8.8.	Immunogenicity Assessments.....	60
8.9.	Health Economics OR Medical Resource Utilization and Health Economics.....	60
<b>9.</b>	<b>Statistical Considerations.....</b>	<b>61</b>
9.1.	Statistical Hypotheses .....	61
9.1.1.	Multiplicity Adjustment.....	61
9.2.	Analyses Sets .....	61
9.3.	Statistical Analyses .....	62
9.3.1.	General Considerations .....	62
9.3.2.	Primary Analysis.....	64
9.3.3.	Analysis of Key Secondary Endpoints .....	64
9.3.4.	Treatment Group Comparability.....	65
9.3.5.	Safety Analyses.....	65
9.3.6.	Evaluation of Immunogenicity .....	66
9.3.7.	Other Analyses.....	67
9.4.	Interim Analysis.....	67
9.5.	Sample Size Determination .....	68
<b>10.</b>	<b>Supporting Documentation and Operational Considerations.....</b>	<b>69</b>
10.1.	Appendix 1: Regulatory, Ethical, and Study Oversight Considerations .....	69
10.1.1.	Regulatory and Ethical Considerations.....	69
10.1.2.	Financial Disclosure.....	69
10.1.3.	Informed Consent Process .....	70
10.1.4.	Data Protection.....	70

10.1.5.	Committees Structure.....	70
10.1.6.	Dissemination of Clinical Study Data.....	71
10.1.7.	Data Quality Assurance .....	72
10.1.8.	Source Documents .....	73
10.1.9.	Study and Site Start and Closure .....	74
10.1.10.	Publication Policy .....	74
10.1.11.	Investigator Information .....	74
10.1.12.	Sample Retention.....	75
10.2.	Appendix 2: Clinical Laboratory Tests.....	76
10.2.1.	Laboratory Samples to be Obtained at the Time of a Systemic Hypersensitivity Event.....	78
10.3.	Appendix 3: Adverse Events and Serious Adverse Events: Definitions and Procedures for Recording, Evaluating, Follow-up, and Reporting.....	80
10.3.1.	Definition of AE .....	80
10.3.2.	Definition of SAE .....	81
10.3.3.	Definition of Product Complaints.....	82
10.3.4.	Recording and Follow-Up of AE and/or SAE and Product Complaints .....	83
10.3.5.	Reporting of SAEs .....	84
10.3.6.	Regulatory Reporting Requirements.....	85
10.3.7.	Special Safety Topics.....	85
10.4.	Appendix 4: Contraceptive and Barrier Guidance.....	88
10.4.1.	Definitions.....	88
10.4.2.	Contraception Guidance.....	89
10.5.	Appendix 5: Genetics.....	91
10.6.	Appendix 6: Liver Safety: Suggested Actions and Follow-up Assessments.....	92
10.6.1.	Hepatic Evaluation Testing.....	92
10.6.2.	Close Hepatic Monitoring.....	94
10.6.3.	Comprehensive Hepatic Evaluation.....	94
10.7.	Appendix 7: Medical Device Adverse Events (AEs), Adverse Device Effects (ADEs), Serious Adverse Events (SAEs) and Device Deficiencies: Definition and Procedures for Recording, Evaluating, Follow-up, and Reporting.....	96
10.8.	Appendix 8: Protocol GPIF Standardized Protocols for the Measurement of Height, Weight, Neck Circumference, Waist Circumference, Vital Signs, and Electrocardiogram .....	97
10.8.1.	Measuring Height.....	97
10.8.2.	Measuring Weight.....	97
10.8.3.	Measuring Hip and Waist Circumference.....	97
10.8.4.	Measuring Neck Circumference .....	98
10.8.5.	Vital Sign Measurements (Blood Pressure and Heart Rate).....	98
10.8.6.	Electrocardiogram.....	99
10.9.	Appendix 9: Provisions for Changes in Study Conduct Due to the COVID-19 Pandemic.....	100

10.10. Appendix 10: Abbreviations and Definitions .....104  
10.11. Appendix 11: Country-Specific Requirements.....108  
10.11.1. Germany.....108  
10.12. Appendix 12: Protocol Amendment History .....112  
**11. References.....122**

## 1. Protocol Summary

### 1.1. Synopsis

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**Brief Title:** A Master Protocol for Tirzepatide in Participants with Obstructive Sleep Apnea and Obesity

#### **Rationale:**

Obstructive sleep apnea (OSA) is a breathing disorder associated with significant comorbidity and mortality. Currently available therapeutic approaches have shown moderate success in treating the clinical signs and symptoms of OSA (that is, snoring and excessive daytime sleepiness) but have failed to address the underlying pathophysiology of the disease and, more importantly, the cardiovascular (CV) morbidity and mortality associated with OSA.

Tirzepatide is a dual GIP/GLP-1R agonist that has demonstrated statistically significant and clinically relevant lowering in HbA1c and dose-dependent weight loss in 5 Phase 3 trials that enrolled patients with type 2 diabetes mellitus (T2DM) (Dahl et al. 2021; Frias et al. 2021; Lilly 2021; Ludvik et al. 2021; Rosenstock et al. 2021). Data from clinical studies show that a large proportion of participants have significant body weight reduction with tirzepatide treatment, suggesting that this agent may represent a promising pharmacologic treatment option that could reduce the frequency of apnea and hypopnea events as well as reduce weight, decrease blood pressure, and improve insulin resistance and dyslipidemia, which are all features associated with the increase in CV morbidity and mortality as seen in people living with OSA.

**Objectives, Endpoints, and Estimands:**

The following objectives and endpoints apply to both intervention-specific appendices (ISAs).

Primary Objective	Endpoints
To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for mean decrease in AHI.	Change in AHI from baseline to Week 52 (events per hour).
Key Secondary Objectives (controlled for type I error)	Endpoints
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in AHI</li> <li>• A hierarchical assessment of PROs</li> </ul> <ul style="list-style-type: none"> <li>• Clinically meaningful change in AHI</li> <li>• Achieving OSA remission or mild nonsymptomatic OSA</li> <li>• Change in body weight</li> <li>• Change in inflammatory status</li> <li>• Change in SBP</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Percent change in AHI</li> <li>• A hierarchical combination of the following: <ul style="list-style-type: none"> <li>○ Change in FOSQ-10 score</li> <li>○ Change in FOSQ (30 items) Vigilance domain score</li> <li>○ Change in FOSQ (30 items) Activity Level domain score</li> </ul> </li> <li>• Percent of participants with <math>\geq 50\%</math> AHI reduction</li> <li>• Percent of participants with <ul style="list-style-type: none"> <li>○ AHI <math>&lt; 5</math> or</li> <li>○ AHI 5-14 with ESS <math>\leq 10</math></li> </ul> </li> <li>• Percent change in body weight</li> </ul> <p>From baseline to Week 48<sup>a</sup></p> <ul style="list-style-type: none"> <li>• Change in hsCRP concentration</li> <li>• Change in SBP</li> </ul>

Abbreviations: AHI = Apnea-Hypopnea Index; ESS = Epworth Sleepiness Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; hsCRP = high-sensitivity C reactive protein; MTD = maximum-tolerated dose; OSA = obstructive sleep apnea; PROs = patient-reported outcomes; SBP = systolic blood pressure; QW = once weekly.

<sup>a</sup> BP will be assessed at Week 48 because PAP suspension at Week 52 may confound BP assessment.

For estimands guiding statistical analyses, see Section 9.3.1.

**Overall Design:**

The overall study design consists of 2 components:

- Master Protocol: defines study elements common for both populations.
- ISAs: provide detailed population-specific information.

Study I8F-MC-GPIF (GPIF) is multicenter, randomized, parallel-arm, double-blind, placebo-controlled, Phase 3 study with 52-week treatment duration conducted under a basket-design, which will investigate the effects of treatment with weekly (QW) tirzepatide at the maximum

tolerated dose (MTD) (10 mg or 15 mg), compared with placebo in participants who have moderate-to-severe OSA and obesity.

One master protocol will support 2 studies/ISAs.

- ISA 1 (GPI1) will include participants who are unwilling or unable to use PAP therapy.
- ISA 2 (GPI2) will include participants who are on PAP therapy for at least 3 months at time of screening and plan to continue PAP therapy during the study.

Participants will be assigned to the ISA which reflects their current PAP usage. The participant will then be randomly assigned 1:1 to treatment or placebo.

**Number of Participants:**

Approximately 412 participants will be randomly assigned to study intervention across the entire master protocol, with approximately 206 participants randomly assigned to study intervention in each ISA. See Section 9.5 for additional information.

An upper limit of approximately 70% enrollment of male participants will be used to ensure a sufficiently large sample of female participants.

**Intervention Groups and Duration:**

The study interventions are:

- tirzepatide at the MTD (10 mg or 15 mg) SC QW, or
- placebo.

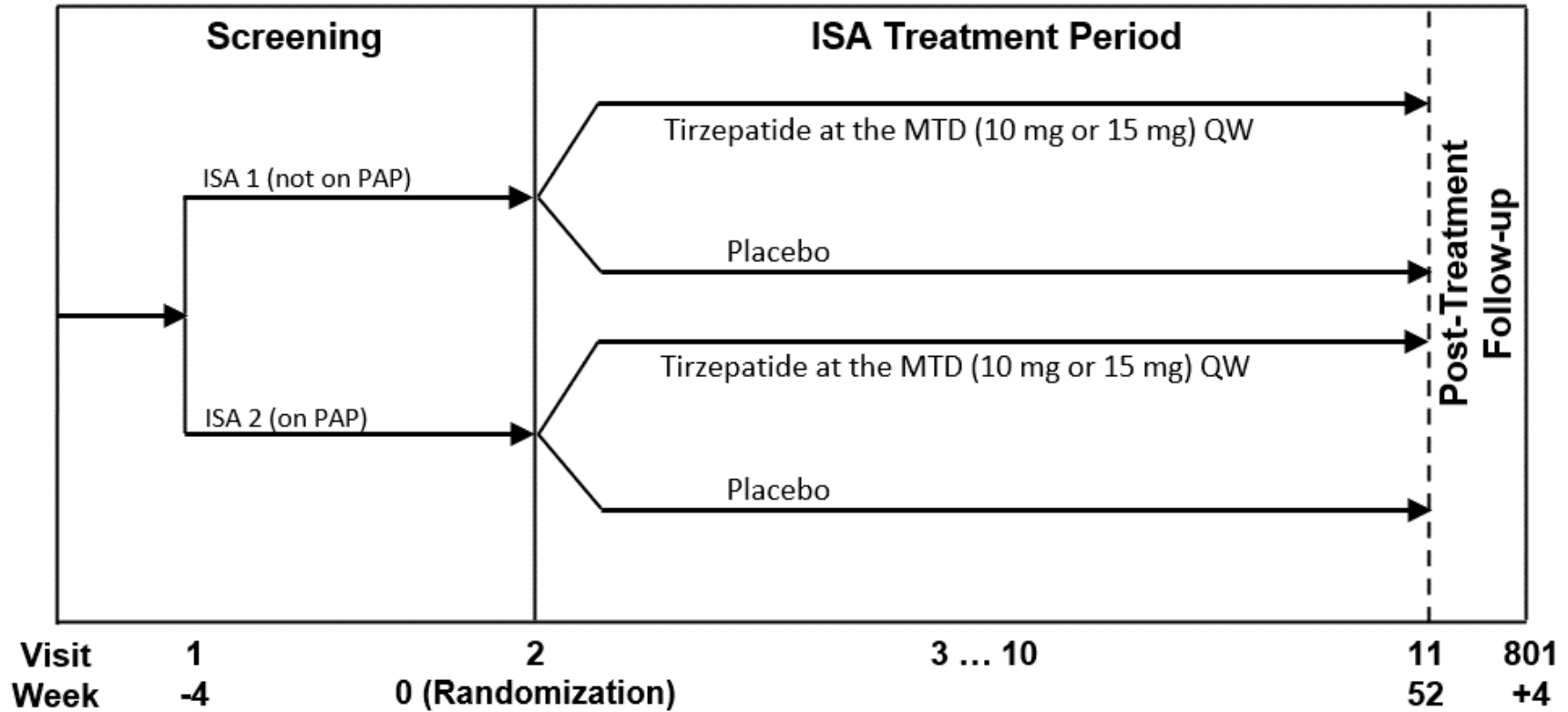
The expected total duration of study participation for each participant, including screening and the post-treatment follow-up periods, is 60 weeks across the following study periods:

- Screening: 4 weeks
- Treatment: 52 weeks
- Post-treatment follow-up: 4 weeks

The maximum duration of treatment is 52 weeks.

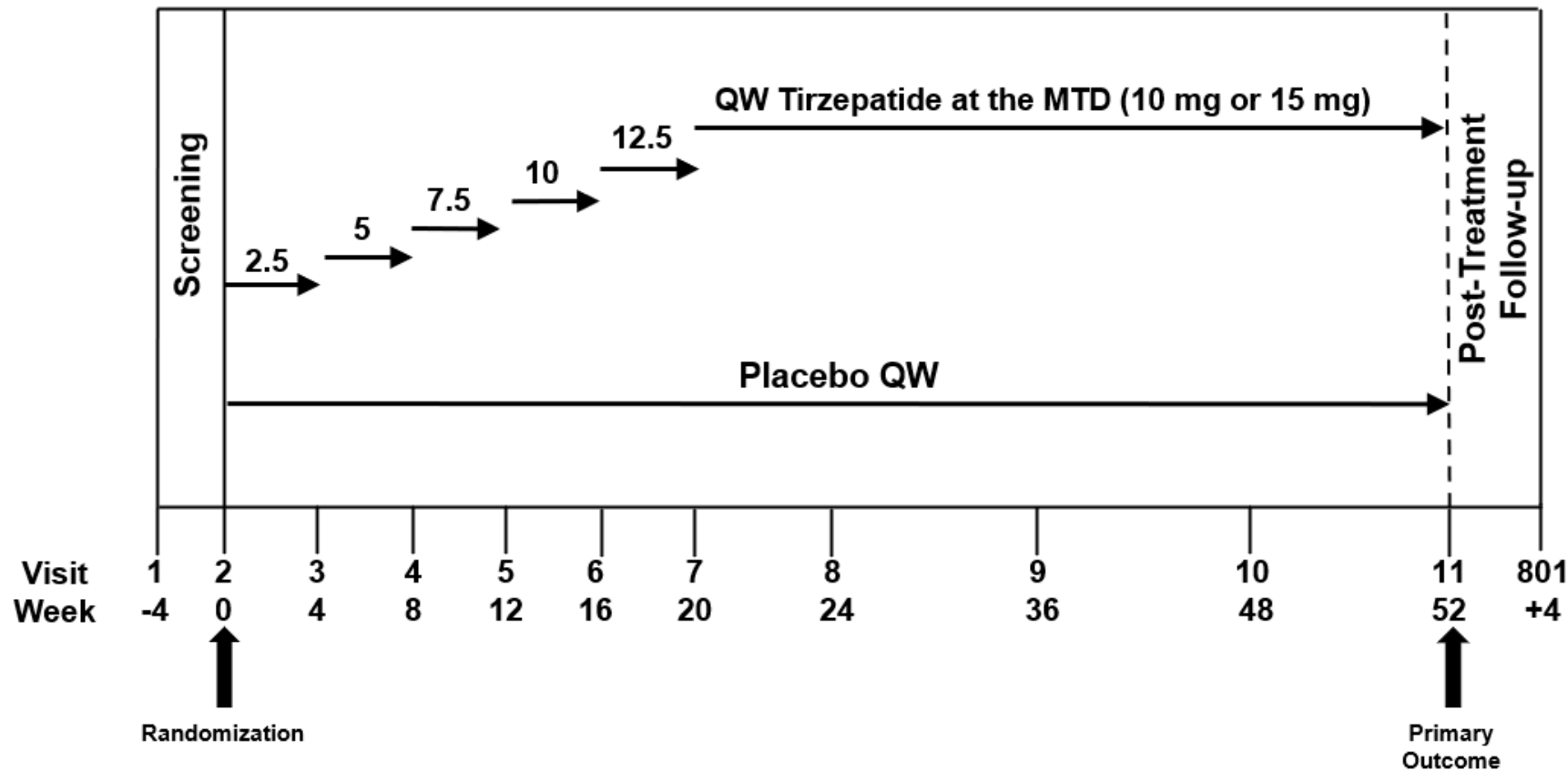
**Data Monitoring Committee: Yes**

**1.2. Schema**  
**Master Protocol and ISA Schema**



Abbreviations: ISA = intervention-specific appendix; MTD = maximum-tolerated dose; PAP = positive airway pressure.

**Dose Escalation and Visit Schema**



Abbreviations: MTD = maximum-tolerated dose; QW = once weekly.



### **1.3. Schedule of Activities (SoA)**

If Study Period I - Screening takes longer or shorter than 4 weeks to complete, it will not be considered a protocol deviation. Study Period I - Screening should not exceed 6 weeks unless approval from the sponsor is received.

Visit procedures may be conducted over more than 1 day as long as all activities are completed within the allowable visit tolerance period of each visit.

Study Period II - Treatment: For early discontinuations (ED) from study that occur before the last visit in treatment period, see the activities listed for ED in this table.

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
		2	3	4	5	6	7	8	9	10	11	ED		
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b.
Informed consent	X													The informed consent must be signed before any protocol-specific tests/procedures are performed.
Inclusion and exclusion criteria, review and confirm	X	X												Inclusion/Exclusion criteria should be confirmed prior to drug assignment and administration of first dose of study intervention.
Demographics	X													Includes ethnicity, year of birth, gender (sex), and race.
Preexisting conditions and medical history, including relevant surgical history	X													All conditions ongoing and relevant past surgical and medical history should be collected.
Prespecified medical history	X													Should include, but not be limited to, collecting diagnosis of OSA and obesity-related health problems.
Prior treatments for indication	X													Include OSA and obesity.
Substance use (alcohol, caffeine, tobacco use)	X													
Concomitant medications	X	X	X	X	X	X	X	X	X	X	X	X	X	
AEs	X	X	X	X	X	X	X	X	X	X	X	X	X	Any events that occur after signing the informed consent are considered AEs as defined in Section 8.3.1. Additional data are collected for certain AEs.

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b.
<b>Physical Evaluation</b>														
Height	X													
Weight	X	X	X	X	X	X	X	X	X	X	X	X	X	Weight measurements should be obtained per the instructions in Section 10.8.
Waist circumference		X					X				X	X	X	Waist circumference should be obtained per the instructions in Section 10.8.
Hip circumference		X					X				X	X	X	Hip circumference should be obtained per the instructions in Section 10.8.
Neck circumference		X					X				X	X	X	Neck circumference should be obtained per the instructions in Section 10.8.
Vital signs	X	X	X	X	X	X	X	X	X	X	X	X	X	Includes PR and BP. Measured after participant has been sitting at least 5 minutes. Vital-sign measurements should be taken before obtaining an ECG tracing and before collection of blood samples for laboratory testing. See Sections 8.2.2 and 10.8.
Complete physical examination	X													The complete physical examination is performed (excludes pelvic, rectal, and breast examinations unless clinically indicated).
Symptom-directed physical assessment		X (refer to Comment)												Symptom-directed physical assessment from V2 through V11 and ED will be conducted at the discretion of the PI, as indicated based on participant status and standard of care.

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
		2	3	4	5	6	7	8	9	10	11	ED		
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b.
														May be performed by qualified personnel per local regulations.
12-lead ECG (local)	X										X	X		ECG measurements should be obtained per the instructions in Section 8.2.3.
<b>Wearable Devices and PSG Assessments</b>														
Schedule Sleep Center Study for PSG	X					X				X				PSG results must be reviewed to confirm eligibility prior to randomization.
Sleep Center Study for PSG	X						X				X			PSG at V7 and 11 may be scheduled for any day +/- 14 days.
Participant wears the WatchPat300	X		X		X		X				X			Applicable only to participants in GPII (off PAP). Training documents will detail information on the dispensation, wearing, and return process.
Participant wears actigraphy (AX6) device	X		X		X		X				X			Training documents will detail information on the dispensation, wearing, and return process.
<b>Participant Education and Supplies</b>														
eDiary education	X													Additional training can be repeated, as needed.

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
		2	3	4	5	6	7	8	9	10	11	ED		
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b.
Train participant on study intervention administration		X												Re-training is available anytime.
Review Lifestyle Program instructions		X	X	X	X	X	X	X	X	X	X	X	X	Diet and exercise goals established during the lifestyle consultation and the importance of adherence to the lifestyle component of the trial will be reinforced at each trial contact by study staff.
<b>Participant Diary</b>														
Participant diary dispensed	X													Includes the following: Electronic diaries: Sleep diary; Dosing diary; Hypoglycemic events (when applicable), PAP adherence (for GPI2 only). Paper diary: Patient Diet and Exercise diary
Diary compliance check		X	X	X	X	X	X	X	X	X	X	X		
Diary return											X	X		
<b>Patient-Reported Outcomes (Electronic)</b> <i>Complete prior to any clinical-administered assessments</i>														
ESS	X		X		X		X				X	X		When the PROs are scheduled for visits at which the PSG will be done, they should be completed on the same day as PSG and in the following order (FOSQ, ESS, PROMIS Short Form v1.0 Sleep Disturbance 8b, PROMIS Short Form v1.0 Sleep-related Impairment 8a, PGIS, PGIC, SF-36v2 acute form, and EQ-5D-
EQ-5D-5L	X										X	X		
FOSQ	X				X		X				X	X		
PGIS (OSA Symptom Scales)	X		X		X		X				X	X		

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
		2	3	4	5	6	7	8	9	10	11	ED		
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b.
														5L). Completing the PRO assessments on the next day after PSG is not considered a protocol violation.
PGIC (OSA Symptom Scales)			X		X		X				X	X		
PROMIS Short Form v1.0 Sleep Disturbance 8b	X		X		X		X				X	X		
PROMIS Short Form v1.0 Sleep-related Impairment 8a	X		X		X		X				X	X		
SF-36v2, acute	X						X				X	X		
PHQ-9	X	X					X				X	X		The PHQ-9 should be administered after assessment of AEs, if both collected on the same day. PHQ-9 at V7 and V11 may be scheduled for any day +/- 14 days.
<b>Clinician-Administered Assessments (Paper)</b>														
C-SSRS screening/baseline	X													The C-SSRS should be administered after assessment of AEs, if both collected on the same day. The C-SSRS since last visit is adapted for the assessment of the ideation and behavior categories only. The Intensity of Ideation and Lethality of Behavior sections are removed.
C-SSRS (since last visit version)		X					X				X	X		
<b>Laboratory Tests and Sample Collections</b>														

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
		2	3	4	5	6	7	8	9	10	11	ED		
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b.
Hematology	X				X			X	X		X	X	X	
HbA1c	X	X			X			X	X		X	X	X	
Clinical chemistry (includes glucose)	X	X			X			X	X		X	X	X	
Lipid panel		X						X			X	X	X	
hsCRP		X						X			X	X		
Serum pregnancy	X													Only for WOCBP and females with a history of tubal ligation. See Section 10.4, Appendix 4.
Urine pregnancy (local)		X			X			X	X		X	X		A urine pregnancy test must be performed at V2 with the result available prior to first dose/injection of study intervention for WOCBP only. Additional pregnancy tests (beyond those required per the SoA) should be performed at any time during the trial if a menstrual period is missed, there is clinical suspicion of pregnancy, or as required by local law or regulation.
FSH	X													Optional; performed as needed to confirm postmenopausal status. See Section 10.4, Appendix 4.
Insulin		X						X			X	X	X	
C-Peptide		X						X			X	X	X	
Free fatty acids		X						X			X	X	X	

	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
		2	3	4	5	6	7	8	9	10	11	ED		
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b.
Cystatin-C	X							X			X	X	X	
Calcitonin	X										X	X	X	
Pancreatic amylase	X				X			X			X	X	X	
Lipase	X				X			X			X	X	X	
TSH	X													
eGFR	X	X			X			X	X		X	X	X	Calculated using CKD-EPI method.
UACR	X	X			X			X	X		X	X	X	
PK samples		X	X		X			X			X	X	X	PK samples to be collected predose and close to ADA samples.
Immunogenicity (ADA) samples		X	X		X			X			X	X	X	In the event of systemic drug hypersensitivity reactions (immediate or nonimmediate), additional unscheduled samples should be collected as detailed in Section 10.3.7.2 (Hypersensitivity Reactions). Immunogenicity samples and PK samples for immunogenicity must be predose.
<b>Stored Samples</b>														
Genetics sample		X												Sample can be obtained at or after the specified visit.
Exploratory biomarker samples		X						X			X	X	X	
<b>Randomization and Dosing</b>														
Register visit with IWRS	X	X	X	X	X	X	X	X	X	X	X	X	X	



	Study Period I - Screening	Study Period II - Treatment											Study Period III - Post-Treatment Follow-up	Comments
		2	3	4	5	6	7	8	9	10	11	ED		
Visit number	1	2	3	4	5	6	7	8	9	10	11	ED	801	
Weeks from randomization	-4	0	4	8	12	16	20	24	36	48	52	-	See footnote a	
Visit interval tolerance (days)	-14 to +21	-	±3	±3	±3	±3	±7	±3	±7	±3	±7	±7	±3	
Fasting Visit		X	X	X	X	X	X	X	X	X	X	X	X	See footnote b.
ISA assignment via IWRS	X													
ISA treatment randomization via IWRS		X												
Observe participant administer study intervention		X												Participants should administer their first dose of study intervention at the end of the V2, after other study procedures are completed.
Dispense study drug via IWRS		X	X	X	X	X	X	X	X	X				
Dispense study drug to participant (for at home dosing)		X	X	X	X	X	X	X	X	X				
Dispense ancillary supplies to participant		X												Dispensation of ancillary supplies may vary beyond V2 based on expiry dating of applicable supplies and/or participant needs.
Participant returns all unused study intervention			X	X	X	X	X	X	X	X	X	X		
Assess study intervention compliance			X	X	X	X	X	X	X	X	X	X		

Abbreviations: ADA = antidrug antibody; AEs = adverse events; AX6 = Axivity 6; CKD EPI = Chronic Kidney Disease Epidemiology; hsCRP = high sensitivity C-reactive protein; C-SSRS = Columbia-Suicide Severity Rating Scale; ECG = electrocardiogram; ED = early discontinuation; eGFR = estimated glomerular filtration rate; EQ-5D-5L = EuroQol - 5 Dimension - 5 Level; ESS = Epworth Sleepiness Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; FSH = follicle-stimulating hormone; HbA1c = hemoglobin A1c; ISA = intervention-specific appendix; IWRS = interactive web-response system; OSA = obstructive sleep apnea; PHQ = Patient Health Questionnaire-9; PGIC = Patient Global Impression of Change; PGIS = Patient Global Impression of Status; PK = pharmacokinetic; PSG = polysomnography; PR = pulse rate; PRO = patient reported outcome; PROMIS = Patient-Reported Outcomes Measurement Information System; SF-36v2 = Short-Form 36 version 2; SoA = schedule of activities; TSH = thyroid-stimulating hormone; UACR = urinary albumin/creatinine ratio; V = visit; WOCBP = woman of childbearing potential.

- a Post-treatment follow-up occurs approximately 4 weeks after the participant's final treatment period visit.
- b Fasting visit: On all office visits, study participants should be reminded to report to the site before taking study intervention in a fasting condition, after a period of approximately 8 hours without eating, drinking (except minimal amount of water, as needed), or any significant physical activity.

## **2. Introduction**

### **2.1. Study Rationale**

Obstructive sleep apnea is a breathing disorder associated with significant comorbidity and mortality. Currently available therapeutic approaches have shown moderate success in treating the clinical signs and symptoms of OSA (that is, snoring and excessive daytime sleepiness) but have failed to address the underlying pathophysiology of the disease and, more importantly, the CV morbidity and mortality associated with OSA.

Tirzepatide is a dual GIP/GLP-1R agonist that has demonstrated statistically significant and clinically relevant lowering in HbA1c and dose-dependent weight loss in 5 Phase 3 trials that enrolled patients with T2DM (Dahl et al. 2021; Frias et al. 2021; Lilly 2021; Ludvik et al. 2021; Rosenstock et al. 2021). Data from clinical studies show that a large proportion of participants have significant body weight reduction with tirzepatide treatment, suggesting that this agent may represent a promising pharmacologic treatment option that could reduce the frequency of apnea and hypopnea events as well as reduce weight, decrease BP, and improve insulin resistance and dyslipidemia, which are all features associated with the increase in CV morbidity and mortality as seen in people living with OSA.

### **2.2. Background**

Obstructive sleep apnea is a serious medical condition with a limited number of therapeutic options and high unmet need. OSA is a well-established risk factor for morbidity and mortality from CV and other metabolic diseases; further, it negatively affects daily life of patients (Tietjens et al. 2019; Gottlieb and Punjabi 2020).

Treatments for OSA include behavioral measures like weight loss programs, medical devices such as PAP, oral appliances or hypoglossal nerve stimulation, and surgery, including bariatric surgery in people living with obesity and OSA. Weight loss through caloric restriction and lifestyle interventions improves AHI, cardiometabolic comorbidities, and quality of life. Although a limited number of studies have investigated the effect of pharmacological weight-loss therapy, their results support that adding weight-loss medication to lifestyle intervention reduces AHI and improves sleep quality and possibly other OSA-related outcomes. As such, standard of care recommends a prioritization of weight loss for patients with OSA and obesity or who are overweight, including consideration for FDA-approved antiobesity pharmacotherapy (Hudgel et al. 2018). Bariatric surgery is available for patients with severe obesity and can dramatically improve OSA (Currie et al. 2021). However, such surgical procedures are associated with greater risk for perioperative and postoperative complications (Opperer et al. 2016).

Positive airway pressure is the first-line treatment for moderate-to-severe and symptomatic mild OSA. Despite good clinical efficacy for the signs and symptoms of OSA, randomized controlled studies in PAP therapy have failed to show improvements in nonsleep-related outcomes linked to OSA such as stroke, heart attack, diabetes and depression (AHRQ 2021). It is thought inadequate PAP therapy adherence may explain why randomized clinical trials have failed to demonstrate benefit (Gottlieb and Punjabi 2020). Nonadherence to PAP therapy is a major clinical challenge

and studies have shown wide variability in adherence, ranging from 29% to 83% (Weaver and Grunstein 2008). Currently, there are only few pharmacologic alternatives indicated for symptoms in patients with OSA and none are disease modifying.

There is significant unmet need in treatment of patients with OSA. Tirzepatide, a GIP and GLP-1R dual agonist, has the potential to provide benefit to patients with OSA who have obesity by reducing weight, which may reduce the frequency of apnea and hypopnea events, decreasing BP, and improving insulin resistance and dyslipidemia.

### **2.3. Benefit/Risk Assessment**

More detailed information about the known and expected benefits and risks and reasonably expected AEs of tirzepatide may be found in the IB.

### 3. Objectives, Endpoints, and Estimands

The following objectives and endpoints apply to both ISAs.

Primary Objective	Endpoints
To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for mean decrease in AHI.	Change in AHI from baseline to Week 52 (events per hour).
Key Secondary Objectives (controlled for type I error)	Endpoints
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in AHI</li> <li>• A hierarchical assessment of PROs</li>   <li>• Clinically meaningful change in AHI</li> <li>• Achieving OSA remission or mild nonsymptomatic OSA</li> <li>• Change in body weight</li>   <li>• Change in inflammatory status</li> <li>• Change in SBP</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Percent change in AHI</li> <li>• A hierarchical combination of the following: <ul style="list-style-type: none"> <li>○ Change in FOSQ-10 score</li> <li>○ Change in FOSQ (30 items) Vigilance domain score</li> <li>○ Change in FOSQ (30 items) Activity Level domain score</li> </ul> </li> <li>• Percent of participants with <math>\geq 50\%</math> AHI reduction</li> <li>• Percent of participants with <ul style="list-style-type: none"> <li>○ AHI <math>&lt; 5</math> or</li> <li>○ AHI 5-14 with ESS <math>\leq 10</math></li> </ul> </li> <li>• Percent change in body weight</li>   <li>• Change in hsCRP concentration</li> </ul> <p>From baseline to Week 48<sup>a</sup></p> <ul style="list-style-type: none"> <li>• Change in SBP</li> </ul>
Other Secondary Objectives	Endpoints
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in excessive daytime sleepiness</li> <li>• Change in patient-reported functional status as assessed by FOSQ (30 items)</li> <li>• Change in body weight</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in ESS score</li> <li>• Change in all other FOSQ domain scores</li>   <li>• Percent of participants who achieve <ul style="list-style-type: none"> <li>○ <math>\geq 10\%</math> body weight reduction</li> <li>○ <math>\geq 15\%</math> body weight reduction</li> <li>○ <math>\geq 20\%</math> body weight reduction</li> </ul> </li> </ul>

<ul style="list-style-type: none"> <li>• Change in lipid parameters</li>   <li>• Change in PROs</li>   <li>• Insulin</li> <li>• Hypoxic burden</li> <li>• Change in DBP</li> </ul>	<ul style="list-style-type: none"> <li>• Change in             <ul style="list-style-type: none"> <li>○ HDL-cholesterol</li> <li>○ non-HDL-cholesterol</li> <li>○ triglycerides</li> </ul> </li>   <li>• Change in:             <ul style="list-style-type: none"> <li>○ PROMIS Sleep-related impairment short form 8a score</li> <li>○ PROMIS Sleep disturbance short form 8b score</li> <li>○ SF-36v2 acute form domain scores</li> </ul> </li> <li>• Percent of participants with improved categorical shift in:             <ul style="list-style-type: none"> <li>○ PGIS-OSA Sleepiness</li> <li>○ PGIS-OSA Fatigue</li> <li>○ PGIS-OSA Snoring</li> </ul> </li> <li>• Change in fasting insulin</li> <li>• Change in SASHB (% min/hour)</li> </ul> <p>From baseline to Week 48<sup>a</sup></p> <ul style="list-style-type: none"> <li>• Change in DBP</li> </ul>
<p><b>Exploratory Objectives</b></p>	<p><b>Endpoints</b></p>
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in PROs</li>   <li>• To evaluate the effect of tirzepatide on sleep parameters as measured by Actigraphy (AX6)</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in             <ul style="list-style-type: none"> <li>○ EQ-5D-5L utility index</li> <li>○ EQ-VAS scores</li> </ul> </li> <li>• Percent of participants with improved categorical shift in:             <ul style="list-style-type: none"> <li>○ PGIC-OSA Sleepiness</li> <li>○ PGIC-OSA Fatigue</li> <li>○ PGIC-OSA Sleep quality</li> <li>○ PGIC-OSA Snoring</li> </ul> </li> <li>• Mean change from baseline to endpoint assessment in             <ul style="list-style-type: none"> <li>○ Daytime sleep duration</li> <li>○ Daily step counts</li> <li>○ Average acceleration</li> </ul> </li> </ul>

Abbreviations: AHI = Apnea-Hypopnea Index; AX6 = Axivity 6; DBP = diastolic blood pressure; ESS = Epworth Sleepiness Scale; EQ-5D-5L = EuroQol - 5 Dimension - 5 Level; EQ-VAS = EuroQol Visual Analogue Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; HDL = high-density lipoprotein; hsCRP = high-sensitivity C reactive protein; MTD = maximum-tolerated dose; OSA = obstructive sleep apnea; PGIC-OSA = Patient Global Impression of Change – Obstructive Sleep Apnea; PGIS-OSA = Patient Global Impression of Status – Obstructive Sleep Apnea; PROs = patient-reported outcomes; PROMIS = Patient-Reported Outcomes Measurement Information System; SBP = systolic blood pressure; SF-36v2 = Short-Form 36 version 2, SASHB = sleep apnea-specific hypoxic burden; QW = once weekly.

<sup>a</sup> BP will be assessed at Week 48 because PAP suspension at Week 52 may confound BP assessment.

For estimands guiding statistical analyses, see Section 9.3.1.

## 4. Study Design

### 4.1. Overall Design

Study I8F-MC-GPIF (GPIF) is a multicenter, randomized, parallel-arm, double-blind, placebo-controlled Phase 3 study to evaluate the efficacy and safety of tirzepatide at the MTD (10 mg or 15 mg) QW versus placebo in participants who have obesity and moderate-to-severe OSA.

This basket-type master protocol will investigate 2 participant populations, described in 2 ISAs:

- GPI1 will include participants who are unwilling or are unable to use PAP therapy.
- GPI2 will include participants who are on PAP therapy for at least 3 months at time of screening and plan to continue PAP therapy during the study.

Participants to be assigned to whichever ISA they qualify for. Participants will then be randomly assigned to:

- tirzepatide at the MTD (10 mg or 15 mg) SC QW, or
- placebo.

The expected total duration of study participation for each participant, including screening and the post-treatment follow-up periods, is 60 weeks across the following study periods:

- Screening: 4 weeks
- Treatment: 52 weeks
- Post-treatment follow-up: 4 weeks

The maximum duration of treatment is 52 weeks. Procedures and assessments for each visit are presented in the SoA, Section 1.3.

### 4.2. Scientific Rationale for Study Design

The basket trial design employs a single overarching trial structure as a means to implement multiple investigations (Woodcock and LaVange 2017). Such approaches often allow for more streamlined and coordinated clinical trial logistics and consistency in data collection methods.

The 2 studies associated with the master protocol, also called “Intervention-Specific Appendices” (ISAs), will describe any objectives, endpoints and efficacy assessments specific to each study population. ISA refers to specific population and background intervention, as defined by inclusion criteria. Used together, the master protocol and the 2 ISAs will describe the investigations to be conducted.

ISA 1 and ISA 2 represent different populations with different treatment needs, and it is anticipated that tirzepatide may meet the needs of each participant group.

Inclusion of a placebo treatment arm is acceptable because there are no currently effective disease-modifying treatments for OSA; this approach is in agreement with the use of placebo described in the Declaration of Helsinki (WMA 2013). The use of a placebo comparator in Study GPIF is needed to determine the efficacy and safety of tirzepatide therapy.



### 4.3. Justification for Dose

Tirzepatide doses of up to 15 mg administered SC QW will be evaluated in this study. Participants may be treated with lower maintenance dose of 10 mg if they do not achieve full dose escalation to 15 mg and/or do not tolerate 15 mg.

These doses and associated escalation schemes were selected based on assessment of safety, efficacy (weight loss), and GI tolerability data in Phase 1, 2, and 3 studies in patients with T2DM, followed by exposure-response modeling of the data that predicted weight loss in patients with overweight or obesity.

Dosing algorithms starting at a low dose of 2.5 mg accompanied by dose escalation of 2.5 mg increments every 4 weeks should permit time for development of tolerance to GI events and are predicted to minimize GI tolerability concerns.

Similar to the GLP-1RA class, most of the tirzepatide AEs were dose-dependent and GI-related, consisting mainly of nausea, vomiting, and diarrhea. In general, these events were mild or moderate in severity, with few severe episodes, and transient.

Tirzepatide doses of 10 mg and 15 mg as MTD were selected based principally on the following criteria:

- each dose provides robust weight loss relative to placebo
- the percent of participants achieving  $\geq 10\%$  weight loss is higher with 15 mg than 10 mg, and
- safety and tolerability were supported by Phase 3 results in T2DM.

The proposed tirzepatide maintenance dose of up to 15 mg using a dose-escalation regimen is expected to safely maximize the potential for weight loss while minimizing GI tolerability concerns as has been achieved for approved GLP-1 RA drugs. The subsequent benefit of weight loss should result in clinically meaningful improvements in AHI.

### 4.4. End of Study Definition

A participant is considered to have completed the study if they have completed all required phases of the study including the last visit or the last scheduled procedure shown in the SoA. The end of the study for each ISA is defined as the date of the last visit of the last participant in the study or last scheduled procedure shown in the SoA for the last participant in the trial globally. The end of study timing for each ISA is independent of the other ISA end of study timing.

## 5. Study Population

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

### 5.1. Inclusion Criteria

Participants are eligible to be included in the study only if all of the following criteria apply:

#### Age

1. Participant must be at least 18 (or the legal age of consent in the jurisdiction in which the study is taking place) years of age, inclusive, at the time of signing the informed consent.

#### Type of Participant and Disease Characteristics

2. Previously diagnosed moderate-to-severe OSA with an AHI  $\geq 15$ , as diagnosed with PSG, home sleep apnea test (HSAT), or other method that meets local guidelines prior to Visit 1. See Section 10.10 for definitions of apnea and hypopnea.
3. AHI  $\geq 15$  on PSG as part of the trial at Visit 1.
4. In the investigator's opinion, are well-motivated, capable, and willing to
  - learn how to self-inject study intervention, as required for this protocol (visually impaired persons who are not able to perform the injections must have the assistance of a sighted individual trained to inject study intervention; persons with physical limitations who are not able to perform the injections must have the assistance of an individual trained to inject study intervention)
  - inject study intervention (or receive an injection from a trained individual if visually impaired or with physical limitations), and
  - follow study procedures for the duration of the study, including, but not limited to: follow lifestyle advice (for example, dietary restrictions and exercise plan), maintain a study diary, and complete required questionnaires.

#### Weight

5. BMI  $\geq 30$  kg/m<sup>2</sup>.
6. Have a history of at least 1 self-reported unsuccessful dietary effort to lose body weight.

#### Sex and Contraceptive/Barrier Requirements

7. Males and females may participate in this trial.

Female participants must not be pregnant, intending to be pregnant, breastfeeding, or intending to breastfeed.

Contraceptive use by participants should be consistent with local regulations regarding the methods of contraception for those participating in clinical studies. For definitions and the contraception requirements of this protocol, see Appendix 4 (Section 10.4) .

#### Informed Consent

8. Capable of giving signed informed consent as described in Appendix 1 (Section 10.1), which includes compliance with the requirements and restrictions listed in the ICF and in this protocol.

## 5.2. Exclusion Criteria

Participants are excluded from the study if any of the following criteria apply:

### Medical Conditions

#### *Diabetes-related*

9. Have T1DM or T2DM, history of ketoacidosis, or hyperosmolar state/coma.
10. HbA1c  $\geq$  6.5% ( $\geq$  48 mmol/mol) at Visit 1.

#### *OSA-related*

11. Any previous or planned surgery for sleep apnea or major ear, nose or throat surgery, including tonsillectomy and adenoidectomy that still may affect breathing at time of Visit 1. Inclusion of a participant with more minor ear, nose or throat surgery (for example, deviated septum) will be at the investigator's discretion.
12. Significant craniofacial abnormalities that may affect breathing at time of Visit 1.
13. Diagnosis of Central or Mixed Sleep Apnea with % of mixed or central apneas/hypopneas  $\geq$ 50%, or diagnosis of Cheyne Stokes Respiration.
14. Diagnosis of Obesity Hypoventilation Syndrome or daytime hypercapnia.
15. Active device treatment of OSA other than PAP therapy (for example, dental appliance), or other treatment, that in the opinion of the investigator, may interfere with study outcomes, unless willing to stop treatment at Visit 1 and throughout the study.
16. Respiratory and neuromuscular diseases that could interfere with the results of the trial in the opinion of the investigator.

#### *Obesity-related*

17. Have a self-reported change in body weight  $>$ 5 kg within 3 months prior to screening.
18. Have a prior or planned surgical treatment for obesity (excluding liposuction or abdominoplasty if performed more than 1 year prior to screening).
19. Have or plan to have endoscopic and/or device-based therapy for obesity or have had device removal within the last 6 months (for example, mucosal ablation, gastric artery embolization, intragastric balloon, and duodenal-jejunal bypass sleeve).

#### *Other medical*

20. History of clinically relevant medical, behavioral, or psychiatric disorder, other than OSA, that is associated with insomnia or excessive sleepiness.
21. Impaired renal function, defined as eGFR  $<$ 30 mL/min/1.73 m<sup>2</sup>.
22. Have a known clinically significant gastric emptying abnormality (for example, severe gastroparesis or gastric outlet obstruction) or chronically take drugs that directly affect GI motility.
23. History of chronic or acute pancreatitis.
24. Thyroid-stimulating hormone outside of the range of 0.4 to 6.0 mIU/L at the screening visit.

**Note:** Participants receiving treatment for hypothyroidism may be included, provided their thyroid hormone replacement dose has been stable for at least 3 months.

**Note:** TSH values above the normal range can, in some patients, suggest subclinical hypothyroidism. If, in the investigator's opinion, the participant has subclinical

hypothyroidism and may require initiation of thyroid hormone replacement during the course of the study, the participant should be excluded from the study.

25. Have obesity induced by other endocrinologic disorders (for example, Cushing Syndrome) or diagnosed monogenetic or syndromic forms of obesity (for example, Melanocortin 4 Receptor deficiency or Prader-Willi Syndrome).
26. Are, in the judgment of the investigator, actively suicidal and therefore deemed to be at significant risk for suicide.
27. Have answered “yes” to either Question 4 or Question 5 on the “Suicidal Ideation” portion of the C-SSRS **or** have answered “yes” to any of the suicide-related behaviors on the “suicidal behavior” portion of the C-SSRS, **and** the ideation or behavior occurred within the past month.
28. PHQ-9 score of 15 or more at Visit 1 or 2, prior to randomization.
29. Uncontrolled hypertension (SBP  $\geq$ 160 mmHg and/or DBP  $\geq$ 100 mmHg) at Visit 1.
30. Any of the following CV conditions less than 3 months prior to randomization: acute MI, cerebrovascular accident (stroke), unstable angina, or hospitalization due to congestive heart failure.
31. History of (less than 3 months prior to Visit 1) or planned CV procedure.
32. Heart failure, including New York Heart Association Functional Classification Class IV
33. Have acute or chronic hepatitis, signs and symptoms of any other liver disease other than nonalcoholic fatty liver disease, or any of the following, as determined by the central laboratory during screening:
  - ALT level  $>3.0X$  the ULN for the reference range
  - ALP level  $>1.5X$  the ULN for the reference range, or
  - TBL level  $>1.2X$  the ULN for the reference range (except for cases of known Gilbert’s Syndrome).

*Note:* Participants with nonalcoholic fatty liver disease are eligible to participate in this trial if their ALT level is  $\leq 3.0X$  the ULN for the reference range.

34. Have a calcitonin level (at Visit 1) of:
  - a.  $\geq 20$  ng/L at Visit 1, if eGFR  $\geq 60$  mL/min/1.73 m<sup>2</sup>
  - b.  $\geq 35$  ng/L at Visit 1, if eGFR  $< 60$  mL/min/1.73 m<sup>2</sup>
35. Have a family or personal history of medullary thyroid carcinoma or multiple endocrine neoplasia syndrome type 2.
36. Have a history of an active or untreated malignancy or are in remission from a clinically significant malignancy (other than basal- or squamous-cell skin cancer, in situ carcinomas of the cervix, or in situ prostate cancer) for less than 5 years.
37. Have any other condition not listed in this section (for example, hypersensitivity or intolerance) that is a contraindication to GLP-1R agonists.
38. Have a history of any other condition (such as known drug or alcohol abuse, diagnosed eating disorder, or other psychiatric disorder) that, in the opinion of the investigator, may preclude the participant from following and completing the protocol.
39. Have history of use of marijuana less than 3 months of V1 and unwillingness to abstain from marijuana use during the trial. Participants should also refrain from use of cannabidiol oil for the duration of the study.

40. Have had a transplanted organ (corneal transplants [keratoplasty] allowed) or awaiting an organ transplant.
41. Requires the use of supplemental oxygen.

**Prior/Concomitant Therapy**

42. Are receiving or have received within 3 months prior to screening chronic (>2 weeks or 14 days) systemic glucocorticoid therapy (excluding topical, intraocular, intranasal, intra-articular, or inhaled preparations) or have evidence of a significant, active autoimmune abnormality (for example, lupus or rheumatoid arthritis) that has required (within the last 3 months) or is likely to require, in the opinion of the investigator, concurrent treatment with systemic glucocorticoids (excluding topical, intraocular, intranasal, intra-articular, or inhaled preparations) in the next 12 months.
43. Have current or history of (less than 3 months prior to Visit 1) treatment with medications that may cause significant weight gain, including but not limited to: tricyclic antidepressants, atypical antipsychotic and mood stabilizers, for example:
  - imipramine
  - amitriptyline
  - mirtazapine
  - paroxetine
  - phenelzine
  - chlorpromazine
  - thioridazine
  - clozapine
  - olanzapine
  - valproic acid and its derivatives
  - lithium

**Note:** Selective serotonin reuptake inhibitors are permitted, except for paroxetine.

44. Have taken, less than 3 months prior to Visit 1, medications (prescribed or over-the-counter) or alternative remedies intended to promote weight loss. Examples include, but are not limited to:
  - Saxenda® (liraglutide 3.0 mg)
  - Xenical®/Alli® (orlistat)
  - Meridia® (sibutramine)
  - Acutrim® (phenylpropanolamine)
  - Sanorex® (mazindol)
  - Adipex® (phentermine)
  - BELVIQ® (lorcaserin)
  - Qsymia® (phentermine/topiramate combination)
  - Contrave® (naltrexone/bupropion)
  - Pramlintide
  - Zonisamide
  - Topiramate

- Wegovy®
- Any glucose-lowering medication, including metformin

**Note:** Use of metformin or any other glucose-lowering medication, whether prescribed for polycystic ovary syndrome or diabetes prevention is not permitted.

45. Use of stimulants less than 3 months prior to Visit 1 (for example, modafinil, armodafinil, solriamfetol, pitolisant, amphetamine, dextroamphetamine, dexmethylphenidate, methylphenidate, and lisdexamfetamine).
46. Use of hypnotics, mirtazapine, opioids, trazodone less than 3 months prior to Visit 1.
47. Use of GLP-1 RA less than 3 months prior to Visit 1.
48. Criterion 48 is deleted.
49. Use of any over-the-counter or prescription medications that could affect the evaluation of excessive sleepiness, per investigator discretion.
50. Unwillingness to discontinue over-the-counter (herbal or supplemental) medication that, in the opinion of the investigator, can interfere with the study.

#### **Prior/Concurrent Clinical Study Experience**

51. Are currently enrolled in any other clinical study involving an investigational product or any other type of medical research judged not to be scientifically or medically compatible with this study.
52. Previously randomly assigned to study intervention in this study or any other study investigating tirzepatide.
53. Have participated, within the last 30 days in a clinical trial involving a study intervention. If the previous study intervention is scientifically or medically incompatible with this study and has a long half-life, 3 months or 5 half-lives (whichever is longer) should have passed prior to screening (participation in observational studies may be permitted upon review of the observational study protocol and approval by the sponsor).
54. Are Lilly employees or employees of third-party organizations involved with the study.
55. Are investigator site personnel directly affiliated with this study and/or their immediate families. Immediate family is defined as a spouse, parent, child, or sibling, whether biological or legally adopted.
56. Have any medical condition that, in the opinion of the investigator, would be a contraindication to participation in the trial

### 5.3. Lifestyle Considerations

Per the SoA (Section 1.3), participants will consult with study personnel experienced in diet and exercise counseling to receive lifestyle program instructions at timepoints indicated in the SoA.

Diet and exercise goals and the importance of adherence to the lifestyle program will be reinforced at each trial contact by study staff.

#### 5.3.1. Meals and Dietary Restrictions

At Visit 2 and subsequent visits, participants will consult with study personnel experienced in diet and exercise counseling to receive lifestyle program instructions at timepoints indicated in the SoA. Dietary counseling will consist of advice on healthy food choices and focus on calorie restriction using a hypocaloric diet with macronutrient composition of:

- maximum 30% of energy from fat
- approximately 20% of energy from protein
- approximately 50% of energy from carbohydrates
- an energy deficit of approximately 500 kcal/day compared to the participant's estimated TEE

To encourage adherence, it is recommended that a 3-day diet and exercise diary be completed prior to each counseling visit. During each visit, the participant's diet is reviewed and advice to maximize adherence is provided if needed.

The hypocaloric diet is continued after randomization and throughout the treatment period. If a BMI  $\leq 22$  kg/m<sup>2</sup> is reached, the recommended energy intake should be recalculated with no kcal deficit for the remainder of the trial.

Total energy expenditure is calculated by multiplying the estimated Basal Metabolic Rate (BMR) (see table below) with a Physical Activity Level value of 1.3 (FAO/WHO/UNU 2004), which reflects an inactive lifestyle. This calculation provides a conservative estimate of caloric requirements:

$$\text{TEE (kcal/day)} = \text{BMR} \times 1.3$$

#### Equations for estimating BMR in kcal/day<sup>a</sup>

Sex	Age	BMR (kcal/day)
Men	18-30 years	15.057 X actual weight in kg + 692.2
	31-60 years	11.472 X actual weight in kg + 873.1
	>60 years	11.711 X actual weight in kg + 587.7
Women	18-30 years	14.818 X actual weight in kg + 486.6
	31-60 years	8.126 X actual weight in kg + 845.6
	>60 years	9.082 X actual weight in kg + 658.5

Abbreviations: BMR = basal metabolic rate; WHO = World Health Organization.

<sup>a</sup>Revised WHO equations (Adapted from: FAO/WHO/UNU 2004).

#### 5.3.2. Physical Activity

At Visit 2 and all subsequent visits, participants will be advised to increase their physical activity to at least 150 minutes per week.

#### **5.4. Screen Failures**

A screen failure occurs when a participant who consents to participate in the clinical study is not subsequently randomly assigned to study intervention in the study. A minimal set of screen failure information is required to ensure transparent reporting of screen failure participants to meet the Consolidated Standards of Reporting Trials (CONSORT) publishing requirements and to respond to queries from regulatory authorities. Minimal information includes demography, screen failure details, eligibility criteria, and any SAE.

Individuals who have history of marijuana use less than 3 months of Visit 1 may be rescreened once provided the individual is willing to abstain from marijuana use during the trial. Rescreened participants should be assigned a new participant number for every screening/rescreening event.

See Section 5.4 of the respective ISA's for additional rescreening guidelines.

#### **5.5. Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant**

Not applicable.



## 6. Study Intervention(s) and Concomitant Therapy

Study intervention is defined as any investigational intervention(s), marketed product(s), placebo, or medical device(s) intended to be administered to/used by a study participant according to the study protocol.

### 6.1. Study Interventions Administered

<b>Intervention Name</b>	Tirzepatide (LY3298176)	Placebo
<b>Type</b>	Drug	Drug
<b>Dose Formulation</b>	Single-dose pen	Single-dose pen
<b>Dosage Level(s)</b>	10 mg QW, 15 mg QW	Not applicable
<b>Route of Administration</b>	SC	SC
<b>Use</b>	Experimental	Placebo
<b>IMP and NIMP</b>	IMP	IMP
<b>Sourcing</b>	Provided centrally by the sponsor and dispensed via IWRS	
<b>Packaging and Labeling</b>	Study intervention will be provided in autoinjectors (single-dose pens), packaged in cartons to be dispensed. Clinical study materials will be labeled according to country regulatory requirements	

Abbreviations: IMP = investigational medicinal product; IWRS = interactive web-response system; NIMP = non-investigational medicinal product; QW = once-weekly; SC = subcutaneous.

There are no restrictions on the time of day each weekly dose of study intervention is given, but it is advisable to administer the SC injections on the same day and same time each week. The actual date, time, and injection details of all dose administrations will be recorded in the diary by the participant. If a dose of study intervention is missed, the participant should take it as soon as possible unless it is within 72 hours of the next dose, in which case that dose should be skipped, and the next dose should be taken at the appropriate time. The day of weekly administration can be changed if necessary, as long as the last dose was administered 72 or more hours before.

All participants will inject study intervention subcutaneously in the abdomen or thigh using the injection supplies provided; a caregiver may administer the injection in the participant's upper arm. A new autoinjector will be used for each injection. If study intervention is to always be injected in the same body region, participants should be advised to alternate injection sites each week.

### **6.1.1. Medical Devices**

The combination products provided for use in the study are tirzepatide autoinjector (or matching placebo). Any medical-device incidents, including those resulting from malfunctions of the device, must be detected, documented, and reported by the investigator throughout the study (see Section 10.3).

## **6.2. Preparation, Handling, Storage, and Accountability**

- The investigator or designee must confirm appropriate storage conditions have been maintained during transit for all study intervention received and any discrepancies are reported and resolved before use of the study intervention.
- Only participants enrolled in the study may receive study intervention. Only authorized study personnel may supply study intervention. All study intervention must be stored in a secure, environmentally controlled, and monitored (manual or automated) area in accordance with the labeled storage conditions with access limited to the investigator and authorized study personnel.
- The investigator or authorized study personnel are responsible for study intervention accountability, reconciliation, and record maintenance (that is, receipt, reconciliation, and final disposition records).
- Further guidance and information for the final disposition of unused study interventions are in the provided instructions.

## **6.3. Measures to Minimize Bias: Randomization and Blinding**

All participants will be centrally randomly assigned to study intervention using an IWRS. Before the study is initiated, the log in information and directions for the IWRS will be provided to each site.

Study intervention will be dispensed at the study visits summarized in SoA.

Returned study intervention should not be re-dispensed to the participants.

This is a double-blind study in which participants and study personnel, are blinded to study intervention. The IWRS will be programmed with blind-breaking instructions. In case of an emergency, the investigator has the sole responsibility for determining if unblinding of a participant's intervention assignment is warranted. Participant safety must always be the first consideration in making such a determination. If the investigator decides that unblinding is warranted, the investigator should make every effort to contact the sponsor prior to unblinding a participant's intervention assignment unless this could delay emergency treatment for the participant. If a participant's intervention assignment is unblinded, the sponsor must be notified immediately within 24 hours of this occurrence. The date and reason that the blind was broken must be recorded.

If an investigator, site personnel performing assessments, or participant is unblinded, the participant must be discontinued from the study. In cases where there are ethical reasons to have the participant remain in the study, the investigator must obtain specific approval from a sponsor clinical research physician (CRP) for the participant to continue in the study.

## Stratification

Participants will be stratified at randomization per ISA by country/geographic region, baseline AHI (moderate/severe), and gender.

### 6.4. Study Intervention Compliance

A record of the number of single-dose pens dispensed to and taken by each participant must be maintained and reconciled with study intervention and compliance records. Intervention start and stop dates, including dates for intervention delays and/or dose reductions will also be recorded.

Treatment compliance for each visit interval is defined as taking at least 75% of the required doses of study intervention. Similarly, a participant will be considered significantly noncompliant if he or she is judged by the investigator to have intentionally or repeatedly taken more than the prescribed amount of medication (more than 125%).

Participants considered to be poorly compliant with their medication and/or the study procedures will receive additional training and instruction, as required, and will be reminded of the importance of complying with the protocol.

### 6.5. Dose Modification

Dose modification is permitted for management of intolerable GI symptoms during the first 24 weeks of the treatment period (Section 6.5.1).

Participants who do not tolerate at least 10 mg even after the described measures, including 1 de-escalation and re-escalation attempt, will be discontinued from the study intervention but remain in the study for continued follow-up.

Interventions to optimize study intervention tolerance and adherence may be employed throughout the study and include, but are not limited to, brief temporary interruptions (Section 7.1.2) and use of additional medications to manage GI symptoms (for example, nausea, vomiting, and diarrhea).

#### 6.5.1. Management of Participants with Gastrointestinal Symptoms

In participants who experience intolerable GI symptoms (for example, nausea, vomiting, or diarrhea) at any time during the study, the following measures are recommended:

- counselling on dietary behaviors that may help mitigate nausea and vomiting, (for example, eating smaller meals, splitting 3 daily meals into 4 or more smaller ones, and stopping eating when they feel full).
- if symptoms persist despite #1, prescribing symptomatic medication (for example, antiemetic or antidiarrheal medication), at the investigator's discretion.
- if symptoms persist despite #1 and #2, interrupting study intervention for 1 dose, provided the participant has taken the last 3 weekly doses. Study treatment should be resumed at the assigned dose immediately, either alone or in combination with symptomatic medication, which can also be utilized to manage symptoms (Section 6.8).

During the first 24 weeks of the treatment period (20-week dose escalation plus 4 weeks), participants unable to tolerate 2.5 mg or 5 mg (despite the interventions mentioned above) will be discontinued from the study intervention but remain in the study.

For participants unable to tolerate any dose between 7.5 mg and 15 mg inclusive, despite the above measures, the investigator should contact Lilly to consider a dose de-escalation step with subsequent re-escalation by 2.5 mg every 4 weeks in a blinded fashion, to reach either the 10-mg or 15-mg dose as described below.

Only 1 cycle of dose de-escalation and re-escalation is permitted during the first 24 weeks of the treatment period. MTD is either 10 mg or 15 mg. Dose modifications after the first 24 weeks of the treatment period are not permitted. For temporary study treatment discontinuation, see Section 7.1.2.

Participants who tolerate

- 10 mg, but do not tolerate 12.5 mg or 15 mg even following the above measures, including 1 de-escalation and re-escalation attempt, will continue on 10 mg as their MTD dose.
- 12.5 mg, but do not tolerate 15 mg even after the above measures, including 1 de-escalation and re-escalation attempt, will continue on 10 mg as their MTD dose.
- 15 mg will continue on 15 mg as their MTD dose.

## 6.6. Continued Access to Study Intervention after the End of the Study

Tirzepatide will not be made available to participants after conclusion of the study.

## 6.7. Treatment of Overdose

Study intervention overdose (more than the specified number of injections in less than 72 hours) will be reported as an AE.

In the event of an overdose, the treating physician should:

- Contact the medical monitor immediately.
- Evaluate the participant to determine, in consultation with the medical monitor, whether study intervention should be interrupted or whether the dose should be reduced.
- Closely monitor the participant for any AE/SAE and laboratory abnormalities until study intervention no longer has a clinical effect.

## 6.8. Concomitant Therapy

Participants will be permitted to use concomitant medications that they require during the study, except certain medications (Section 6.8.2) that may interfere with the assessment of efficacy and safety characteristics of the study treatments.

Investigative-site staff will inform participants that they must consult with the investigator or a designated site staff member upon being prescribed any new medications during the study. This may not be possible when initiated for treatment of medical emergencies, in which case, the participant will inform the investigator or a designated site staff member as soon as possible.

Nonstudy medications taken by participants who are screened but not randomly assigned to study intervention will not be reported to Lilly unless an SAE or AE occurs that the investigator believes may have been caused by a study procedure.

Any medication or vaccine (including over-the-counter or prescription medicines, vitamins, and/or herbal supplements, or other specific categories of interest) that the participant is receiving at the time of screening or receives during the study must be recorded along with:

- Reason for use
- Dates of administration including start and end dates

The medical monitor should be contacted if there are any questions regarding concomitant or prior therapy.

### **6.8.1. Management of Incident Diabetes**

Incident diabetes is defined when any 1 of the following occur after randomization (American Diabetes Association 2020):

- unequivocal hyperglycemia (random glucose  $\geq 200$  mg/dL) with signs or symptoms of hyperglycemia
- within a 4-week period, any 2 of the following criteria are observed or 1 abnormal value is observed and confirmed:
  - HbA1c  $\geq 6.5\%$  (48 mmol/mol)
  - FG or 0-hour serum glucose from 2-hour OGTT  $\geq 126$  mg/dL (7.0 mmol/L)
  - 2-hour glucose  $\geq 200$  mg/dL (11.1 mmol/L) by a 2-hour OGTT
  - initiation of any medication for the treatment of diabetes

Participants who develop type 2 diabetes during the study will be

- provided with and trained to use a glucometer
- educated on the signs and symptoms of hypoglycemia and its treatment (see Section 10.3.7.1), and
- provided a diary to record hypoglycemic episodes per Section 10.3.7.1.

Participants will be referred to their usual care provider and provided with a letter showing the study results indicative of diabetes. The decision to further evaluate, to initiate antihyperglycemic therapy, and the choice of antihyperglycemic medication will be at the discretion of the participant's usual care provider, with the exception of use of DPP-4 inhibitors and open-label GLP-1R agonists, which are prohibited in the study (Section 6.8.2). Monitoring for hypoglycemia includes capture of events as defined in Section 10.3 (Appendix 3). Date of diabetes diagnosis will be captured in the AE CRF.

Initiation of metformin for the treatment of diabetes is permitted, but metformin should not be initiated during the study for the treatment of other metabolic conditions (for example, polycystic ovary syndrome and diabetes prevention).

### 6.8.2. Prohibited Concomitant Medications

The following medications are prohibited during the study:

- DPP-4 inhibitors
- Open-label GLP-1R agonists
- Stimulants (for example, modafinil, armodafinil, solriamfetol, pitolisant, amphetamine, dextroamphetamine, dexamethylphenidate, methylphenidate, and lisdexamfetamine)
- medications that may cause significant weight gain (such as, but not limited to, paroxetine, tricyclic antidepressants, atypical antipsychotic and mood stabilizers).
- Medications that may cause weight loss (such as, but not limited to, liraglutide, semaglutide, orlistat, sibutramine, phenylpropanolamine, mazindol, phentermine, lorcaserin, naltrexone/bupropion, phentermine/topiramate combination, pramlintide, zonisamide, and topiramate)
- OTC, herbal, or supplemental medications that, in the opinion of the investigator, may interfere with the study
- hypnotics, mirtazapine, opioids, trazodone, pramlintide, sibutramine, orlistat, and zonisamide
- Systemic glucocorticoid therapy, per discussion with sponsor
- Use of any over-the-counter or prescription medications that could affect the evaluation of excessive sleepiness, per investigator discretion (such as, but not limited to, CBD oil, THC, etc.)
- Any glucose-lowering medication, including metformin

In addition,

- Active device treatment of OSA other than PAP therapy (for example, dental appliance), or other treatment that, in the opinion of the investigator, may interfere with study outcomes.

Participants who initiate and will not or cannot discontinue a prohibited medication, device, or other treatment will be permanently discontinued from study intervention per Section 7.1.

## 7. Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal

Discontinuation of specific sites or of the study as a whole are handled as part of Section 10.1, Appendix 1.

### 7.1. Discontinuation of Study Intervention

When necessary, a participant may be permanently discontinued from study intervention. If so, the participant will remain in the study and follow procedures for remaining study visits, as shown in the SoA.

A participant who prematurely discontinues study intervention is strongly encouraged to remain in the study for safety and efficacy assessments through the treatment period and post-treatment follow-up.

If a participant who discontinues the double-blind study treatment prematurely declines to complete the remaining scheduled study visits, then the participant should complete the early ED procedures indicated in the SoA. Participants should be encouraged to come back for the last treatment visit (Visit 11) to complete the end-of-treatment phase procedures and return for the post-treatment follow-up Visit 801.

Possible reasons leading to permanent discontinuation of study intervention:

- **participant decision**
  - the participant requests to discontinue study intervention
- **clinical considerations**
  - initiation of a prohibited medication (Section 6.8.2), if participants will not or cannot discontinue them
  - BMI  $\leq 18.5$  kg/m<sup>2</sup> is reached at any time during the treatment period, study intervention discontinuation should be considered
  - TEAE
    - intolerable GI symptoms despite management (Section 6.5.1)  
*Note:* The investigator should contact the Sponsor CRP to discuss whether it is medically appropriate for the participant to continue study treatment.
    - significant elevation of calcitonin

- If the investigator determines that a systemic hypersensitivity reaction has occurred related to study intervention administration, the participant may be permanently discontinued from the study intervention, and the sponsor's designated medical monitor should be notified. If the investigator is uncertain about whether a systemic hypersensitivity reaction has occurred and whether discontinuation of study intervention is warranted, the investigator may consult the sponsor
- occurrence of any other TEAE, SAE, or clinically significant finding for which the investigator believes that permanent study intervention discontinuation is the appropriate measure to be taken
- Diagnosis of
  - T1DM
  - Thyroid C-cell hyperplasia, MTC or MEN Syndrome type 2 after randomization
  - acute or chronic pancreatitis
  - an active or untreated malignancy (other than basal or squamous cell skin cancer, in situ carcinomas of the cervix, or in situ prostate cancer) after randomization
- onset of pregnancy in a female participant (see Sections 7.2 and 8.3.2)
- Suicidal Ideation and Behavior
  - PHQ-9 score  $\geq 15$ 
    - Participants should be referred to a Mental Health Professional (MHP) to assist in deciding whether the participant should be discontinued from study intervention. If a participant's psychiatric disorder can be adequately treated with psycho- and/or pharmacotherapy, then the participant, at the discretion of the Investigator (in agreement with the MHP), may be continued in the trial on randomly assigned therapy.
  - in addition, study intervention may be discontinued if participants:
    - answered "yes" to Question 4 or Question 5 on the "Suicidal Ideation" portion of the C-SSRS, **or**
    - answered "yes" to any of the suicide-related behaviors on the Suicidal Behavior portion of the C-SSRS.

A psychiatrist or appropriately trained professional may assist in the decision to discontinue the participant. The participant should be referred to a MHP for further evaluation and care.



### 7.1.1. Liver Chemistry Stopping Criteria

The study intervention should be interrupted or discontinued if one or more of these conditions occur:

Elevation	Exception
ALT or AST >8x ULN	
ALT or AST >5x ULN for more than 2 weeks	
ALT or AST >3x ULN and either TBL >2x ULN or INR >1.5	In participants with Gilbert's syndrome, doubling of direct bilirubin should be used for drug interruption/ discontinuation decisions rather than TBL>2x ULN.
ALT or AST >3x ULN with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%)	
ALP >3x ULN, when the source of increased ALP is the liver	
ALP >2.5x ULN and TBL > 2x ULN	In participants with Gilbert's syndrome, doubling of direct bilirubin should be used for drug interruption/ discontinuation decisions rather than TBL>2x ULN.
ALP >2.5x ULN with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%)	

Source: FDA 2009 and other consensus guidelines, with minor modifications

Abbreviations: ALT = alanine aminotransferase; AST = aspartate aminotransferase; ALP = alkaline phosphatase; INR = international normalized ratio; TBL = total bilirubin level; ULN = upper limit of normal.

Resumption of the study intervention can be considered only in consultation with the Lilly-designated medical monitor and only if the liver test results return to baseline and if a self-limited non-drug etiology is identified.

Participants who are discontinued from study intervention due to a hepatic event or liver test abnormality should have additional hepatic safety data collected as described in Section 10.6, Appendix 6.

### 7.1.2. Temporary Discontinuation

In certain situations, after randomization, the investigator may need to temporarily interrupt study intervention. Every effort should be made by the investigator to maintain participants on study intervention and to restart study intervention after any temporary interruption, as soon as it is safe to do so. Distribution of study intervention at the correct dose will be per IWRS instructions.

Investigators should inform the sponsor that study intervention has been temporarily interrupted. The data related to temporary interruption of study treatment will be documented in source documents and entered on the CRF.

If study intervention interruption is...	then...
2 consecutive doses or less	participant restarts study intervention at last administered dose, as per escalation schedule.
3 consecutive doses or more	participant restarts study intervention (at 5 mg, managed by IWRS) and repeats dose escalation scheme.
Due to an AE	the event is to be documented and followed according to the procedures in Section 8.3 of this protocol.
Due to intolerable persistent GI AE	participants should be treated as suggested in Section 6.5.1.

Abbreviations: AE = adverse event; GI = gastrointestinal; IWRS = interactive web-response system.

## 7.2. Participant Discontinuation/Withdrawal from the Study

Discontinuation is expected to be uncommon.

A participant may withdraw from the study:

- at any time at the participant's own request
- at the request of the participant's designee (for example, parents or legal guardian)
- a participant will be withdrawn from the study in case of inadvertent enrollment
- at the discretion of the investigator for safety, behavioral, compliance, or administrative reasons
  - If the participant becomes pregnant during the study (see Section 8.3.2 for additional details)
- if enrolled in any other clinical study involving an investigational product, or enrolled in any other type of medical research judged not to be scientifically or medically compatible with this study
- if the participant, for any reason, requires treatment with a therapeutic agent that is prohibited by the protocol and has been demonstrated to be effective for treatment of the study indication. In this case, discontinuation from the study occurs prior to introduction of the new agent.

At the time of discontinuing from the study, if possible, the participant will complete procedures for an ED visit and post-treatment follow-up, if applicable, as shown in the SoA. If the participant has not already discontinued the study intervention, the participant will be permanently discontinued from the study intervention at the time of the decision to discontinue the study.

If the participant withdraws consent for disclosure of future information, the sponsor may retain and continue to use any data collected before such a withdrawal of consent. If a participant withdraws from the study, the participant may request destruction of any samples taken and not tested, and the investigator must document this in the site study records.

### **7.3. Lost to Follow up**

A participant will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site. Site personnel or designee are expected to make diligent attempts to contact participants who fail to return for a scheduled visit or were otherwise unable to be followed up by the site.

## 8. Study Assessments and Procedures

Study procedures and their timing are summarized in the SoA.

Immediate safety concerns should be discussed with the sponsor immediately upon occurrence or awareness to determine if the participant should continue or discontinue study intervention.

Adherence to the study design requirements, including those specified in the SoA, is essential and required for study conduct.

All screening evaluations must be completed and reviewed to confirm that potential participants meet all eligibility criteria. The investigator will maintain a screening log to record details of all participants screened and to confirm eligibility or record reasons for screening failure, as applicable.

### 8.1. Efficacy Assessments

#### 8.1.1. Primary Efficacy Assessment

##### 8.1.1.1. Polysomnography

The primary efficacy assessment in this study is AHI. AHI measurements will be collected via polysomnography.

Polysomnography assessments (including AHI, blood oxygen saturation parameters, PR, sleep parameters) will be performed during 1-night, overnight clinic stays, per the SoA. Data from the PSGs will be read and scored centrally using the AASM 1B hypopnea scoring method (when there is  $\geq 4\%$  oxygen desaturation from pre-event baseline; see Section 10.10, Appendix 10 for definitions) (Hamilton et al. 2021).

The eligibility criteria and AHI related endpoints of the study will be assessed based on central reading of the PSG.

#### 8.1.2. Secondary Efficacy Assessments

##### 8.1.2.1. Patient-Reported Outcomes

###### 8.1.2.1.1. Functional Outcomes of Sleep Questionnaire (FOSQ)

The FOSQ will be included to assess change in FOSQ domains and total score from baseline to Week 52. The FOSQ is a 30-item sleep-specific, participant-completed questionnaire used to assess the effect of disorders associated with excessive daytime sleepiness (EDS) on daily functioning in adults. It assesses the following 5 domains of

- General productivity (8 items)
- Activity level (9 items)
- Vigilance (7 items)
- Social outcomes (2 items)
- Intimate and sexual relationships (4 items)

The FOSQ items assess participant's current status with each item rated on a scale of 1 (extreme difficulty) to 4 (no difficulty), with an additional not applicable (0 = "I don't do this activity for

other reasons”) also available. Individual domain scores are calculated by taking the mean of answered, non-zero items within each domain and a total score can be calculated by first computing the mean score for each domain, then multiplying the mean of the domain scores by 5 (Weaver et al. 1997). The total score for the FOSQ 10-item short form (FOSQ-10) can also be calculated.

#### **8.1.2.1.2. *PROMIS Short Form v1.0 Sleep-related Impairment 8a***

The PROMIS Short Form v1.0 Sleep-related Impairment 8a assesses self-reported perceptions of alertness, sleepiness, and tiredness during usual waking hours, and the perceived functional impairments associated with sleep problems or impaired alertness. The PROMIS Short Form v1.0 Sleep-related Impairment 8a consists of 8 items each rated on a 5-point scale ranging from “not at all” to “very much.” Items have a recall period of “in the past 7 days.” Individual item scores are totaled to obtain a raw score, with higher scores indicating more sleep-related impairment. Raw scores can be converted to a T-score, which is standardized with a mean of 50 and a SD of 10. (Northwestern, 2016a)

#### **8.1.2.1.3. *PROMIS Short Form v1.0 Sleep Disturbance 8b***

The PROMIS Short Form v1.0 Sleep Disturbance 8b assesses self-reported perceptions of sleep quality, sleep depth, and restoration associated with sleep, including perceived difficulties and concerns with getting to sleep or staying asleep, as well as perceptions of the adequacy of and satisfaction with sleep. The PROMIS Short Form v1.0 Sleep Disturbance 8b consists of 8 items each rated on a 5-point scale ranging from “not at all” to “very much,” “never” to “always,” or “very poor” to “very good.” Items have a recall period of “in the past 7 days.” Individual item scores are totaled to obtain a raw score, with higher scores indicating more sleep disturbance. Raw scores can be converted to a T-score, which is standardized with a mean of 50 and a SD of 10. (Northwestern, 2016b)

#### **8.1.2.1.4. *Epworth Sleepiness Scale***

The ESS will be included to assess improvements in excessive daytime sleepiness from baseline to Week 52. The ESS is an 8-item participant-completed measure that asks the participant to rate on a scale of 0 (would never doze) to 3 (high chance of dozing), their usual chances of dozing in 8 different daytime situations, with a recall period of “in recent times.” The ESS total score is the sum of the 8-item scores and ranges from 0 to 24, with higher scores indicating greater daytime sleepiness (Johns 1991).

#### **8.1.2.1.5. *Short-Form 36 Version 2 Health Survey, Acute Form, 1-week Recall Version***

The SF-36v2 will be included to assess health-related quality of life from baseline to Week 52. The SF-36v2 acute form, 1-week recall version is a 36-item generic, participant-completed measure designed to assess the following 8 domains over “the past week.”

- Physical functioning
- Role-physical
- Bodily pain
- General health
- Vitality
- Social functioning

- Role-emotional, and
- Mental health

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

#### **8.1.2.1.6. *Patient Global Impression of Status – Obstructive Sleep Apnea (PGIS-OSA) Symptoms Scales***

Three patient global impression of status scales will be included to assess categorical shift in participant self-rated assessment of their OSA symptom severity from baseline to Week 52.

##### **PGIS-OSA Fatigue**

This is a single-item, participant self-rated assessment of their overall level of fatigue due to OSA, “over the past 7 days.” The item is rated on a 4-point scale ranging from “No fatigue” to “Severe fatigue.”

##### **PGIS-OSA Sleepiness**

This is a single-item, participant self-rated assessment of their overall level of sleepiness due to OSA during waking hours, “over the past 7 days.” The item is rated on a 4-point scale ranging from “Not at all sleepy” to “Very sleepy.”

##### **PGIS-OSA Snoring**

The PGIS-OSA Snoring scale consists of two items. The first item is a participant self-rated assessment of their overall perception of the severity of their snoring due to OSA, “over the past 7 days,” with respect to how much their snoring has affected their sleep. The item is rated on a 4-point scale ranging from “Not at all affected” to “Very affected.” For the second item, participants will be asked on a 3-point scale (“Not at all” to “All the time”) if they have ever been told by someone else that they snore in their sleep.

### **8.1.3. Exploratory Efficacy Assessments**

#### **8.1.3.1. Actigraphy**

The actigraphy device (AX6) will be utilized in the OSA trial to objectively evaluate the effect of tirzepatide on various sleep and physical activity parameters including but not limited to change of daytime sleep time, sleep efficiency, and change in daytime physical activity from baseline. The actigraphy device is a data logger capable of recording raw data from a suite of integrated sensors and can be configured to collect movement-relevant data in an uninterrupted fashion for up to 2 months, thus is ideal for collecting longitudinal movement data (for example, physical activity, sleep, etc.) in real-world health research and well-defined clinical trials. The actigraphy device meets the CE (Conformite Europeenne) mark requirements. Participants should wear the actigraphy device for 7 consecutive days, 5 times during the study, per the SoA.

Lack of participation in actigraphy collections at any time is not considered a protocol deviation.

### **8.1.3.2. Patient Global Impression of Change – Obstructive Sleep Apnea (PGIC-OSA) Symptoms Scales**

Four patient global impression of change scales will be included to assess categorical shift in participant self-rated assessment of change in their OSA symptom severity from baseline to Week 52.

#### **PGIC-OSA Fatigue**

This is a single-item, participant self-rated assessment of the change in their overall level of fatigue due to OSA, “since you started taking the study medication.” The item is rated on a 5-point scale ranging from “Much worse” to “Much better.”

#### **PGIC-OSA Sleepiness**

This is a single-item, participant self-rated assessment of the change in their overall level of sleepiness due to OSA during waking hours, “since you started taking the study medication.” The item is rated on a 5-point scale ranging from “Much more sleepy” to “Much less sleepy.”

#### **PGIC-OSA Sleep Quality**

This is a single-item, participant self-rated assessment of the change in their overall sleep quality due to OSA, “since you started taking the study medication.” The item is rated on a 5-point scale ranging from “Much worse” to “Much better.”

#### **PGIC-OSA Snoring**

This is a single-item, participant self-rated assessment of the overall change in how their snoring has affected their sleep, “since you started taking the study medication.” The item is rated on a 5-point scale ranging from “My sleep is much more affected” to “My sleep is much less affected.”

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

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

- mobility
- selfcare
- usual activities
- pain/discomfort, and
- anxiety/depression.

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

- no problems
- slight problems
- moderate problems
- severe problems, and
- unable to perform/extreme problems.

A total of 3125 health states is possible. In addition to the health profile, a single health state index value can be derived based on a formula that attaches weights to each of the levels in each dimension. This index value ranges between less than 0 (where 0 is a health state equivalent to death; negative values are valued as worse than dead) to 1 (perfect health). In addition, the EQ Visual Analog Scale records the respondent's self-rated health status on a vertical graduated (0 to 100) visual analog scale. The participant rates his/her perceived health from 0 (the worst imaginable health) to 100 (the best imaginable health). In conjunction with the health state data, it provides a composite picture of the respondent's health status.

## **8.2. Safety Assessments**

Planned time points for all safety assessments are provided in the SoA.

### **8.2.1. Physical Examinations**

A complete physical examination will include, at a minimum, assessments of the CV, respiratory, gastrointestinal and neurological systems as well as a thyroid examination. Height, weight, and waist circumference will also be measured and recorded.

A symptom-directed physical assessment will be performed as indicated in the SoA, as clinically indicated.

Investigators should pay special attention to clinical signs related to previous serious illnesses.

### **8.2.2. Vital Signs**

For each participant, vital signs measurements should be conducted according to the SoA (Section 1.3).

### **8.2.3. Electrocardiograms (ECG)**

Single 12-lead ECG will be obtained as outlined in the SoA (see Section 1.3) using an ECG machine that automatically calculates the heart rate and measures PR, QRS, QT, and QTc intervals.

A local single 12-lead ECG will be collected at designated visits. ECGs should be collected prior to collection of blood samples for laboratory testing, including PK samples. Participants should be supine for at least 5 minutes before ECG collections and remain supine but awake during the ECG collection. ECGs may be repeated at the investigator's discretion at any visit.

After enrollment, if a clinically significant increase in the QT/QTc interval from baseline or other clinically significant quantitative or qualitative change from baseline is identified, the participant will be assessed by the investigator for symptoms (for example, palpitations, near syncope, and syncope) and to determine whether the participant can continue in the study. The investigator or qualified designee is responsible for determining if any change in participant management is needed and must document his/her review of the ECG printed at the time of evaluation from at least 1 of the replicate ECGs from each time point.

### **8.2.4. Clinical Safety Laboratory Tests**

See Section 10.2, Appendix 2 for the list of clinical laboratory tests to be performed and the SoA for the timing and frequency.



The investigator must review the laboratory results, document this review, and report any clinically relevant changes occurring during the study as an AE. The laboratory results must be retained with source documents unless a Source Document Agreement or comparable document cites an electronic location that accommodates the expected retention duration. Clinically significant abnormal laboratory findings are those which are not associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant's condition.

All laboratory tests with values considered clinically significantly abnormal during participation in the study should be repeated until the values return to normal or baseline or are no longer considered clinically significant by the investigator or medical monitor.

- If such values do not return to normal/baseline within a period of time judged reasonable by the investigator, the etiology should be identified and the sponsor notified.
- All protocol-required laboratory assessments, as defined in Section 10.2, Appendix 2 must be conducted in accordance with the SoA, standard collection requirements, and laboratory manual.

If laboratory values from non-protocol specified laboratory assessments performed at an investigator-designated local laboratory require a change in participant management or are considered clinically significant by the investigator (for example, SAE or AE or dose modification), then report the information as an AE.

#### **8.2.5. Pregnancy Testing**

See the SoA (Section 1.3) for testing for pregnancy and timepoints.

See Appendix 4 (Section 10.4) for additional information for pregnancy.

#### **8.2.6. Suicidal Ideation and Behavior Risk Monitoring**

Patients with obesity may occasionally develop suicidal ideation or behavior.

Participants should be monitored appropriately and observed closely for SIB or any other unusual changes in behavior, especially at the beginning and end of the course of intervention, or at the time of dose changes, either increases or decreases. Participants who experience signs of SIB should undergo a risk assessment. All factors contributing to SIB should be evaluated and consideration should be given to discontinuation of the study intervention.

Baseline assessment of suicidal ideation and behavior/intervention emergent suicidal ideation and behavior will be monitored using the C-SSRS and PHQ-9.

##### **8.2.6.1. C-SSRS**

Columbia Suicide-Severity Rating Scale (C-SSRS) is a scale that captures the occurrence, severity, and frequency of suicidal ideation and behavior during the assessment period via a questionnaire. The scale was developed by the National Institute of Mental Health (NIMH) trial group (TASA) for the purpose of being counterpart to the Columbia Classification Algorithm of Suicide Assessment (C-CASA) categorization of suicidal events. (The Columbia Lighthouse Project, 2013)

For this study, the C-SSRS is adapted for the assessment of the ideation and behavior categories only. The Intensity of Ideation and Lethality of Behavior sections are removed.

### 8.2.6.2. PHQ-9

The PHQ-9 is a validated self-report screening tool that assesses the presence and intensity of depressive symptoms. The PHQ-9, which incorporates the 9 Diagnostic and Statistical Manual-IV depression criteria as “0” (not at all) to “3” (nearly every day), was developed for use in primary care settings (Kroenke et al. 2001).

## 8.3. Adverse Events, Serious Adverse Events, and Product Complaints

The definitions of the following events can be found in Section 10.3, Appendix 3:

- AEs
- SAEs
- PCs

These events will be reported by the participant (or, when appropriate, by a caregiver, surrogate, or the participant’s legally authorized representative).

The investigator and any qualified designees are responsible for detecting, documenting, and recording events that meet these definitions and remain responsible for following up events that are serious, considered related to the study intervention or study procedures, or that caused the participant to discontinue the study intervention before completing the study (see Section 7).

Care will be taken not to introduce bias when detecting events. Open-ended and non-leading verbal questioning of the participant is the preferred method to inquire about event occurrences.

After the initial report, the investigator is required to proactively follow each participant at subsequent visits/contacts. All SAEs and AEs of special interest (as defined in Section 8.3.3) will be followed until resolution, stabilization, the event is otherwise explained, or the participant is lost to follow-up (as defined in Section 7.3). For PCs, the investigator is responsible for ensuring that follow-up includes any supplemental investigations as indicated to elucidate the nature and/or causality. Further information on follow-up procedures is provided in Section 10.3, Appendix 3.

### 8.3.1. Timing and Mechanism for Collecting Events

This table describes the timing, deadlines, and mechanism for collecting events.

Event	Collection Start	Collection Stop	Timing for Reporting to Sponsor or Designee	Mechanism for Reporting	Back-Up Method of Reporting
<b>AE</b>					
AE	Signing of the ICF	participation in study has ended	As soon as possible upon site awareness	AE CRF	N/A

Event	Collection Start	Collection Stop	Timing for Reporting to Sponsor or Designee	Mechanism for Reporting	Back-Up Method of Reporting
<b>SAE</b>					
SAE and SAE updates – prior to start of study intervention <b>and</b> deemed reasonably possibly related to study procedures	Signing of the ICF	start of intervention	Within 24 hours of awareness	SAE CRF	SAE paper form
SAE and SAE updates – after start of study intervention	Start of intervention	participation in study has ended	Within 24 hours of awareness	SAE CRF	SAE paper form
SAE <sup>a</sup> – after participant’s study participation has ended <b>and</b> the investigator becomes aware	After participant’s study participation has ended	N/A	Promptly	SAE paper form	N/A
<b>Pregnancy</b>					
Pregnancy in female participants and female partners of male participants	After the start of study intervention	until 1 month after the last dose of study intervention	Within 24 hours (see Section 8.3.2)	Pregnancy paper form	Pregnancy paper form
<b>PCs</b>					
PC associated with an SAE or might have led to an SAE	Start of study intervention	End of study intervention	Within 24 hours of awareness	PC form	N/A

Event	Collection Start	Collection Stop	Timing for Reporting to Sponsor or Designee	Mechanism for Reporting	Back-Up Method of Reporting
PC not associated with an SAE	Start of study intervention	End of study intervention	Within 1 business day of awareness	PC form	N/A
Updated PC information	—	—	As soon as possible upon site awareness	Originally completed PC form with all changes signed and dated by the investigator	N/A
PC (if investigator becomes aware)	Participation in study has ended	N/A	Promptly	PC form	

Abbreviations: AE = adverse event; CRF = case report form; ICF = informed consent form; N/A = not applicable; PC = product complaint; SAE = serious adverse event.

<sup>a</sup> Serious adverse events should not be reported unless the investigator deems them to be possibly related to study treatment or study participation.

### 8.3.1.1. Adverse Event Monitoring with a Systematic Questionnaire

Nonleading AE collection should occur prior to the collection of the C-SSRS and PHQ-9, if AE and C-SSRS/PHQ-9 collections done on the same day.

If a suicide-related event is discovered *during the C-SSRS or PHQ-9* but was not captured during the nonleading AE collection, sites should not change the AE form.

If an AE is serious or leads to discontinuation, it needs to be included on the AE form and the process for reporting SAEs is followed.

### 8.3.2. Pregnancy

#### Collection of pregnancy information

*Male participants with partners who become pregnant*

- The investigator will attempt to collect pregnancy information on any male participant's female partner who becomes pregnant while the male participant is in this study. This applies only to male participants who receive study intervention.
- After learning of a pregnancy in the female partner of a study participant, the investigator will
  - obtain a consent to release information from the pregnant female partner directly, and
  - within 24 hours after obtaining this consent will record pregnancy information on the appropriate form and submit it to the sponsor.

- The female partner will also be followed to determine the outcome of the pregnancy. Information on the status of the mother and child will be forwarded to the sponsor. Generally, the follow-up will be no longer than 6 to 8 weeks following the estimated delivery date. Any termination of the pregnancy will be reported regardless of gestational age, fetal status (presence or absence of anomalies) or indication for the procedure.

#### *Female participants who become pregnant*

- The investigator will collect pregnancy information on any female participant who becomes pregnant while participating in this study. The initial information will be recorded on the appropriate form and submitted to the sponsor within 24 hours of learning of a participant's pregnancy.
- The participant will be followed to determine the outcome of the pregnancy. The investigator will collect follow-up information on the participant and the neonate and the information will be forwarded to the sponsor. Generally, follow-up will not be required for longer than 6 to 8 weeks beyond the estimated delivery date. Any termination of pregnancy will be reported, regardless of gestational age, fetal status (presence or absence of anomalies) or indication for the procedure.
- While pregnancy itself is not considered to be an AE or SAE, any pregnancy complication or elective termination of a pregnancy for medical reasons will be reported as an AE or SAE.
- A spontaneous abortion (occurring at <20 weeks gestational age) or still birth (occurring at  $\geq$ 20 weeks gestational age) is always considered to be an SAE and will be reported as such.
- Any post-study pregnancy related SAE considered reasonably related to the study intervention by the investigator will be reported to the sponsor as described in protocol Section 8.3.1. While the investigator is not obligated to actively seek this information in former study participants, he or she may learn of an SAE through spontaneous reporting.
- Any female participant who becomes pregnant while participating in the study will discontinue study intervention and be withdrawn from the study. If the participant is discontinued from the study, follow the standard discontinuation process and continue directly to the study follow-up phase. The follow-up on the pregnancy outcome should continue independent of intervention or study discontinuation.

### **8.3.3. Adverse Events of Special Interest**

Adverse events of special interest are prospectively defined to include

- severe hypoglycemia
- MACE (adjudicated); includes, but not limited to CV death, nonfatal MI, nonfatal stroke, hospitalization for unstable angina, and hospitalization for heart failure
- treatment-emergent arrhythmias and cardiac conduction disorders
- hepatobiliary disorders; includes biliary colic, cholecystitis, and other gallbladder disease
- severe GI events
- acute renal events

- Major Depressive Disorder/suicidal behavior and ideation
- pancreatitis (adjudicated)
- c-cell hyperplasia and thyroid malignancies
- allergic/hypersensitivity reactions; includes ISRs and ADA formation

If these events are reported, investigators may be prompted to collect additional data about the event. Sections 10.1.5.1 and 10.1.5.2 (Appendix 1), outline additional information on CV and pancreatic adjudication. Section 10.3.7 (Appendix 3) outline additional information on hypoglycemia, hypersensitivity reactions, and ISRs.

## 8.4. Pharmacokinetics

Blood samples will be obtained from all participants enrolled in the 2 ISAs to enable the characterization of tirzepatide PK and exposure response relationships, as permissible. Samples will be collected with concurrent immunogenicity samples at timepoints indicated in the SoA (Section 1.3).

Blood samples collected from participants in the placebo arms will not be included in the bioanalyses of drug concentrations.

- Plasma samples will be collected for measurement of plasma concentrations of tirzepatide as specified in the SoA
- Instructions for the collection and handling of biological samples will be provided by the sponsor. The actual date and time (24-hour clock time) of each sample will be recorded.

Drug concentration information that may unblind the study will not be reported to investigative sites or blinded personnel.

### 8.4.1. Bioanalytical Methods

Samples will be analyzed at a laboratory approved by the sponsor, and stored at a facility designated by the sponsor. Concentrations of tirzepatide will be assayed using a validated liquid chromatography mass spectrometry method. Analyses of samples collected from placebo-treated participants are not planned. Bioanalytical samples collected to measure tirzepatide concentrations will be retained for a maximum of 1 year following last participant visit for the study. During this time, samples remaining after the bioanalyses may be used for exploratory analyses such as metabolism work, protein binding, and/or bioanalytical method cross-validation.

## 8.5. Pharmacodynamics

Samples to assess the PD properties of tirzepatide are included in the efficacy measures and not applicable in this section.

## 8.6. Genetics

A whole blood sample will be collected for pharmacogenetic analysis where local regulations allow.

## 8.7. Biomarkers

Blood samples will be collected to address questions of relevance to drug disposition, target engagement, PD, mechanism of action, variability of study participant response (including safety), and clinical outcome. Biomarkers will include measurement of biomolecules including proteins, lipids, and other cellular elements.

Samples will be collected according to the schedule described in the SoA and as detailed in the laboratory manual provided separately to sites.

All samples will be coded with the study participant number. These samples and any data generated can be linked back to the study participant only by the investigator site personnel.

Lilly may store samples after the end of the study to achieve study objectives. Additionally, with participant's consent, samples may be used for further research by Lilly or others such as universities or other companies to contribute to the understanding of OSA, obesity, or other diseases, the development of related or new treatments, or research methods.

See Section 10.1.12, Appendix 1 for details related to sample retention.

## 8.8. Immunogenicity Assessments

At the visits and times specified in the SoA (Section 1.3), venous blood samples will be collected for analysis to determine antibody production against tirzepatide. Antibodies may be further characterized for: cross-reactive binding to native GIP and GLP-1, neutralizing activity of tirzepatide on the GIP and GLP-1 receptors, and neutralizing activity to native GIP and/or GLP-1. To interpret the results of immunogenicity, a venous blood sample will be collected at the same time points to determine the serum concentrations of tirzepatide. All samples for immunogenicity should be taken predose when applicable and possible.

Treatment-emergent ADAs are defined in Section 10.3, Appendix 3.

Immunogenicity will be assessed by a validated assay designed to detect ADAs in the presence of tirzepatide at a laboratory approved by the sponsor. The purpose of retention, the maximum duration of retention, and facility for long-term storage of samples is described in Section 10.2, Appendix 2. Samples may also be used for development and control of an immunogenicity assay.

For details related to sample retention, see Section 10.1.12, Appendix 1.

## 8.9. Health Economics OR Medical Resource Utilization and Health Economics

Health economics or medical resource utilization and health economics parameters are not evaluated in this study.

## **9. Statistical Considerations**

This section is a summary of the planned statistical analyses of the most important endpoints, including primary and key secondary endpoints. The SAP will include more technical and detailed description of the statistical analyses described in this section.

Unblinding details will be specified in the unblinding plan section of the SAP or in a separate unblinding plan document.

### **9.1. Statistical Hypotheses**

For each ISA, the primary objective is to demonstrate that tirzepatide at the MTD (10 mg or 15 mg) is superior to Placebo in treating participants with OSA with respect to AHI endpoint. Thus the null and alternative hypotheses will be defined as below.

Null hypothesis: tirzepatide at the MTD (10 mg or 15 mg) is not different from the placebo with respect to the mean percent change from baseline in AHI at 52 weeks.

Alternative hypothesis: tirzepatide at the MTD (10 mg or 15 mg) is superior to the placebo with respect to the mean percent change from baseline in AHI at 52 weeks.

The treatment effect will be defined as the difference between the estimates of the mean change from baseline at 52 weeks for tirzepatide at the MTD (10 mg or 15 mg) and placebo.

#### **9.1.1. Multiplicity Adjustment**

The statistical comparisons for the primary efficacy endpoint and the key secondary endpoints will be carried out based on a graphical approach for multiple comparisons within each ISA (Bretz et al. 2011). The graphical approach is a closed testing procedure; hence, it strongly controls the familywise Type I error rate (2-sided alpha level of 0.05) across the primary and key secondary objectives (Alosh et al. 2014). Details about the graphical approach will be described in the SAP.

### **9.2. Analyses Sets**

This table describes the populations that will be used for statistical analyses within each ISA of the master protocol. Additional intervention-specific populations for analyses may be described in the respective ISA.



<b>Analysis Set or Population</b>	<b>Description</b>
<b>mITT population</b>	All randomly assigned participants who are exposed to at least 1 dose of study intervention.
<b>FAS</b>	Data obtained during treatment period from the mITT population excluding those discontinuing study due to inadvertent enrollment, regardless of adherence to study intervention
<b>EAS</b>	Data obtained during treatment period from the mITT population excluding those discontinuing study due to inadvertent enrollment, excluding data after discontinuation of study intervention (last dose + 7 days)
<b>SS</b>	Data obtained during treatment period plus safety follow-up period from mITT population, regardless of adherence to study intervention

Abbreviations: EAS = efficacy analysis set; FAS = full analysis set; mITT = modified intent-to-treat; SS = safety analysis set.

### **9.3. Statistical Analyses**

#### **9.3.1. General Considerations**

Statistical analysis will be the responsibility of the sponsor or its designee. Statistical analysis for each ISA will be conducted individually and a combined analysis with both ISAs is not planned.

The SAP will be finalized prior to the unblinding of the first ISA.

Changes to the data analysis methods will require an amendment only if a principal feature of the master protocol is changed. Any other change to the data analysis methods, and the justification for making the change, will be described in the SAP or the CSR for each respective ISA.

Additional exploratory analyses of the data will be conducted as deemed appropriate.

Efficacy analyses will be conducted on all participants meeting study eligibility criteria randomly assigned to study intervention according to the treatment to which the participants are assigned and were exposed to at least one dose. The primary and key secondary efficacy analysis will be guided by 2 estimands, the “treatment regimen” estimand and the “efficacy” estimand to support global regulatory submissions and publications. For the “treatment regimen” estimand, the analysis will be conducted using FAS. To minimize missing data, participants randomly assigned to study intervention who prematurely discontinue study treatment will be encouraged to remain in the study, however, some participants may choose to permanently discontinue from the study which will lead to missing endpoints. Details on handling missing values can be found in Missing Value Imputation section. For the “efficacy” estimand, the analysis will be conducted using the EAS.

Safety analysis will be conducted using the SS. Selected safety analyses may be conducted after excluding the data after permanent discontinuation of the study intervention.

Unless otherwise noted, all tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, and the confidence interval will be calculated at 95%, 2-sided. In statistical summaries and analyses, participants will be analyzed as randomized. Countries in similar geographic regions with fewer than 10 participants, based on the all-randomized population, will be pooled to achieve a pooled country of at least 10 participants. All analyses using country in the model will use a pooled country, unless otherwise specified. The final pooling by country and geographic region will be finalized prior to data lock.

Baseline is defined as the last non missing measurement at or before the randomization visit (Visit 2) unless otherwise specified.

Analysis of covariance will be used to analyze continuous variables collected only at baseline and endpoint. The model will include treatment and strata (pooled country/geographic region, AHI stratum [moderate (AHI  $\geq 15$  and AHI  $< 30$ ), severe (AHI  $\geq 30$ )] and gender) as fixed effects and baseline as a covariate. The ANCOVA model for AHI analysis will include baseline AHI instead of the AHI stratum as a fixed covariate.

The MMRM analysis, a restricted-maximum-likelihood-based model, will be used to analyze continuous longitudinal variables. All the longitudinal observations at each scheduled postbaseline visit will be included in the analysis. The model will include the fixed class effects of treatment, strata (pooled country/geographic region and gender), visit, and treatment-by-visit interaction, as well as the continuous, fixed covariate of baseline value. For analyses of variables other than AHI, the AHI stratum will also be included in the model. Significance tests will be based on least-squares means and Type III tests.

For continuous measures, summary statistics may include sample size, mean, standard deviation, median, minimum, and maximum for both the actual and the change from baseline measurements. Least-square means and standard errors derived from the analysis models will also be displayed for the change from baseline measurements. Treatment comparisons will be displayed showing the treatment difference least-square means and the 95% confidence intervals for the treatment differences, along with the p-values for the treatment comparisons.

For categorical measures, summary statistics may include sample size, frequency, and percentages. Fisher's exact test or Pearson's chi-square test will be used for treatment comparisons unless otherwise specified.

Missing values for the primary and multiplicity adjusted endpoints at Week 52 and at Week 48 for SBP will be handled as follows. For efficacy analysis relative to the "efficacy" estimand, no explicit imputation will be performed. For efficacy analysis relative to the "treatment regimen" estimand, missing values will be imputed using multiple imputation based on the reason of intercurrent events. The statistical inference over multiple imputations will be guided by the method proposed by Rubin (1987). The intercurrent events and the resulting missing values will be handled as follows:

<b>Intercurrent events</b>	<b>Strategy to handle intercurrent events</b>	<b>Assumptions for missing values</b>	<b>Methods to handle missing values</b>
----------------------------	---	---------------------------------------	---

Study DC due to the COVID-19 pandemic (after other reasons for missing data are ruled out) or incorrect study procedure leading to invalid measurements ascertained while on treatment	Hypothetical	MAR	Use observed data without any imputation
All other study discontinuations	Treatment policy	MNAR	Multiple imputation. Details will be indicated in SAP

Abbreviations: COVID-19 = coronavirus disease-2019; DC = discontinuation; MAR = missing at random; MNAR = missing not at random.

Handling of missing, unused, and spurious data are addressed prospectively in the overall statistical methods described in the protocol and in the SAP, where appropriate. Adjustments to the planned analyses are described in the final CSR.

### 9.3.2. Primary Analysis

The primary objective of this study is to test the hypothesis that tirzepatide at the MTD (10 mg or 15 mg) is superior to placebo for participants with moderate to severe OSA on the mean AHI reduction from baseline to Week 52.

The primary analysis guided by the “treatment regimen” estimand will use FAS. Missing values will be imputed based on the strategy to handle intercurrent events described in Section 9.3.1, and more details about the imputation methods will be provided in the SAP. After the imputation, the primary efficacy comparison will be based on the contrast between tirzepatide at the MTD (10 mg or 15 mg) and placebo from the ANCOVA analysis of mean change from baseline to Week 52 in AHI using FAS as described in Section 9.3.1.

The primary analysis guided by the “efficacy” estimand will use EAS. The efficacy comparison will be based on the contrast between tirzepatide at the MTD (10 mg or 15 mg) and placebo at Week 52 (Visit 11) from the MMRM analysis of mean change from baseline in AHI using EAS as described in Section 9.3.1. More details about the model will be provided in the SAP.

Analysis aligned to each estimand will be evaluated at the full significance level of 0.05.

### 9.3.3. Analysis of Key Secondary Endpoints

The analyses for key secondary endpoints as in Section 3 will be performed for both the “treatment regimen” and “efficacy” estimand as described in Section 9.3.1 using the graphical testing scheme. The details of graphical testing scheme will be described in the SAP.

The analysis of the hierarchical combination PRO endpoint will be performed with the Finkelstein-Schoenfeld method, and the win ratio (Pocock et al. 2012) will be reported as the measure of treatment effect.

The last measurement prior to randomization for all three PRO endpoints will be used as baseline. For the “treatment regimen” estimand, missing values at Week 52 will be imputed through multiple imputations based on the reason of missingness as described in Section 9.3.1.

Analysis of percent change in AHI, percent change from baseline in body weight, CRP at the 52-week visit and change in SBP at the 48-week visit will be conducted in a manner similar to the primary efficacy analyses with baseline AHI stratum added in the model, and baseline of the corresponding variable as a covariate. Analysis of percent change in AHI will not include the baseline AHI stratum.

Comparisons at the 52-week visit between the treatments relative to the proportion of participants achieving  $\geq 50\%$  AHI reduction and  $AHI < 5$  or ( $AHI \geq 5$  and  $AHI \leq 14$  and  $ESS \leq 10$ ) will be conducted using logistic regression analysis including terms for treatment, pooled country, gender, and baseline AHI as a covariate.

Analysis aligned to each estimand will be evaluated at the full significance level of 0.05 contingent on reaching statistical significance of the primary objective.

#### **9.3.4. Treatment Group Comparability**

##### **9.3.4.1. Participant Disposition**

Participants who discontinue from the study will be identified, and the extent of their participation in the study will be reported for each ISA. A detailed description of participant disposition and the reasons for discontinuation will be summarized by treatment group for each ISA at the end of the study. Intervention-specific analyses will be detailed in each respective ISA.

##### **9.3.4.2. Participant Characteristics**

Participant characteristics and baseline clinical measures will be summarized for each treatment . For all participant characteristics, the summaries will include descriptive statistics for continuous measures (for example, sample size, mean, standard deviation, median, minimum, and maximum) and for categorical measures (for example, sample size, counts and percentages). Additional intervention-specific analyses will be specified in each respective ISA.

##### **9.3.4.3. Concomitant Therapy**

Concomitant medications used during the study will be summarized for each ISA.

#### **9.3.5. Safety Analyses**

Unless specified otherwise, safety assessments will compare safety of tirzepatide at the MTD (10 mg or 15 mg) with placebo irrespective of adherence to study intervention. Thus, safety analyses will be conducted using the SS.

**9.3.5.1. Adverse Events**

Adverse events will be classified by system organ class and preferred term as defined by the Medical Dictionary for Regulatory Activities.

All conditions existing prior to randomization at Visit 2 will be used as baseline. The postbaseline visits during the placebo-controlled phase will be included as the postbaseline period for analysis.

For events that are gender specific, the denominator and computation of the percentage will only include participants of the given gender.

**9.3.5.2. Hypoglycemic Events**

Incidence of documented symptomatic hypoglycemia events and severe hypoglycemia will be summarized and compared between tirzepatide at the MTD (10 mg or 15 mg) and placebo. Rate of hypoglycemic episodes will also be analyzed. Some analyses may be conducted excluding data after introducing another antihyperglycemic therapy.

**9.3.5.3. Gastrointestinal Events**

Summaries and analyses for incidence and severity of nausea, vomiting, and diarrhea will be provided by each treatment.

**9.3.5.4. Adjudicated Cardiovascular Events**

Listings of deaths, myocardial infarctions, strokes, and hospitalizations for unstable angina or heart failure confirmed by an independent Clinical Endpoint Committee (CEC) will be provided.

**9.3.5.5. Central Laboratory Measures and Vital Signs**

Values and change from baseline to postbaseline values of central laboratory measures and vital signs will be summarized and compared between tirzepatide at the MTD (10 mg or 15 mg) and placebo at each scheduled visit.

**9.3.5.6. Analysis of C-SSRS Data**

Suicide-related thoughts and behaviors occurring during treatment will be summarized based on responses to the C-SSRS consistent with the C-SSRS Scoring and Data Analysis Guide (The Columbia Lighthouse Project, 2013).

**9.3.6. Evaluation of Immunogenicity**

The frequency and percentage of participants with preexisting ADA and with TE ADA+ to tirzepatide will be tabulated. Treatment-emergent ADAs are defined as those with a titer 2-fold (1 dilution) greater than the minimum required dilution (1:10) of the ADA assay if no ADAs were detected at baseline (treatment-induced ADA), or those with a 4-fold (2 dilutions) increase in titer compared with baseline if ADAs were detected at baseline (treatment-boosted ADA). For the TE ADA+ participants, the distribution of maximum titers will be described. The frequency of neutralizing antibodies may also be tabulated in TE ADA+ participants. The relationship between the presence of antibodies and the PK parameters and PD response including safety and efficacy to tirzepatide may be assessed.

### 9.3.7. Other Analyses

#### 9.3.7.1. Health Economics

Analyses of actual and change from baseline in PRO scores will be conducted using linear models with baseline PRO scores, treatment, stratification factors and other factors that may be considered relevant. These variables will be specified in the SAP.

#### 9.3.7.2. Subgroup Analyses

The following subgroups will be analyzed using the “efficacy” estimand on percent change in AHI values from baseline to 52 week visit if there are sufficient numbers of participants in each treatment by subgroup (for example, 10%):

- Age (<50 years, ≥50 years),
- Baseline OSA severity (Moderate, Severe),
- Race,
- Ethnicity,
- Country,
- Gender (Male or Female),
- Baseline BMI (<30, ≥30 and <35, ≥35 and <40, ≥40 kg/m<sup>2</sup>), and
- Baseline ESS (ESS≤10, ESS>10).

Analyses for percent change from baseline in AHI will be performed using an MMRM model that includes the same fixed effects given for the primary analysis model plus factors of subgroup, 2-way interaction of subgroup and treatment, 2-way interaction of subgroup and visit, and 3-way interaction of treatment, visit and subgroup. The interaction of subgroup and treatment at the primary endpoint (Week 52) will be evaluated to assess the treatment by subgroup interaction. When analyzing OSA severity (Moderate, Severe) as a subgroup, the baseline AHI will be not be included as a covariate to avoid confounding. Additional subgroup analyses may also be performed.

More details on other analyses will be described in the SAP.

## 9.4. Interim Analysis

Based on the projected enrollment, approximately 3 interim analyses of safety will be conducted. The first interim analysis is planned to occur when approximately 20% of the anticipated number of participants are randomly assigned to study intervention or 6 months after the first participant is randomly assigned to study intervention, whichever occurs later, from one of the ISAs. Subsequent reviews will follow approximately every 6 months throughout the study.

Only the DMC is authorized to evaluate unblinded interim analyses. Study sites will receive information about interim results only if they need to know for the safety of their participants.

Unblinding details are specified in SAP and/or a separate unblinding plan document.

The DMC charter will describe the specific parameters of the planned interim analysis.

See Section 10.1.5, Appendix 1 And the DMC Charter for details related to the DMC.

The timing of dissemination of data summaries based on interim analyses is addressed in Section 10.1.6, Appendix 1.

## 9.5. Sample Size Determination

Approximately 206 participants per ISA will be randomly assigned to either tirzepatide or placebo in a 1:1 ratio (approximately 103 participants per treatment arm), and the statistical power is evaluated for the primary efficacy endpoint and key secondary combination PRO endpoint at a 2-sided significance level of 0.05. This sample size will provide

- At least 90% power to demonstrate superiority of tirzepatide at the MTD (10 mg or 15 mg) to placebo for the mean percent change from baseline in AHI, assuming 50% improvement, with a common standard deviation of 50%, and a dropout rate of 25%
- at least 90% power to demonstrate superiority of tirzepatide at the MTD (10 mg or 15 mg) to placebo for the hierarchical combination PRO endpoint using Finkelstein-Schoenfeld method (Finkelstein and Schoenfeld 1999) with a dropout rate of 25%

An upper limit of approximately 70% enrollment of male participants will be used to ensure a sufficiently large sample of female participants.

The sample size and power for the key secondary hierarchical combination PRO endpoint is estimated through simulations under the following assumptions (Blackman et al. 2016, Weaver et al. 2021):

Variable	Placebo	Tirzepatide at the MTD (10 mg or 15 mg)	Definition of Win
FOSQ-10 – change from baseline (mean ± SD)	1.24 ± 1.84	1.95 ± 1.84	If the change from baseline FOSQ-10 for TZP at MTD (10 mg or 15 mg) minus change from baseline FOSQ-10 for placebo $\geq 2.2$
FOSQ Vigilance – change from baseline (mean ± SD)	0.3 ± 0.5	0.5 ± 0.5	If the change from baseline FOSQ Vigilance for TZP at MTD (10 mg or 15 mg) minus change from baseline FOSQ Vigilance for placebo $\geq 0.44$
FOSQ Activity Level – change from baseline (mean ± SD)	0.3 ± 0.5	0.5 ± 0.5	If the change from baseline FOSQ Activity Level for TZP at MTD (10 mg or 15 mg) minus change from baseline FOSQ Activity Level for placebo $\geq 0.44$

Abbreviations: FOSQ = Functional Outcomes of Sleep Questionnaire; FOSQ-10 = FOSQ 10-item short form; MTD = maximum tolerated dose; SD = standard deviation; TZP = tirzepatide.

## **10. Supporting Documentation and Operational Considerations**

### **10.1. Appendix 1: Regulatory, Ethical, and Study Oversight Considerations**

#### **10.1.1. Regulatory and Ethical Considerations**

This study will be conducted in accordance with the protocol and with the following:

- Consensus ethical principles derived from international guidelines including the Declaration of Helsinki and Council for International Organizations of Medical Sciences International Ethical Guidelines
- Applicable ICH GCP Guidelines
- International Organization for Standardization (ISO) 14155
- Applicable laws and regulations

The protocol, protocol amendments, ICFs, IB, and other relevant documents (for example, advertisements) must be submitted to an IRB/IEC by the investigator and reviewed and approved by the IRB/IEC before the study is initiated.

Any amendments to the protocol will require IRB/IEC approval before implementation of changes made to the study design, except for changes necessary to eliminate an immediate hazard to study participants.

Protocols and any substantial amendments to the protocol will require health authority approval prior to initiation except for changes necessary to eliminate an immediate hazard to study participants.

The investigator will be responsible for the following:

- Providing written summaries of the status of the study to the IRB/IEC annually or more frequently in accordance with the requirements, policies, and procedures established by the IRB/IEC
- Notifying the IRB/IEC of SAEs or other significant safety findings as required by IRB/IEC procedures
- Providing oversight of study conduct for participants under their responsibility and adherence to requirements of 21 Code of Federal Regulations, ICH guidelines, the IRB/IEC, European regulation 536/2014 for clinical studies (if applicable), and all other applicable local regulations

Investigator sites are compensated for participation in the study as detailed in the clinical trial agreement.

#### **10.1.2. Financial Disclosure**

Investigators and sub-investigators will provide the sponsor with sufficient, accurate financial information as requested to allow the sponsor to submit complete and accurate financial certification or disclosure statements to the appropriate regulatory authorities. Investigators are responsible for providing information on financial interests during the course of the study and for 1 year after completion of the study.



### **10.1.3. Informed Consent Process**

The investigator or the investigator's representative will explain the nature of the study, including the risks and benefits, to the participant or the participant's legally authorized representative and answer all questions regarding the study.

Participants must be informed that their participation is voluntary. Participants or their legally authorized representatives will be required to sign a statement of informed consent that meets the requirements of 21 Code of Federal Regulations 50, local regulations, ICH guidelines, privacy and data protection requirements, where applicable, and the IRB/IEC or study center.

The medical record must include a statement that written informed consent was obtained before the participant was entered in the study and the date the written consent was obtained. The authorized person obtaining the informed consent must also sign the ICF.

Participants must be re-consented to the most current version of the ICF(s) during their participation in the study.

A copy of the ICF(s) must be provided to the participant or the participant's legally authorized representative and is kept on file.

Participants who are rescreened are required to sign a new ICF.

### **10.1.4. Data Protection**

Participants will be assigned a unique identifier by the sponsor. Any participant records, datasets or tissue samples that are transferred to the sponsor will contain the identifier only; participant names or any information which would make the participant identifiable will not be transferred.

The participant must be informed that the participant's personal study-related data will be used by the sponsor in accordance with local data protection law. The level of disclosure must also be explained to the participant who will be required to give consent for their data to be used as described in the informed consent.

The participant must be informed that their medical records may be examined by Clinical Quality Assurance auditors or other authorized personnel appointed by the sponsor, by appropriate IRB/IEC members, and by inspectors from regulatory authorities.

The sponsor has processes in place to ensure data protection, information security and data integrity. These processes include appropriate contingency plan(s) for appropriate and timely response in the event of a data security breach.

### **10.1.5. Committees Structure**

Prospective adjudication of major adverse CV events and pancreatic AEs will be performed by the independent adjudication committees for this study.

Sections 10.1.5.1 and 10.1.5.2, Appendix 1 outline additional information on pancreatic and CV adjudication.

An independent DMC for the interim analysis (Section 9.4) will include members with no conflict of financial interest. An independent DMC with members all external to the sponsor will be used to monitor participant safety in an unblinded fashion. For details on the DMC, refer to the DMC charter.

### 10.1.5.1. Cardiovascular Adjudicated Events

Deaths and nonfatal CV AEs will be adjudicated by a committee blinded to treatment assignment. The nonfatal CV AEs to be adjudicated include:

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

### 10.1.5.2. Pancreatitis Adjudicated Event

Acute pancreatitis is defined as an AE of special interest in this trial (Section 8.3.3).

The diagnosis of acute pancreatitis requires 2 of the following 3 features:

- abdominal pain, characteristic of acute pancreatitis (generally located in the epigastrium and radiates to the back in approximately half the cases) (Banks and Freeman 2006; Koizumi et al. 2006); the pain is often associated with nausea and vomiting
- serum amylase (total and/or pancreatic) and/or lipase  $\geq 3X$  ULN, and
- characteristic findings of acute pancreatitis on CT scan or MRI.

All suspected cases of acute or chronic pancreatitis will be adjudicated by a committee blinded to treatment assignment. In addition, AEs of severe or serious abdominal pain of unknown etiology will also be submitted to the adjudication committee to assess for possible pancreatitis or other pancreatic disease. Relevant data from participants with acute or chronic pancreatitis and those with severe or serious abdominal pain will be entered into a specifically designed CRF page. The adjudication committee representative will enter the results of adjudication in a corresponding CRF page.

### 10.1.6. Dissemination of Clinical Study Data

#### Reports

The sponsor will disclose a summary of study information, including tabular study results, on publicly available websites where required by local law or regulation.

The summary of results will be posted within the time frame specified by local law or regulation. If the study remains ongoing in some countries and a statistical analysis of an incomplete data set would result in analyses lacking scientific rigor (for example, underpowered) or compromise the integrity of the overall analyses (for example, trial not yet unblinded), the summary of results will be submitted within 1 year after the end of the study globally or as soon as available, whichever is earlier.

The publication policy for Study GPIF is outlined in Section 10.1.10, Appendix 1 and further described in the Clinical Trial Agreement.

**Data**

The sponsor provides access to all individual participant data collected during the trial, after anonymization, with the exception of PK or genetic data.

Data are available to request 6 months after the indication studied has been approved in the US and EU and after primary publication acceptance, whichever is later. No expiration date of data requests is currently set once data are made available.

Access is provided after a proposal has been approved by an independent review committee identified for this purpose and after receipt of a signed data sharing agreement.

Data and documents, including the study protocol, SAP, CSR, and blank or annotated CRFs, will be provided in a secure data sharing environment for up to 2 years per proposal.

For details on submitting a request, see the instructions provided at [www.vivli.org](http://www.vivli.org).

**10.1.7. Data Quality Assurance**

All participant data relating to the study will be recorded on printed or electronic CRFs unless transmitted to the sponsor or designee electronically (for example, laboratory data). The investigator is responsible for verifying that data entries are accurate and correct by physically or electronically signing the CRF.

The investigator must maintain accurate documentation (source data) that supports the information entered in the CRF.

The investigator must permit study-related monitoring, audits, IRB/IEC review, and regulatory agency inspections and provide direct access to source data documents.

Quality tolerance limits (QTLs) will be pre-defined to identify systematic issues that can impact participant safety and/or reliability of study results. These pre-defined parameters will be monitored during the study and important excursions from the QTLs and remedial actions taken will be summarized in the CSR.

Monitoring details describing strategy (for example, risk-based initiatives in operations and quality such as risk management and mitigation strategies and analytical risk-based monitoring), methods, responsibilities and requirements, including handling of noncompliance issues and monitoring techniques are provided in the Monitoring Plan.

The sponsor or designee is responsible for the data management of this study including quality checking of the data.

The sponsor assumes accountability for actions delegated to other individuals (for example, contract research organizations).

Study monitors will perform ongoing source data verification to confirm that data transcribed into the CRF by authorized site personnel are accurate, complete, and verifiable from source documents; that the safety and rights of participants are being protected; and that the study is being conducted in accordance with the currently approved protocol and any other study agreements, ICH GCP, and all applicable regulatory requirements.

Records and documents, including signed ICFs, pertaining to the conduct of this study must be retained by the investigator for the time period outlined in the Clinical Trial Agreement unless

local regulations or institutional policies require a longer retention period. No records may be destroyed during the retention period without the written approval of the sponsor. No records may be transferred to another location or party without written notification to the sponsor.

In addition, sponsor or its representatives will periodically check a sample of the participant data recorded against source documents at the study site. The study may be audited by sponsor or its representatives, and/or regulatory agencies at any time. Investigators will be given notice before an audit occurs.

### **Data Capture System**

The investigator is responsible for ensuring the accuracy, completeness, legibility, and timeliness of the data reported to the sponsor.

An EDC system will be used in this study for the collection of CRF data. The investigator maintains a separate source for the data entered by the investigator or designee into the sponsor-provided EDC system. The investigator is responsible for the identification of any data to be considered source and for the confirmation that data reported are accurate and complete by signing the CRF.

Additionally, COA data (clinician-reported outcome instrument) for suicidality assessments will be collected by the authorized study personnel, via a paper source document and will be transcribed by the authorized study personnel into the EDC system.

Additionally, electronic COA data (participant-focused outcome instrument) will be directly recorded by the participant, into an instrument (for example, handheld smart phone or tablet). The electronic COA data will serve as the source documentation and the investigator does not maintain a separate, written or electronic record of these data.

Data collected via the sponsor-provided data capture system(s) will be stored at third parties. The investigator will have continuous access to the data during the study and until decommissioning of the data capture system(s). Prior to decommissioning, the investigator will receive or access an archival copy of pertinent data for retention.

Data managed by a central vendor, such as laboratory test data, will be stored electronically in the central vendor's database system and reports/electronic transfers will be provided to the investigator for review and retention. Data will subsequently be transferred from the central vendor to the sponsor data warehouse.

Data collected on the actigraphy device will be transferred electronically to the sponsor data warehouse, via a third party.

Data from complaint forms submitted to the sponsor will be encoded and stored in the global PC management system.

#### **10.1.8. Source Documents**

Source documents provide evidence for the existence of the participant and substantiate the integrity of the data collected. Source documents are filed at the investigator's site.

Data reported on or entered in the CRF and are transcribed from source documents must be consistent with the source documents or the discrepancies must be explained. The investigator

may need to request previous medical records or transfer records, depending on the study. Also, current medical records must be available.

Definition of what constitutes source data can be found in Section 10.1.7, Appendix 1.

### **10.1.9. Study and Site Start and Closure**

#### **First Act of Recruitment**

The study start date is the date on which the clinical study will be open for recruitment of participants. The first act of recruitment is the first site open and will be the study start date.

#### **Study or Site Termination**

The sponsor or sponsor's designee reserves the right to close the study site or terminate the study at any time for any reason at the sole discretion of the sponsor. Study sites will be closed upon study completion. A study site is considered closed when all required documents and study supplies have been collected and a study-site closure visit has been performed.

The investigator may initiate study-site closure at any time, provided there is reasonable cause and sufficient notice is given in advance of the intended termination.

Reasons for the early closure of a study site by the sponsor or investigator may include but are not limited to:

- For study termination:
  - Discontinuation of further study intervention development
- For site termination:
  - Failure of the investigator to comply with the protocol, the requirements of the IRB/IEC or local health authorities, the sponsor's procedures, or GCP guidelines
  - Inadequate recruitment (evaluated after a reasonable amount of time) of participants by the investigator
  - Total number of participants included earlier than expected.

If the study is prematurely terminated or suspended, the sponsor shall promptly inform the investigators, the IECs/IRBs, the regulatory authorities, and any contract research organization(s) used in the study of the reason for termination or suspension, as specified by the applicable regulatory requirements. The investigator shall promptly inform the participant and should assure appropriate participant therapy and/or follow-up.

### **10.1.10. Publication Policy**

In accordance with the sponsor's publication policy, the results of this study will be submitted for publication by a peer-reviewed journal.

### **10.1.11. Investigator Information**

Researchers with appropriate education, training, and experience, as determined by the sponsor, will participate as investigators in this clinical trial.

**10.1.12. Sample Retention**

Sample retention enables use of new technologies, response to regulatory questions, and investigation of variable response that may not be observed until later in the development of tirzepatide or after tirzepatide becomes commercially available.

<b>Sample Type</b>	<b>Custodian</b>	<b>Retention Period After Last Participant Visit<sup>a</sup></b>
Exploratory Biomarker Samples	Sponsor or Designee	15 years
PK Samples	Sponsor or Designee	2 years
Genetics	Sponsor or Designee	15 years
Immunogenicity (ADA) Samples	Sponsor or Designee	15 years

Abbreviations: ADA = antidrug antibody; PK = pharmacokinetic.

<sup>a</sup>Retention periods may differ locally.

## 10.2. Appendix 2: Clinical Laboratory Tests

The tests detailed in the table below will be performed by the central laboratory.

Local laboratory results are only required in the event that the central laboratory results are not available in time for either study intervention administration and/or response evaluation. If a local sample is required, it is important that the sample for central analysis is obtained at the same time.

In circumstances where the sponsor approves local laboratory testing in lieu of central laboratory testing (in the table below), the local laboratory must be qualified in accordance with applicable local regulations.

Protocol-specific requirements for inclusion or exclusion of participants are detailed in Section 5 of the protocol.

Additional tests may be performed at any time during the study as determined necessary by the investigator or required by local regulations.

Investigators must document their review of the laboratory safety results.

Laboratory results that could unblind the study will not be reported to investigative sites or other blinded personnel.

Clinical Laboratory Tests	Comments
<b>Hematology</b>	Assayed by Lilly-designated laboratory
Hemoglobin	
Hematocrit	
Erythrocyte count (RBCs)	
Mean cell volume	
Mean cell hemoglobin	
Mean cell hemoglobin concentration	
Leukocytes (WBCs)	
Differential	
Percent and/or Absolutes Count of:	
Neutrophils, segmented	
Bands	
Lymphocytes	
Monocytes	
Eosinophils	
Basophils	
Platelets	
Cell morphology (RBCs and WBCs)	
<b>Clinical Chemistry</b>	Assayed by Lilly-designated laboratory
Sodium	
Potassium	
Chloride	
Bicarbonate	

<b>Clinical Laboratory Tests</b>	<b>Comments</b>
Total bilirubin	
Direct bilirubin	
ALP	
ALT	
AST	
BUN	
Creatinine	
CK	
Uric acid	
Albumin	
Calcium	
Glucose	
<b>Lipid Panel</b>	Assayed by Lilly designated laboratory.
Cholesterol	
Triglycerides	
HDL-C	.
LDL-C	Generated by Lilly-designated laboratory. If Triglycerides are >400; direct LDL will be measured.
VLDL-C	Generated by Lilly-designated laboratory.
<b>Hormones (female)</b>	
Serum Pregnancy	Assayed by Lilly-designated laboratory.
Urine Pregnancy	Assayed and Evaluated locally
FSH	Assayed by Lilly-designated laboratory.
<b>Urine Chemistry</b>	Assayed by Lilly-designated laboratory.
Albumin	
Creatinine	
<b>Calculations</b>	Generated by Lilly-designated laboratory.
eGFR (CKD-EPI)	
UACR	
<b>PK Samples – Tirzepatide</b>	Assayed by Lilly-designated laboratory. Results will not be provided to the investigative sites.
<b>Immunogenicity (ADA) Samples</b>	Assayed by Lilly-designated laboratory. Results will not be provided to the investigative sites.
Anti-tirzepatide antibodies OR Tirzepatide antibodies	
Anti-tirzepatide antibodies neutralization OR Tirzepatide antibodies neutralization	
<b>Additional Testing</b>	Assayed by Lilly-designated laboratory.
HbA1c	
Calcitonin	
Pancreatic Amylase	
Lipase	



Clinical Laboratory Tests	Comments
Insulin	Results will not be provided to the investigative sites.
C-Peptide	Results will not be provided to the investigative sites.
Free Fatty Acids	Results will not be provided to the investigative sites.
hsCRP	Results will not be provided to the investigative sites.
Cystatin-C	
TSH	
<b>Stored Samples</b>	Assayed by Lilly-designated laboratory. Results will not be provided to the investigative sites.
Genetics Sample	
Exploratory Biomarker Samples	
Serum	
Plasma (EDTA)	
Plasma (P800)	

Abbreviations: ADA = antidrug antibody; ALP = alkaline phosphatase; ALT = alanine aminotransferase;; AST = aspartate aminotransferase; BUN = blood urea nitrogen; CK = creatinine kinase, CKD-EPI = Chronic Kidney Disease-Epidemiology; EDTA = ethylenediaminetetraacetic acid; eGFR = estimated glomerular filtration rate; FSH = follicle-stimulating hormone; HbA1c = hemoglobin A1c; HDL-C = high-density lipoprotein cholesterol; hsCRP = C-Reactive Protein, high-sensitivity; IWRS = interactive web-response system; LDL-C = low-density lipoprotein cholesterol; PK = pharmacokinetics; RBC = red blood cells; TSH = thyroid-stimulating hormone; UACR = urine albumin/creatinine ratio; VLDL-C = very-low-density lipoprotein cholesterol; WBC = white blood cells.

### 10.2.1. Laboratory Samples to be Obtained at the Time of a Systemic Hypersensitivity Event

#### Purpose of collecting samples after a systemic hypersensitivity event

The samples listed in this appendix are not collected for acute study participant management. The sponsor will use the laboratory tests results from these samples to characterize hypersensitivity events across the clinical development program.

#### When to collect samples after a systemic hypersensitivity event occurs

Collect the samples listed below if a systemic hypersensitivity event is suspected. The timing should be as designated in the table, assuming the participant has been stabilized.

Obtain follow-up pre-dose samples at the next regularly scheduled laboratory sample collection (ideally prior to the next dose after the event) to assess post-event return-to-baseline values.

Timing	Sample Type	Laboratory Test <sup>a</sup>
Collect from 30 minutes to 4 hours after the start of the event.  ● Note: The optimal collection time is from 1 to 2 hours after the start of event.	Serum	total tryptase
	Serum	complements (C3, C3a, and C5a)
	Serum	cytokine panel (IL-6, IL-1 $\beta$ , IL-10 or any cytokine panel that includes these 3 cytokines)
Collect only if not already collected on the same day as the event.  ● Note: If collecting, collect up to 12 hours after the start of the event.	Serum	Tirzepatide ADA
	Serum/plasma	Tirzepatide concentration

Abbreviations: ADA = anti-drug antibodies; IL = interleukin.

<sup>a</sup> All samples for hypersensitivity testing will be assayed by Lilly-designated laboratory. Results will not be provided to the study site. If samples are not collected or are collected outside the specified time period, this will not be considered a protocol deviation.

### What information to record

Record the date and time when the samples are collected.

### Allowed additional testing for participant management

The investigator may perform additional tests locally, if clinically indicated, for acute study participant management.

### 10.3. Appendix 3: Adverse Events and Serious Adverse Events: Definitions and Procedures for Recording, Evaluating, Follow-up, and Reporting

- The definitions and procedures detailed in this appendix are in accordance with International Organization for Standardization (ISO) 14155.
- Both the investigator and the sponsor will comply with all local medical device reporting requirements.
- The detection and documentation procedures described in this protocol apply to all sponsor medical devices provided for use in the study. See Section 6.1.1 for the list of sponsor medical devices).

#### 10.3.1. Definition of AE

AE Definition
<ul style="list-style-type: none"> <li>• An AE is any untoward medical occurrence in a participant administered a pharmaceutical product and which does not necessarily have a causal relationship with the study intervention. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product.</li> <li>• An AE is any untoward medical occurrence, unintended disease or injury, or untoward clinical signs (including abnormal laboratory finding) in study participants, users, or other persons, whether or not related to the investigational medical device. This definition includes events related to the investigational medical device or comparator and events related to the procedures involved except for events in users or other persons, which only include events related to investigational devices.</li> </ul>

Events Meeting the AE Definition
<ul style="list-style-type: none"> <li>• Any abnormal laboratory test results (hematology, clinical chemistry, or urinalysis) or other safety assessments (for example, ECG, radiological scans, vital signs measurements), including those that worsen from baseline, considered clinically significant in the medical and scientific judgment of the investigator (that is, not related to progression of underlying disease).</li> <li>• Exacerbation of a chronic or intermittent pre-existing condition including either an increase in frequency and/or intensity of the condition.</li> <li>• New condition detected or diagnosed after study intervention administration even though it may have been present before the start of the study.</li> <li>• Signs, symptoms, or the clinical sequelae of a suspected drug-drug interaction.</li> <li>• Medication error, misuse, or abuse of study intervention, including signs, symptoms, or clinical sequelae.</li> <li>• Lack of efficacy or failure of expected pharmacological action per se will not be reported as an AE or SAE. Such instances will be captured in the efficacy assessments.</li> </ul>

However, the signs, symptoms, and/or clinical sequelae resulting from lack of efficacy will be reported as AE or SAE if they fulfill the definition of an AE or SAE.

<b>Events NOT Meeting the AE Definition</b>
<ul style="list-style-type: none"> <li>• Any clinically significant abnormal laboratory findings or other abnormal safety assessments that are associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant’s condition.</li> <li>• The disease/disorder being studied or expected progression, signs, or symptoms of the disease/disorder being studied, unless more severe than expected for the participant’s condition.</li> <li>• Medical or surgical procedure (for example, endoscopy, appendectomy): the condition that leads to the procedure is the AE.</li> <li>• Situations in which an untoward medical occurrence did not occur (social and/or convenience admission to a hospital).</li> <li>• Anticipated day-to-day fluctuations of pre-existing disease(s) or condition(s) present or detected at the start of the study that do not worsen.</li> </ul>

**10.3.2. Definition of SAE**

<b>An SAE is defined as any untoward medical occurrence that, at any dose, meets one or more of the criteria listed:</b>
<b>a. Results in death</b>
<b>b. Is life-threatening</b> The term <i>life-threatening</i> in the definition of <i>serious</i> refers to an event in which the participant was at risk of death at the time of the event. It does not refer to an event, which hypothetically might have caused death, if it were more severe.
<b>c. Requires inpatient hospitalization or prolongation of existing hospitalization</b> <ul style="list-style-type: none"> <li>• In general, hospitalization signifies that the participant has been admitted to hospital or emergency ward (usually involving at least an overnight stay) for observation and/or treatment that would not have been appropriate in the physician’s office or outpatient setting. Complications that occur during hospitalization are AEs. If a complication prolongs hospitalization or fulfills any other serious criteria, the event is serious. When in doubt as to whether hospitalization occurred or was necessary, the AE should be considered serious.</li> <li>• Hospitalization for elective treatment of a pre-existing condition that did not worsen from baseline is not considered an AE.</li> </ul>
<b>d. Results in persistent disability/incapacity</b> <ul style="list-style-type: none"> <li>• The term disability means a substantial disruption of a person’s ability to conduct normal life functions.</li> <li>• This definition is not intended to include experiences of relatively minor medical significance such as uncomplicated headache, nausea, vomiting, diarrhea, influenza,</li> </ul>

<p>and accidental trauma (for example, sprained ankle) which may interfere with or prevent everyday life functions but do not constitute a substantial disruption.</p>
<p><b>e. Is a congenital anomaly/birth defect</b></p> <ul style="list-style-type: none"> <li>Abnormal pregnancy outcomes (for example, spontaneous abortion, fetal death, stillbirth, congenital anomalies, ectopic pregnancy) are considered SAEs.</li> </ul>
<p><b>f. Other situations:</b></p> <ul style="list-style-type: none"> <li>Medical or scientific judgment should be exercised by the investigator in deciding whether SAE reporting is appropriate in other situations such as important medical events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the participant or may require medical or surgical intervention to prevent one of the other outcomes listed in the above definition. These events should usually be considered serious.</li> <li>Examples of such events include invasive or malignant cancers, intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias or convulsions that do not result in hospitalization, or development of drug dependency or drug abuse.</li> </ul>
<p><b>g.</b> Resulted in medical or surgical intervention to prevent life-threatening illness or injury or permanent impairment to a body structure or a body function.</p>

**10.3.3. Definition of Product Complaints**

<b>Product Complaint</b>
<ul style="list-style-type: none"> <li>A PC is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, durability, reliability, safety, effectiveness or performance of a study intervention. When the ability to use the study intervention safely is impacted, the following are also PCs:                     <ul style="list-style-type: none"> <li>Deficiencies in labeling information, and</li> <li>Use errors for device or drug-device combination products due to ergonomic design elements of the product.</li> </ul> </li> <li>Product complaints related to study interventions used in clinical trials are collected in order to ensure the safety of participants, monitor quality, and to facilitate process and product improvements.</li> <li>Investigators will instruct participants to contact the site as soon as possible if he or she has a PC or problem with the study intervention so that the situation can be assessed.</li> <li>An event may meet the definition of both a PC and an AE/SAE. In such cases, it should be reported as both a PC and as an AE/SAE.</li> </ul>

**10.3.4. Recording and Follow-Up of AE and/or SAE and Product Complaints****AE, SAE, and PC Recording**

- When an AE/SAE/PC occurs, it is the responsibility of the investigator to review all documentation (for example, hospital progress notes, laboratory reports, and diagnostics reports) related to the event.
- The investigator will then record all relevant AE/SAE/PC information in the participant's medical records, in accordance with the investigator's normal clinical practice. AE/SAE information is reported on the appropriate CRF page and PC information is reported on the PC Form.

Note: An event may meet the definition of both a PC and an AE/SAE. In such cases, it should be reported as both a PC and as an AE/SAE.

- It is **not** acceptable for the investigator to send photocopies of the participant's medical records to sponsor or designee in lieu of completion of the CRF page for AE/SAE and the PC Form for product complaints.
- There may be instances when copies of medical records for certain cases are requested by sponsor or designee. In this case, all participant identifiers, with the exception of the participant number, will be redacted on the copies of the medical records before submission to sponsor or designee.
- The investigator will attempt to establish a diagnosis of the event based on signs, symptoms, and/or other clinical information. Whenever possible, the diagnosis (not the individual signs/symptoms) will be documented as the AE/SAE.

**Assessment of Intensity**

The investigator will make an assessment of intensity for each AE and SAE reported during the study and assign it to one of the following categories:

- **Mild:** A type of AE that is usually transient and may require only minimal treatment or therapeutic intervention. The event does not generally interfere with usual activities of daily living.
- **Moderate:** A type of AE that is usually alleviated with additional specific therapeutic intervention. The event interferes with usual activities of daily living, causing discomfort but poses no significant or permanent risk of harm to the research participant.
- **Severe:** A type of AE that interrupts usual activities of daily living, or significantly affects clinical status, or may require intensive therapeutic intervention. An AE that is assessed as severe should not be confused with an SAE. Severe is a category utilized for rating the intensity of an event; and both AEs and SAEs can be assessed as severe.

An event is defined as "serious" when it meets at least one of the predefined outcomes as described in the definition of an SAE, NOT when it is rated as severe.

### Assessment of Causality

- The investigator is obligated to assess the relationship between study intervention and each occurrence of each AE/SAE. The investigator will use clinical judgment to determine the relationship.
- A “reasonable possibility” of a relationship conveys that there are facts, evidence, and/or arguments to suggest a causal relationship, rather than a relationship cannot be ruled out.
- Alternative causes, such as underlying disease(s), concomitant therapy, and other risk factors, as well as the temporal relationship of the event to study intervention administration will be considered and investigated.
- The investigator will also consult the IB in their assessment.
- For each AE/SAE, the investigator **must** document in the medical notes that he/she has reviewed the AE/SAE and has provided an assessment of causality.
- There may be situations in which an SAE has occurred and the investigator has minimal information to include in the initial report to sponsor or designee. However, it is very important that the investigator always make an assessment of causality for every event before the initial transmission of the SAE data to sponsor or designee.
- The investigator may change their opinion of causality in light of follow-up information and send an SAE follow-up report with the updated causality assessment.
- The causality assessment is one of the criteria used when determining regulatory reporting requirements.

### Follow-Up of AEs and SAEs

- The investigator is obligated to perform or arrange for the conduct of supplemental measurements and/or evaluations as medically indicated or as requested by sponsor or designee to elucidate the nature and/or causality of the AE or SAE as fully as possible. This may include additional laboratory tests or investigations, histopathological examinations, or consultation with other health care professionals.
- If a participant dies during participation in the study or during a recognized follow-up period, the investigator will provide sponsor or designee with a copy of any post-mortem findings including histopathology.

### 10.3.5. Reporting of SAEs

#### SAE Reporting via an Electronic Data Collection Tool

- The primary mechanism for reporting an SAE will be the electronic data collection tool.
- If the electronic system is unavailable, then the site will use the SAE paper form (see next section) in order to report the event within 24 hours.
- The site will enter the SAE data into the electronic system as soon as it becomes available.

- After the study is completed at a given site, the electronic data collection tool will be taken off-line to prevent the entry of new data or changes to existing data.
- If a site receives a report of a new SAE from a study participant or receives updated data on a previously reported SAE after the electronic data collection tool has been taken off-line, then the site can report this information on an SAE paper form (see next section) or to the sponsor or designee by telephone.
- Contacts for SAE reporting can be found on the SAE form.

#### **SAE Reporting via Paper Form**

- Facsimile transmission of the SAE paper form is the preferred method to transmit this information to the sponsor or designee.
- Initial notification via telephone does not replace the need for the investigator to complete and sign the SAE CRF pages within the designated reporting time frames.
- Contacts for SAE reporting can be found on the SAE form.

### **10.3.6. Regulatory Reporting Requirements**

#### **SAE Regulatory Reporting**

- Prompt notification by the investigator to the sponsor of an SAE is essential so that legal obligations and ethical responsibilities toward the safety of participants and the safety of a study intervention under clinical investigation are met.
- The sponsor has a legal responsibility to notify both the local regulatory authority and other regulatory agencies about the safety of a study intervention under clinical investigation. The sponsor will comply with country-specific regulatory requirements relating to safety reporting to the regulatory authority, IRB/IEC, and investigators.
- An investigator who receives an investigator safety report describing an SAE or other specific safety information (for example, summary or listing of SAEs) from the sponsor will review and then file it along with the IB and will notify the IRB/IEC, if appropriate according to local requirements.

### **10.3.7. Special Safety Topics**

#### **10.3.7.1. Hypoglycemia**

Participants will be trained by authorized study personnel about signs and symptoms of hypoglycemia and how to treat hypoglycemia, and how to collect appropriate information for each episode of hypoglycemia.

Hypoglycemia may be identified by spontaneous reporting of symptoms from participants (whether confirmed or unconfirmed by simultaneous glucose values) or by blood glucose samples collected during study visits.



All participants who develop diabetes during the study will be provided with glucometers. Participants without diabetes may, at the investigator's discretion, be given glucometers to assist in the evaluation of reported symptoms consistent with hypoglycemia. Participants receiving glucometers will be provided a diary to record relevant information (for example, glucose values, symptoms).

Hypoglycemic episodes will be recorded in the eDiary and should not be recorded as AEs unless the event meets serious criteria. If a hypoglycemic event meets severe criteria (see definition below), it should be recorded as serious on the AE and SAE CRFs, and reported to Lilly as an SAE.

Investigators should use the following classification of hypoglycemia (ADA 2020):

### **Level 1 hypoglycemia:**

**Glucose <70 mg/dL (3.9 mmol/L) and  $\geq$  54 mg/dL (3.0 mmol/L):** Level 1 hypoglycemia can alert a person to take action such as treatment with fast-acting carbohydrates. Providers should continue to counsel participants to treat hypoglycemia at this glucose alert value.

### **Level 2 hypoglycemia:**

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

### **Level 3 hypoglycemia:**

**Severe hypoglycemia (in adults):** A severe event characterized by altered mental and/or physical status requiring assistance for treatment of hypoglycemia. For example, participants had altered mental status, and could not assist in their own care, or were semiconscious or unconscious, or experienced coma with or without seizures, and the assistance of another person was needed to actively administer carbohydrate, glucagon, or other resuscitative actions. Glucose measurements may not be available during such an event, but neurological recovery attributable to the restoration of glucose concentration to normal is considered sufficient evidence that the event was induced by a low glucose concentration.

- The determination of a hypoglycemic event as an episode of severe hypoglycemia, as defined above, is made by the investigator based on the medical need of the participant to have required assistance and is not predicated on the report of a participant simply having received assistance.
- If a hypoglycemic event meets the criteria of severe hypoglycemia, the investigator must record the event as serious on the AE CRF and report it to Lilly as an SAE.

### **Nocturnal hypoglycemia:**

Nocturnal hypoglycemia is a hypoglycemia event (including severe hypoglycemia) that **occurs at night** and presumably during sleep.

To avoid duplicate reporting, all consecutive BG values <70 mg/dL (<3.9 mmol/L) occurring within a 1-hour period may be considered to be a single hypoglycemic event (Weinberg et al. 2010; Danne et al. 2013).

#### **10.3.7.2. Hypersensitivity Reactions**

Many drugs, including oral agents and biologic agents, carry the risk of systemic hypersensitivity reactions. If such a reaction occurs, additional data should be provided to the sponsor in the designated CRFs.

Sites should have appropriately trained medical staff and appropriate medical equipment available when study participants are receiving study intervention. It is recommended that participants who experience a systemic hypersensitivity reaction be treated per national and international guidelines.

In the case of a suspected systemic hypersensitivity event, additional blood samples should be collected as described in Section 10.2.1. Laboratory results are provided to the sponsor via the central laboratory.

#### **10.3.7.3. Injection-Site Reactions**

Symptoms and signs of a local ISR may include erythema, induration, pain, pruritus, and edema.

If an ISR is reported by a participant or parent or guardian or site staff, the ISR CRF will be used to capture additional information about this reaction, for example, injection site pain, degree and area of erythema, induration, pruritis and edema.

At the time of AE occurrence, samples will be collected for measurement of tirzepatide ADAs and tirzepatide concentration.

## 10.4. Appendix 4: Contraceptive and Barrier Guidance

### 10.4.1. Definitions

Word/Phrase	Definition
Women of child bearing potential	<p>Females are considered a WOCBP if</p> <ul style="list-style-type: none"> <li>• they have had at least 1 cycle of menses, or</li> <li>• they have Tanner 4 breast development.</li> </ul> <p>Any amount of spotting should be considered menarche. If Tanner Staging of breasts is performed as part of study procedures, please refer to the Reproductive, Pregnancy and Pediatrics Safety Committee Safety Guidance for Children in Clinical Trial regarding Tanner staging.</p>
Women not of child bearing potential	<p>Females are considered women not of child bearing potential if</p> <ul style="list-style-type: none"> <li>• they have a congenital anomaly such as Mullerian agenesis</li> <li>• they are infertile due to surgical sterilization, or</li> <li>• they are post-menopausal.</li> </ul> <p>Examples of surgical sterilization include: hysterectomy, bilateral oophorectomy, or tubal ligation.</p>
Post-menopausal state	<p>The post-menopausal state should be defined as:</p> <ol style="list-style-type: none"> <li>1. A woman at any age at least 6 weeks post-surgical bilateral oophorectomy with or without hysterectomy, confirmed by operative note; or</li> <li>2. A woman at least 40 years of age and up to 55 years old with an intact uterus, not on hormone therapy, who has had cessation of menses for at least 12 consecutive months without an alternative medical cause, AND with a follicle-stimulating hormone &gt;40 mIU/mL; or</li> <li>3. A woman 55 or older not on hormone therapy, who has had at least 12 months of spontaneous amenorrhea; or</li> <li>4. A woman at least 55 years of age with a diagnosis of menopause prior to starting hormone replacement therapy</li> </ol> <p>* Women should not be taking medications during amenorrhea such as oral contraceptives, hormones, gonadotropin-releasing hormone, anti-estrogens, SERMs, or chemotherapy that could induce transient amenorrhea.</p>

Abbreviations: SERM = selective estrogen receptor modulators; WOCBP = woman of child bearing potential.

**10.4.2. Contraception Guidance**

**10.4.2.1. Females**

**WOCBP who are completely abstinent as their preferred and usual lifestyle, or in a same sex relationship, as part of their preferred and usual lifestyle**

Must...	Must not...
<ul style="list-style-type: none"> <li>• agree to either remain abstinent, or</li> <li>• stay in a same sex relationship without sexual relationships with males</li> </ul>	<ul style="list-style-type: none"> <li>• use periodic abstinence methods                             <ul style="list-style-type: none"> <li>○ calendar</li> <li>○ ovulation</li> <li>○ symptothermal, or</li> <li>○ post-ovulation</li> </ul> </li> <li>• declare abstinence just for the duration of a trial, or</li> <li>• use the withdrawal method</li> </ul>

**WOCBP who are NOT completely abstinent as their preferred and usual lifestyle, or in a same sex relationship, as part of their preferred and usual lifestyle**

Topic	Condition
Pregnancy testing	Negative serum result at screening, followed by a negative urine result within 24 hours prior to treatment exposure.
Contraception	Agree to use 2 forms of effective contraception, where at least one form must be highly effective (less than 1% failure rate) for the duration of the trial and for 30 days thereafter.

**Examples of different forms of contraception:**

Methods	Examples
Highly effective contraception	<ul style="list-style-type: none"> <li>• combination oral contraceptive pill and mini-pill</li> <li>• implanted contraceptives</li> <li>• injectable contraceptives</li> <li>• contraceptive patch (only women &lt;198 pounds or 90 kg)</li> <li>• total abstinence</li> <li>• vasectomy (if only sexual partner)</li> <li>• fallopian tube implants (if confirmed by hysterosalpingogram)</li> <li>• combined contraceptive vaginal ring, or</li> <li>• intrauterine devices</li> </ul>
Effective contraception	<ul style="list-style-type: none"> <li>• male or female condoms with spermicide</li> <li>• diaphragms with spermicide or cervical sponges</li> <li>• barrier method with use of a spermicide                             <ul style="list-style-type: none"> <li>○ condom with spermicide</li> </ul> </li> </ul>

	<ul style="list-style-type: none"> <li>○ diaphragm with spermicide, or</li> <li>○ female condom with spermicide</li> </ul> <p>Note: The barrier method must include use of a spermicide (that is, condom with spermicide, diaphragm with spermicide, female condom with spermicide) to be considered effective.</p>
Ineffective forms of contraception	<ul style="list-style-type: none"> <li>● spermicide alone</li> <li>● immunocontraceptives</li> <li>● periodic abstinence</li> <li>● fertility awareness (calendar method, temperature method, combination of above 2, cervical mucus, symptothermal)</li> <li>● withdrawal</li> <li>● post coital douche</li> <li>● lactational amenorrhea</li> </ul>

**10.4.2.2. Males**

The table below describes contraception guidance for all men.

Topic	Guidance
For all men	should refrain from sperm donation for the duration of the study and for 120 days (4 months) from the last dose of study intervention received
Contraception for men with partners of childbearing potential	<ul style="list-style-type: none"> <li>● either remain abstinent (if this is their preferred and usual lifestyle), or</li> <li>● must use condoms during intercourse for the duration of the study, <b>and</b></li> <li>● for 120 days (4 months) from the last dose of study intervention received</li> </ul>
Contraception for men in exclusively same sex relationships, as their preferred and usual lifestyle	Are not required to use contraception

Examples of highly effective, effective and unacceptable methods of contraception are listed in Section 10.4.2.1, Appendix 4.

## 10.5. Appendix 5: Genetics

### Use/Analysis of DNA

Genetic variation may impact a participant's response to study intervention, susceptibility to, and severity and progression of disease. Variable response to study intervention may be due to genetic determinants that impact drug absorption, distribution, metabolism, and excretion; mechanism of action of the drug; disease etiology; and/or molecular subtype of the disease being treated. Therefore, where local regulations and IRB/IEC allow, a blood sample will be collected for DNA analysis from consenting participants.

DNA samples will be used for research related to tirzepatide, OSA, obesity and related diseases. They may also be used to develop tests/assays including diagnostic tests related to tirzepatide, OSA, and obesity. Genetic research may consist of the analysis of 1 or more candidate genes or the analysis of genetic markers throughout the genome or analysis of the entire genome (as appropriate).

The samples may be analyzed as part of a multi-study assessment of genetic factors involved in the response to tirzepatide or study interventions of this class to understand study disease or related conditions.

The results of genetic analyses may be reported in the CSR or in a separate study summary.

The sponsor will store the DNA samples in a secure storage space with adequate measures to protect confidentiality.

The samples will be retained while research on tirzepatide, OSA, and obesity continues but no longer than the sample retention limits described in Section 10.1.12, Appendix 1.

## 10.6. Appendix 6: Liver Safety: Suggested Actions and Follow-up Assessments

### 10.6.1. Hepatic Evaluation Testing

See Sections 10.6.2 and 10.6.3, Appendix 6 for guidance on appropriate test selection.

The Lilly-designated central laboratory must complete the analysis of all selected testing except for microbiology testing.

Local testing may be performed *in addition to central testing* when necessary for immediate participant management.

Results will be reported if a validated test or calculation is available.

<b>Hematology</b>	<b>Clinical Chemistry</b>
Hemoglobin	Total bilirubin
Hematocrit	Direct bilirubin
Erythrocytes (RBCs – red blood cells)	Alkaline phosphatase (ALP)
Leukocytes (WBCs – white blood cells)	Alanine aminotransferase (ALT)
Differential:	Aspartate aminotransferase (AST)
Neutrophils, segmented	Gamma-glutamyl transferase (GGT)
Lymphocytes	Creatine kinase (CK)
Monocytes	<b>Other Chemistry</b>
Basophils	Acetaminophen
Eosinophils	Acetaminophen protein adducts
Platelets	Alkaline phosphatase isoenzymes
Cell morphology (RBC and WBC)	Ceruloplasmin
	Copper
<b>Coagulation</b>	Ethyl alcohol (EtOH) (quantitative)
Prothrombin time, INR (PT-INR)	Haptoglobin
<b>Serology</b>	Immunoglobulin A (IgA) (quantitative)
Hepatitis A virus (HAV) testing:	Immunoglobulin G (IgG) (quantitative)
HAV total antibody	Immunoglobulin M (IgM) (quantitative)
HAV IgM antibody	Phosphatidylethanol (Peth)
Hepatitis B virus (HBV) testing:	<b>Urine Chemistry</b>
Hepatitis B surface antigen (HbsAg)	Drug screen
Hepatitis B surface antibody (anti-HBs)	Ethyl glucuronide (EtG)
Hepatitis B core total antibody (anti-HBc)	<b>Other Serology</b>
Hepatitis B core IgM antibody	Anti-nuclear antibody (ANA)

Hepatitis B core IgG antibody	Anti-smooth muscle antibody (ASMA) <sup>a</sup>
HBV DNA <sup>b</sup>	Anti-actin antibody <sup>c</sup>
Hepatitis C virus (HCV) testing:	Epstein-Barr virus (EBV) testing:
HCV antibody	EBV antibody
HCV RNA <sup>b</sup>	EBV DNA <sup>b</sup>
Hepatitis D virus (HDV) testing:	Cytomegalovirus (CMV) testing:
HDV antibody	CMV antibody
Hepatitis E virus (HEV) testing:	CMV DNA <sup>b</sup>
HEV IgG antibody	Herpes simplex virus (HSV) testing:
HEV IgM antibody	HSV (Type 1 and 2) antibody
HEV RNA <sup>b</sup>	HSV (Type 1 and 2) DNA <sup>b</sup>
<b>Microbiology <sup>d</sup></b>	Liver kidney microsomal type 1 (LKM-1) antibody
Culture:	
Blood	
Urine	

<sup>a</sup> Not required if anti-actin antibody is tested.

<sup>b</sup> Reflex/confirmation dependent on regulatory requirements, testing availability, or both.

<sup>c</sup> Not required if anti-smooth muscle antibody (ASMA) is tested.

<sup>d</sup> Assayed ONLY by investigator-designated local laboratory; no central testing available.



### 10.6.2. Close Hepatic Monitoring

Laboratory tests (Section 10.2, Appendix 2), including ALT, AST, ALP, TBL, direct bilirubin, gamma-glutamyl transferase, and creatine kinase, should be repeated within 48 to 72 hours to confirm the abnormality and to determine if it is increasing or decreasing, if one or more of these conditions occur:

If a participant with baseline results of...	develops the following elevations:
ALT or AST <1.5x ULN	ALT or AST ≥3x ULN
ALP <1.5x ULN	ALP ≥2x ULN
TBL <1.5x ULN	TBL ≥2x ULN (except for participants with Gilbert's syndrome)
ALT or AST ≥1.5x ULN	ALT or AST ≥2x baseline
ALP ≥1.5x ULN	ALP ≥2x baseline
TBL ≥1.5x ULN	TBL ≥1.5x baseline (except for participants with Gilbert's syndrome)

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; TBL = total bilirubin level; ULN = upper limit of normal.

If the abnormality persists or worsens, clinical and laboratory monitoring, and evaluation for possible causes of abnormal liver tests should be initiated by the investigator in consultation with the Lilly-designated medical monitor. At a minimum, this evaluation should include physical examination and a thorough medical history, including symptoms, recent illnesses (for example, heart failure, systemic infection, hypotension, or seizures), recent travel, history of concomitant medications (including over-the-counter), herbal and dietary supplements, history of alcohol drinking and other substance abuse.

Initially, monitoring of symptoms and hepatic biochemical tests should be done at a frequency of 1 to 3 times weekly, based on the participant's clinical condition and hepatic biochemical tests. Subsequently, the frequency of monitoring may be lowered to once every 1 to 2 weeks, if the participant's clinical condition and lab results stabilize. Monitoring of ALT, AST, ALP, and total bilirubin level should continue until levels normalize or return to approximate baseline levels.

### 10.6.3. Comprehensive Hepatic Evaluation

A comprehensive evaluation should be performed to search for possible causes of liver injury if one or more of these conditions occur:

If a participant with baseline results of..	develops the following elevations:
ALT or AST <1.5x ULN	ALT or AST $\geq 3x$ ULN with hepatic signs/symptoms <sup>a</sup> , or ALT or AST $\geq 5x$ ULN
ALP <1.5x ULN	ALP $\geq 3x$ ULN
TBL <1.5x ULN	TBL $\geq 2x$ ULN (except for participants with Gilbert's syndrome)
ALT or AST $\geq 1.5x$ ULN	ALT or AST $\geq 2x$ baseline with hepatic signs/symptoms <sup>a</sup> , or ALT or AST $\geq 3x$ baseline
ALP $\geq 1.5x$ ULN	ALP $\geq 2x$ baseline
TBL $\geq 1.5x$ ULN	TBL $\geq 2x$ baseline (except for participants with Gilbert's syndrome)

Abbreviations: ALP = alkaline phosphatase; ALT= alanine aminotransaminase; AST = aspartate aminotransferase; TBL = total bilirubin level; ULN = upper limit of normal.

a Hepatic signs/symptoms are severe fatigue, nausea, vomiting, right upper quadrant abdominal pain, fever, rash, and/or eosinophilia >5%.

At a minimum, this evaluation should include physical examination and a thorough medical history, as outlined above, as well as tests for PT-INR; tests for viral hepatitis A, B, C, or E; tests for autoimmune hepatitis; and an abdominal imaging study (for example, ultrasound or CT scan).

Based on the participant's history and initial results, further testing should be considered in consultation with the Lilly-designated medical monitor, including tests for hepatitis D virus, cytomegalovirus, Epstein-Barr virus, acetaminophen levels, acetaminophen protein adducts, urine toxicology screen, Wilson's disease, blood alcohol levels, urinary ethyl glucuronide, and blood phosphatidylethanol. Based on the circumstances and the investigator's assessment of the participant's clinical condition, the investigator should consider referring the participant for a hepatologist or gastroenterologist consultation, magnetic resonance cholangiopancreatography, endoscopic retrograde cholangiopancreatography, cardiac echocardiogram, or a liver biopsy.

#### **Additional hepatic data collection (hepatic safety CRF) in study participants who have abnormal liver tests during the study**

Additional hepatic safety data collection in hepatic safety CRFs should be performed in study participants who meet 1 or more of the following 5 conditions:

1. Elevation of serum ALT to  $\geq 5x$  ULN on 2 or more consecutive blood tests (if baseline ALT <1.5x ULN)
  - In participants with baseline ALT  $\geq 1.5x$  ULN, the threshold is ALT  $\geq 3x$  baseline on 2 or more consecutive tests
2. Elevated TBL to  $\geq 2x$  ULN (if baseline TBL <1.5x ULN) (except for cases of known Gilbert's syndrome)
  - In participants with baseline TBL  $\geq 1.5x$  ULN, the threshold should be TBL  $\geq 2x$  baseline
3. Elevation of serum ALP to  $\geq 2x$  ULN on 2 or more consecutive blood tests (if baseline ALP <1.5x ULN)
  - In participants with baseline ALP  $\geq 1.5x$  ULN, the threshold is ALP  $\geq 2x$  baseline on 2 or more consecutive blood tests
4. Hepatic event considered to be an SAE
5. Discontinuation of study intervention due to a hepatic event

**Note:** the interval between the 2 consecutive blood tests should be at least 2 days.

**10.7. Appendix 7: Medical Device Adverse Events (AEs), Adverse Device Effects (ADEs), Serious Adverse Events (SAEs) and Device Deficiencies: Definition and Procedures for Recording, Evaluating, Follow-up, and Reporting**

Refer to Appendix 3 for definitions and procedures for recording, evaluating, follow-up, and reporting of all events.

## **10.8. Appendix 8: Protocol GPIF Standardized Protocols for the Measurement of Height, Weight, Neck Circumference, Waist Circumference, Vital Signs, and Electrocardiogram**

The following information has been adapted from standardized physical measurement protocols for the World Health Organization's STEPwise approach to Surveillance (STEPS) (WHO 2017).

### **10.8.1. Measuring Height**

**Step 1.** Ask the participant to remove their footwear and any headgear (light headgear worn for religious reasons can remain, but this should be worn by the participant at every clinic visit when their height is measured).

**Step 2.** Ask the participant to stand on the calibrated height measuring board (stadiometer) or against a wall with their feet together and their knees straight with their heels against the backboard, the stadiometer, or the wall.

**Step 3.** Ask the participant to look straight ahead without tilting their head up.

**Step 4.** Ask the participant to breathe in and stand tall. Measure and record the participant's height in centimeters to 1 decimal place.

### **10.8.2. Measuring Weight**

- Body weight measurements should be done in a consistent manner using a calibrated electronic scale capable of measuring weight in kilograms to 1 decimal place.
- All weights for a given participant should be measured using the same scale, whenever possible, at approximately the same time in the morning after evacuation of bladder contents.
- Body weight will be measured in fasting state at all visits except Visit 1. If the participant is not fasting, the participant should be called in for a new visit within the visit window to have the fasting body weight measured.

**Step 1.** Ask the participant to empty their pockets, remove their footwear, outerwear (coat, jacket, etc.), and any headgear (light headgear worn for religious reasons can remain, but this should be worn by the participant at every clinic visit when weight is measured).

**Step 2.** Make sure the scale is placed on a firm, flat, even surface (not on carpet, on a sloping surface, or a rough, uneven surface).

**Step 3.** Ask the participant to step onto the scale with 1 foot on each side of the scale.

**Step 4.** Ask the participant to stand still with arms by sides and then record weight in kilograms to the nearest one-tenth kilogram.

### **10.8.3. Measuring Hip and Waist Circumference**

- Hip circumference measurements should be obtained with the participant in the standing position. The hip circumference should be measured at the maximal circumference of the buttocks.

- Waist circumference should be measured in the horizontal plane and at the midpoint between the lower margin of the last palpable rib and the top of the iliac crest.
- Measurements should be taken at the end of a normal expiration using a nonstretchable measuring tape. The tape should lie flat against the skin without compressing the soft tissue.
- The waist circumference should be measured twice, rounded to the nearest 0.5 cm. The measuring tape should be removed between the 2 measurements. Both measurements will be recorded in the CRF. If the difference between the 2 measurements exceeds 1 cm, this set of measurements should be discarded and the 2 measurements repeated.

**Step 1.** Ask the participant to wear light clothing (if available, patient gowns could also be used).

**Step 2.** Ask the participant to stand with their feet close together, arms at their side, body weight evenly distributed.

**Step 3.** Ask the participant to relax and measure the participant's waist circumference.

#### **10.8.4. Measuring Neck Circumference**

- Participants should look straight ahead during the measurement, with shoulders down (not hunched).
- Measure the neck circumference at a point just below the larynx (Adam's Apple) and perpendicular to the long axis of the neck.
- Do not place the tape measure over the Adam's Apple.
- The tape will be as close to horizontal as anatomically feasible (the tape line in the front of the neck should be at the same height as the tape line in the back of the neck).
- Care should be taken so as not to involve the shoulder/neck muscles (trapezius) in the measurement.
- Round neck measurement up to the nearest half centimeter.

#### **10.8.5. Vital Sign Measurements (Blood Pressure and Heart Rate)**

- Vital sign measurements (BP and heart rate, measured by pulse) should be taken before obtaining an ECG tracing and before collection of blood samples for laboratory testing
- The participant should sit quietly for at least 5 minutes before vital signs measurements are taken
- For each parameter, 3 measurements will be taken using the same arm, preferably the nondominant arm
- The recordings should be taken at least 1 minute apart. Each measurement of sitting pulse and BP needs to be recorded in the CRF
- Blood pressure must be taken with an automated BP instrument
- If BP and pulse measurements are taken separately, pulse should be taken prior to BP.

*Note:* In the event pulse measurement cannot be taken via an automated BP instrument, the preferred location for measurement of pulse is the radial artery.

#### **10.8.6. Electrocardiogram**

- All digital ECGs will be obtained using local ECG machines.
- 12-lead ECGs should be obtained after the participant has rested in a supine position for at least 5 minutes.
- Electrocardiograms should be collected prior to collection of blood samples for laboratory testing, including PK samples.

## **10.9. Appendix 9: Provisions for Changes in Study Conduct Due to the COVID-19 Pandemic**

### **Implementation of this appendix**

The changes to procedures described in this appendix are temporary measures intended to be used only during specific time periods as directed by the sponsor in partnership with the investigator.

### **Study disruptions due to the COVID-19 pandemic**

Individual, site, or regional restrictions due to the COVID-19 pandemic may cause disruptions to the conduct of the study. These disruptions may limit the ability of the investigators or participants, or both, to attend on-site visits or to conduct planned study procedures.

### **Implementing changes due to the COVID-19 pandemic**

After receiving the sponsor's written approval, sites may implement changes if permitted by local regulations.

After approval by local ERBs/IRBs, regulatory bodies and any other relevant local authorities, implementation of these changes will not typically require additional notification to these groups, unless they have specific requirements in which notification is required (for example, upon implementation and suspension of changes). All approvals and notifications must be retained in the study records.

If the sponsor grants written approval for changes in study conduct, the sponsor will also provide additional written guidance, if needed.

### **Considerations for making a change**

The prevailing consideration for making a change is ensuring the safety of study participants. Additional important considerations for making a change are compliance with GCP, enabling participants to continue safely in the study and maintaining the integrity of the study.

### **Informed consent**

Additional consent from the participant will be obtained, if required, for:

- participation in remote visits, as defined in Section "Remote Visits"
- dispensation of additional study intervention during an extended treatment period
- alternate delivery of study intervention and ancillary supplies, and
- provision of their personal or medical information required prior to implementation of these activities.

### **Changes in study conduct due to the COVID-19 pandemic**

Changes in study conduct not described in this appendix, or not consistent with applicable local regulations, are not allowed.

The following changes in study conduct will not be considered protocol deviations.

***Remote visits****Types of remote visits***Telephone/Telemedicine**

Telephone or technology-assisted virtual visits, or both, are acceptable to complete appropriate assessments. Assessments to be completed in this manner include, but are not limited to, collection of AEs and PCs, concomitant medications review, review study participant diary (including study intervention compliance), review diet and exercise goals, C-SSRS (Since Last Visit Assessed) and PHQ-9. PROs will be completed by the participant on the provisioned device per the SoA.

**Mobile healthcare visit**

Healthcare visits may be performed by a mobile healthcare provider at locations other than the study site when participants cannot travel to the site due to the COVID pandemic, if written approval is provided by the sponsor. Procedures performed at such visits include, but are not limited to, weight measurement, blood sample collection, vital signs (temperature, PR, BP), concomitant medication review, conducting physical assessments, collection of AEs and PCs, and collecting health information. PROs will be completed by the participant on the provisioned device per the SoA.

Every effort should be made for the participant to return to on-site visits as soon as reasonably possible, while ensuring the safety of the participant and investigational site staff.

*Data capture*

In source documents and the CRF, the study site should capture the visit method, with a specific explanation for any data missing because of missed in-person site visits.

*Safety reporting*

Regardless of the type of remote visits implemented, the protocol requirements regarding the reporting of AEs, SAEs, and PCs remain unchanged.

*Return to on-site visits*

Every effort should be made to enable participants to return to on-site visits as soon as reasonably possible, while ensuring the safety of both the participants and the site staff.

***Local laboratory testing option***

Local laboratory testing may be conducted in lieu of central laboratory testing. However, central laboratory testing must be retained for PK, immunogenicity, hsCRP, insulin and lipid samples. The local laboratory must be qualified in accordance with applicable local regulations.

***Study intervention and supplies***

When a participant is unable to go to the site to receive study supplies during normal on-site visits, the site should work with the sponsor to determine appropriate actions. These actions may include:

- asking the participant to go to the site and receive study supplies from site staff without completion of a full study visit



- asking the participant's designee to go to the site and receive study supplies on a participant's behalf, and
- arranging delivery of study supplies.

These requirements must be met before action is taken:

- Alternate delivery of study intervention should be performed in a manner that does not compromise treatment blinding and ensures product integrity. The existing protocol requirements for product accountability remain unchanged, including verification of participant's receipt of study supplies.
- When delivering supplies to a location other than the study site (for example, participant's home), the investigator, sponsor, or both should ensure oversight of the shipping process to ensure accountability and product quality (that is, storage conditions maintained and intact packaging upon receipt).
- Instructions may be provided to the participant or designee on the final disposition of any unused or completed study supplies.

### ***Screening period guidance***

To ensure safety of study participants, laboratory values and other eligibility assessments taken at Visit 1 are valid for a maximum of 90 days. The following rules will be applied for active participants not randomly assigned to study intervention whose participation in the study must be paused due to the COVID-19 pandemic:

- If screening is paused for less than 90 days from Visit 1 to Visit 2: the participant will proceed to the next study visit per the usual SoA, provided that Visit 2 must be conducted within 90 days from Visit 1.
  - The site should conduct the next visit if the participant's eligibility criteria are confirmed, and the site should document the reason for delay.
  - Due to the pause in screening, sites should also reconfirm the impacted participant's consent and document this confirmation in the source documentation.
- If screening is paused for more than 90 days from Visit 1 to Visit 2: the participant must be discontinued because of screening interruption due to the COVID-19 pandemic. This is documented as a screen failure in the CRF. The participant can reconsent and be rescreened as a new participant. This rescreen is in addition to the one allowed by the main protocol. The screening procedures per the usual SoA should be followed, starting at Visit 1 to ensure participant eligibility by Visit 2.

### ***Adjustments to visit windows***

Whenever possible and safe to do so, as determined by the investigator's discretion, participants should complete the usual SoA. To maximize the possibility that these visits can be conducted as on-site visits, the windows for visits may be adjusted, upon further guidance from the sponsor. This minimizes missing data and preserves the intended conduct of the study.

This table describes the allowed adjustments to visit windows.

<b>Visit Number</b>	<b>Tolerance</b>
Visit 1 (Screening)	No change
Visit 2 (Randomization)	Within 90 days after Visit 1.
Visits 3 through 6	Within 7 days before or after the intended date.
Visits 7 through 10	Within 14 days before or after the intended date.
Visit 11	Within 14 days before the intended date, or up to 28 days after the intended date.
Visit 801	Up to 28 days after the intended date

For participants whose visits have extended windows, additional study intervention may need to be provided to avoid interruption and maintain overall integrity of the study.

### **Documentation**

*Changes to study conduct will be documented*

Sites will identify and document the details of how participants, visits types, and conducted activities were affected due to the COVID-19 pandemic. Dispensing/shipment records of study intervention and relevant communications, including delegation, should be filed with site study records.

*Source documents at alternate locations*

Source documents generated at a location other than the study site should be part of the investigator's source documentation and should be transferred to the site in a secure and timely manner.

## 10.10. Appendix 10: Abbreviations and Definitions

Term	Definition
<b>AASM</b>	American Academy of Sleep Medicine
<b>Abuse</b>	Use of a study intervention for recreational purposes or to maintain an addiction or dependence
<b>ADA</b>	antidrug antibody
<b>AE</b>	adverse event
<b>AHI</b>	Apnea-Hypopnea Index
<b>ALP</b>	alkaline phosphatase
<b>ALT</b>	alanine aminotransferase
<b>ANCOVA</b>	analysis of covariance
<b>Apnea</b>	decrease in airflow $\geq 90\%$ from baseline for $\geq 10$ seconds
<b>AST</b>	aspartate aminotransferase
<b>AX6</b>	Axivity 6
<b>blinding/masking</b>	A single-blind study is one in which the investigator and/or the investigator's staff are aware of the treatment but the participant is not, or vice versa, or when the sponsor is aware of the treatment but the investigator and/the investigator's staff and the participant are not. A double-blind study is one in which neither the participant nor any of the investigator or sponsor staff who are involved in the treatment or clinical evaluation of the participants are aware of the treatment received.
<b>BMI</b>	body mass index
<b>BP</b>	blood pressure
<b>CBD</b>	cannabidiol
<b>COA</b>	clinical outcome assessment
<b>complaint</b>	A complaint is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, purity, durability, reliability, safety or effectiveness, or performance of a drug or drug delivery system.
<b>compliance</b>	Adherence to all study-related, good clinical practice (GCP), and applicable regulatory requirements.
<b>CRF</b>	case report form; a printed, optical, or electronic document designed to record all of the protocol-required information to be reported to the sponsor for each trial participant.
<b>CSR</b>	clinical study report
<b>C-SSRS</b>	Columbia-Suicide Severity Rating Scale
<b>CV</b>	cardiovascular
<b>Device deficiencies</b>	Equivalent to product complaint
<b>DMC</b>	data monitoring committee. A data monitoring committee, or data monitoring board (DMB) is a group of independent scientists who are appointed to monitor the safety and scientific integrity of a human research intervention, and to make recommendations to the sponsor regarding the stopping of a study for efficacy, or for harms, or for futility. The composition of the committee is dependent upon the scientific skills and knowledge required for monitoring the particular study.
<b>DPP-4</b>	dipeptidyl-peptidase-4
<b>EAS</b>	efficacy analysis set
<b>ECG</b>	electrocardiogram
<b>ED</b>	early discontinuation
<b>EDC</b>	electronic data capture
<b>eGFR</b>	estimated glomerular filtration rate
<b>enroll</b>	The act of assigning a participant to a treatment. Participants who are enrolled in the study are those who have been assigned to a treatment.

<b>enter</b>	Participants entered into a study are those who sign the informed consent form directly or through their legally acceptable representatives.
<b>ERB</b>	ethical review board
<b>EQ-5D-5L</b>	EuroQol - 5 Dimension - 5 Level
<b>ESS</b>	Epworth Sleepiness Scale
<b>FAS</b>	full analysis set
<b>FG</b>	fasting glucose
<b>FOSQ</b>	Functional Outcomes of Sleep Questionnaire
<b>GCP</b>	good clinical practice
<b>GI</b>	gastrointestinal
<b>GIP</b>	glucose-dependent insulinotropic polypeptide
<b>GLP-1</b>	glucagon-like peptide-1
<b>GLP-1R</b>	glucagon-like peptide-1 receptor
<b>GLP-1RA</b>	glucagon like peptide 1 receptor
<b>HbA1c</b>	hemoglobin A1c
<b>HDL</b>	high-density lipoprotein
<b>HSAT</b>	home sleep apnea test
<b>hsCRP</b>	high-sensitivity C-reactive protein
<b>Hypopnea</b>	an abnormal respiratory event lasting $\geq 10$ seconds with $\geq 30\%$ reduction in thoracoabdominal movement or airflow as compared to baseline, and with $\geq 4\%$ oxygen desaturation.
<b>IB</b>	Investigator's Brochure
<b>ICF</b>	informed consent form
<b>ICH</b>	International Council for Harmonisation
<b>IEC</b>	Independent Ethics Committees
<b>IMP</b>	Investigational Medicinal Product (see also "investigational product") A medicinal product which is being tested or used as a reference, including as a placebo, in a clinical trial.
<b>informed consent</b>	A process by which a participant voluntarily confirms their willingness to participate in a particular study, after having been informed of all aspects of the study that are relevant to the participant's decision to participate. Informed consent is documented by means of a written, signed and dated informed consent form.
<b>INR</b>	International normalized ratio
<b>interim analysis</b>	An interim analysis is an analysis of clinical study data, separated into treatment groups, that is conducted before the final reporting database is created/locked.
<b>investigational product</b>	A pharmaceutical form of an active ingredient or placebo being tested or used as a reference in a clinical trial, including products already on the market when used or assembled (formulated or packaged) in a way different from the authorized form, or marketed products used for an unauthorized indication, or marketed products used to gain further information about the authorized form.
<b>IRB</b>	institutional review board
<b>ISA</b>	intervention-specific appendix
<b>ISR</b>	injection site reaction
<b>ITT</b>	intention to treat: The principle that asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a participant (that is, the planned treatment regimen) rather than the actual treatment given. It has the consequence that participant allocated to a treatment group should be followed up, assessed, and analyzed as members of that group irrespective of their compliance to the planned course of treatment.
<b>IWRS</b>	interactive web-response system
<b>MACE</b>	major adverse cardiovascular events

<b>medication error</b>	Errors in the prescribing, dispensing, or administration of a study intervention, regardless of whether or not the medication is administered to the participant or the error leads to an AE. Medication error generally involve a failure to uphold one or more of the five “rights” of medication use: the right participant, the right drug, the right dose, right route, at the right time. In addition to the core five rights, the following may also represent medication errors: <ul style="list-style-type: none"> <li>• dose omission associated with an AE or a product complaint</li> <li>• dispensing or use of expired medication</li> <li>• use of medication past the recommended in-use date</li> <li>• dispensing or use of an improperly stored medication</li> <li>• use of an adulterated dosage form or administration technique inconsistent with the medication's labeling (for example, Summary of Product Characteristics, IB, local label, protocol), or</li> <li>• shared use of cartridges, prefilled pens, or both.</li> </ul>
<b>MEN</b>	multiple endocrine neoplasia
<b>MI</b>	myocardial infarction
<b>misuse</b>	Use of a study intervention for self-treatment that either is inconsistent with the prescribed dosing regimen, indication, or both, or is obtained without a prescription.
<b>MMRM</b>	Mixed model repeated measures
<b>MTC</b>	medullary thyroid cancer
<b>MTD</b>	maximum tolerated dose
<b>OGTT</b>	oral glucose tolerance test
<b>OSA</b>	obstructive sleep apnea
<b>OTC</b>	Over the counter
<b>PAP</b>	positive airway pressure
<b>participant</b>	Equivalent to CDISC term “subject”: an individual who participates in a clinical trial, either as recipient of an investigational medicinal product or as a control.
<b>PC</b>	product complaint
<b>PGIC</b>	Patient Global Impression of Change
<b>PGIS</b>	Patient Global Impression of Status
<b>PHQ-9</b>	Patient Health Questionnaire-9
<b>PK/PD</b>	pharmacokinetics/pharmacodynamics
<b>PR</b>	pulse rate
<b>PRO/ePRO</b>	patient-reported outcomes/electronic patient-reported outcomes
<b>PROMIS</b>	Patient-Reported Outcomes Measurement Information System
<b>PSG</b>	polysomnography
<b>QTc</b>	corrected QT interval
<b>QW</b>	weekly
<b>SAE</b>	serious adverse event
<b>SAP</b>	statistical analysis plan
<b>SASHB</b>	sleep apnea specific hypoxic burden
<b>SBP</b>	systolic blood pressure
<b>SC</b>	subcutaneous
<b>SD</b>	standard deviation
<b>screen</b>	The act of determining if an individual meets minimum requirements to become part of a pool of potential candidates for participation in a clinical study.
<b>SF-36v2</b>	Short-Form 36 version 2
<b>SIB</b>	suicidal ideation and behavior
<b>SoA</b>	Schedule of Activities
<b>SS</b>	safety analysis set
<b>T1DM</b>	Type 1 diabetes mellites

<b>T2DM</b>	Type 2 diabetes mellites
<b>TBL</b>	total bilirubin level
<b>TEE</b>	total energy expenditure
<b>TEAE</b>	treatment-emergent adverse event: An untoward medical occurrence that emerges during a defined treatment period, having been absent pretreatment, or worsens relative to the pretreatment state, and does not necessarily have to have a causal relationship with this treatment.
<b>THC</b>	Tetrahydrocannabinol
<b>TSH</b>	thyroid-stimulating hormone
<b>UACR</b>	urinary albumin/creatinine ratio
<b>ULN</b>	upper limit of normal
<b>WOCBP</b>	woman of childbearing potential

---

## 10.11. Appendix 11: Country-Specific Requirements

### 10.11.1. Germany

This section describes protocol changes applicable for adult participants in study sites in Germany.

This table describes the changes and provides a rationale for the changes.

Protocol Section Number and Name	Description of the Change	Brief Rationale
7.2. Participant Discontinuation/Withdrawal from the Study		
8.3. Adverse Events, Serious Adverse Events, and Product Complaints	Deleted references to “legally authorized representative,” “legal guardian,” “parents”	The German Drug Law (Arzneimittelgesetz – AMG) requires per Paragraph 40 (1-3) and Paragraph 41 (3) that adult participants act on their own behalf and provide their own written informed consent. If written consent is not possible, verbal consent with a witness is acceptable. No legal representative consent is accepted.
10.1.3 Informed Consent Process		
10.10. Abbreviations and Definitions		

The revised text in the following subsections show the changes applicable for adult participants to study sites in Germany. Additions are identified by underline. Deletions are identified by ~~strikethrough format~~.

### 7.2. Participant Discontinuation/Withdrawal from the Study

Discontinuation is expected to be uncommon.

A participant may withdraw from the study:

- at any time at the participant’s own request
- ~~at the request of the participant’s designee (for example, parents or legal guardian)~~
- at the discretion of the investigator for safety, behavioral, compliance, or administrative reasons
  - If the participant becomes pregnant during the study (see Section 8.3.2 for additional details)
- if enrolled in any other clinical study involving an investigational product, or enrolled in any other type of medical research judged not to be scientifically or medically compatible with this study
- if the participant, for any reason, requires treatment with a therapeutic agent that is prohibited by the protocol and has been demonstrated to be effective for treatment of the study indication. In this case, discontinuation from the study occurs prior to introduction of the new agent.

**8.3. Adverse Events, Serious Adverse Events, and Product Complaints**

The definitions of the following events can be found in Section 10.3, Appendix 3:

- AEs
- SAEs
- PCs

These events will be reported by the participant (or, when appropriate, by a caregiver, or surrogate, or the participant's legally authorized representative).



### 10.1.3 Informed Consent Process

The investigator or the investigator's representative will explain the nature of the study, including the risks and benefits, to the participant ~~or the participant's legally authorized representative~~ and answer all questions regarding the study.

Participants must be informed that their participation is voluntary. Participants ~~or their legally authorized representatives~~ will be required to sign a statement of informed consent that meets the requirements of 21 Code of Federal Regulations 50, local regulations, ICH guidelines, privacy and data protection requirements, where applicable, and the IRB/IEC or study center.

The medical record must include a statement that written informed consent was obtained before the participant was entered in the study and the date the written consent was obtained. The authorized person obtaining the informed consent must also sign the ICF.

Participants must be reconsented to the most current version of the ICF(s) during their participation in the study.

A copy of the ICF(s) must be provided to the participant ~~or the participant's legally authorized representative~~ and is kept on file.

Participants who are rescreened are required to sign a new ICF.

**10.10. Appendix 10: Abbreviations and Definitions**

**enter**

Participants entered into a study are those who sign the informed consent form directly ~~or through their legally acceptable representatives.~~

### 10.12. Appendix 12: Protocol Amendment History

The Protocol Amendment Summary of Changes Table for the current amendment is located directly before the Table of Contents (TOC).

#### Amendment [b]: 30-Sep-2022

This amendment is considered to be substantial.

The amendment is considered to be substantial because it is likely to have a significant impact on the

- safety or the rights of the study participants, and
- reliability and robustness of the data generated in the clinical study.

#### Overall Rationale for the Amendment:

This amendment includes the addition of a key secondary objective, “Change in AHI”. The endpoint for key secondary objective “Change in SBP” and other secondary objective “Change in DBP” has been updated to Week 48 from Week 52. Changes have been made to the Schedule of Activities to reflect the synchronization of timing of C-SSRS to PHQ-9.

Protocol changes have been made as outlined in the following table.

Section # and Name	Description of Change	Brief Rationale
1.1. Synopsis and 3. Objectives, Endpoints, and Estimands	Key Secondary Objectives: Added “Change in AHI” as a key secondary objective with endpoint “Change in AHI”	To align with recommendations from health authority.
	Key Secondary Objectives: Moved “change in SBP” to “From baseline to Week 48” section	BP will be assessed at Week 48 because PAP suspension at Week 52 may confound BP assessment.
	Added footnote a “BP will be assessed at Week 48 because PAP suspension at Week 52 may confound BP assessment”.	For additional information.
1.3. Schedule of Activities	Removed text, “Note: Visit 10 is eligible to be conducted remotely (that is, by telephone, IT-assisted virtual visit, or in combination with on-site visit) at the direction of the sponsor, according to local laws and regulations.”	Vital signs will be measured at Week 48, so this visit will be on site.

Section # and Name	Description of Change	Brief Rationale
	“Fasting Visit” row: Removed text, “If V10 is remote, then the visit will not be fasting.”	This instruction is no longer applicable.
	“Weight” row: Removed text, “If V10 is remote, then weight will not be obtained at the visit.”	This instruction is no longer applicable.
	“Vital signs” row: Removed text, “If V10 is remote, vital signs will not be obtained at the visit.”	This instruction is no longer applicable.
	“Symptom-directed physical assessment” row: Removed text, “If V10 is remote, then symptom-directed physical assessment will not be done at the visit.”	This instruction is no longer applicable.
	Removed “Schedule Sleep Center Study for PSG” and “Sleep Center Study for PSG” from ED visit.	Sleep study is not required for participants who discontinue early.
	“Review Lifestyle Program instructions” row: Updated comment to “Diet and exercise goals established during the lifestyle consultation and the importance of adherence to the lifestyle component of the trial will be reinforced at each trial contact by study staff.”	To clarify the requirement.
	“ESS, EQ-5D-5L, FOSQ, and PGIS” rows: Updated comment “When the PROs are scheduled for visits at which the PSG will be done, they should be completed <u>on the same day as PSG and</u> in the following order (FOSQ, ESS, PROMIS Short Form v1.0 Sleep Disturbance 8b, PROMIS Short Form v1.0 Sleep-related Impairment 8a, PGIS, PGIC, SF-36v2 acute form, and EQ-5D-5L) <del>before the</del>	To clarify the requirement.

Section # and Name	Description of Change	Brief Rationale
	<del>PSG is conducted, and should be done at the same time of day for each of these visits. Completing the PRO assessments on the next day after PSG is not considered a protocol violation.</del>	
	“PHQ-9”, “C-SSRS screening/baseline”, and “C-SSRS (since last visit version)” rows: added “if both collected on the same day” to comments.	To clarify timing and allow further flexibility for AE collection independent of PHQ-9 and C-SSRS assessments.
	“C-SSRS (since last visit version)” row: Removed assessment from Visits 3, 4, 5, 6, 8, 9, 10, and 801.	Collection schedule has been updated to align with PHQ-9 assessments.
	“Participant returns all unused study intervention” row: Removed text, “If V10 is remote, participant will return any unused study intervention dispensed at V9 during the V11 visit.”	This instruction is no longer applicable.
	“Assess study intervention compliance” row: Removed text, “If V10 is remote, the next study intervention compliance will be done at V11.”	This instruction is no longer applicable.
3. Objectives, Endpoints, and Estimands	Other Secondary Objectives: Moved “change in DBP” to “From baseline to Week 48” section	BP will be assessed at Week 48 because PAP suspension at Week 52 may confound BP assessment.
5.2. Exclusion Criteria	Exclusion Criterion 42: Updated formatting.	For clarification.
	Exclusion Criterion 53: Removed “(4 months for studies conducted in Japan, 3 months for studies conducted in the United Kingdom).”	This information is no longer applicable.
6.5. Dose Modification	Modified text, “Dose modification is permitted for	Clarification.

Section # and Name	Description of Change	Brief Rationale
	management of intolerable GI symptoms <u>during the first 24 weeks of the treatment period</u> (Section 6.5.1).	
	Modified text, “Participants who do not tolerate at least 10 mg even after <u>the described measures, including 1 de-escalation and re-escalation attempt</u> , will be discontinued from the study intervention but remain in the study for continued follow-up.”	Clarification.
6.5.1. Management of Participants with Gastrointestinal Symptoms	Modified text, “For participants unable to tolerate any dose <del>escalation</del> between 7.5 mg and 15 mg inclusive, <u>despite the above measures, the investigator should contact Lilly to consider a dose de-escalation step with subsequent re-escalation by 2.5 mg every 4 weeks up to MTD will be allowed</u> in a blinded fashion, to reach either the 10-mg or 15-mg dose as described below.”	Direct the investigator to contact sponsor to ensure that dose modifications are completed according to protocol.
	Added text “Dose modifications after the first 24 weeks of the treatment period are not permitted. For temporary study treatment discontinuation see Section 7.1.2.”	Clarification.
6.5.1. Management of Participants with Gastrointestinal Symptoms	Modified text, “Participants who tolerate <ul style="list-style-type: none"> <li>• 10 mg, but do not tolerate 12.5 mg or 15 mg even following <u>the above measures, including 1 de-escalation and re-escalation attempt</u>, will continue on 10 mg as their MTD dose.</li> </ul>	Clarification.

Section # and Name	Description of Change	Brief Rationale
	<ul style="list-style-type: none"> <li>12.5 mg, but do not tolerate 15 mg even after <u>the above measures, including</u> 1 de-escalation and re-escalation attempt, will continue on 10 mg as their MTD dose.”</li> </ul>	
7.1. Discontinuation of Study Intervention	Removed text in clinical considerations section, “inadvertent enrollment if continued treatment with study intervention would not be medically appropriate”	Participants who do not meet enrollment criteria will discontinue from the study.
	Added text, “The participant should be referred to a MHP for further evaluation and care.”	To provide further detail for participant care.
7.2. Participant Discontinuation/Withdrawal from the Study	Added text, “A participant will be withdrawn from the study in case of inadvertent enrollment”	Text updated to align with relevant edits in Section 7.1.
8.2.3. Electrocardiograms (ECG)	<p>Modified instructions for timing of ECG collection.</p> <p>“ECGs should be collected <del>at least 30 minutes</del> prior to collection of blood samples for laboratory testing, including PK samples. Participants should be supine for <del>at least approximately</del> 5 to 10 minutes before ECG collections and remain supine but awake during the ECG collection. ECGs may be repeated at the investigator’s discretion at any visit.”</p>	Clarification.
8.3.1.1. Adverse Event Monitoring with a Systematic Questionnaire	Added text, “if AE and C-SSRS/PHQ-9 collections done on the same day”.	To clarify timing and allow further flexibility for AE collection independent of PHQ-9 and C-SSRS assessments.
9.2. Analyses Sets	FAS and EAS descriptions updated to include “excluding those discontinuing study due to inadvertent enrollment”	Text revised to reflect updated definitions for FAS and EAS.

Section # and Name	Description of Change	Brief Rationale
	Added “(last dose +7 days)” to the description of data excluded due to the discontinuation of study intervention in the EAS population.	Clarification.
9.3. Statistical Analysis	Modified text for “hybrid” estimand, changed “hybrid” to “treatment regimen” throughout Section 9.3.	To keep terminology consistent with the literature.
9.3.1. General Considerations	Modified text, “Efficacy analyses will be conducted on all participants randomly assigned to study intervention according to the treatment to which the participants are assigned <u>and were exposed to at least one dose.</u> ”	To add details for clarification.
	Modified text, “For the “efficacy” estimand, the analysis will <del>include data collected prior to permanent discontinuation of study intervention and</del> will be conducted using the EAS.”	This was duplicate information.
	“Missing Value Imputation” information moved and modified.	To reflect that the endpoint for SBP is at Week 48 instead of Week 52 and to clarify that no explicit imputation will be performed for efficacy analysis relative to “efficacy” estimand.
	Moved description of handling of missing data for “efficacy” estimand earlier in the paragraph.	Clarification.
9.3.2. Primary Analysis	Deleted text, “There will be 2 primary analysis methods, each tested at the full significance level of 0.05”	This information has been moved to the last sentences of the primary analysis section and key secondary analysis section for clarification.



Section # and Name	Description of Change	Brief Rationale
	Modified text, “Missing values will be imputed <del>following imputation methods based on the strategy to handle intercurrent events</del> described in Section 9.3.1”	Clarification.
	Added text, “Analysis aligned to each estimand will be evaluated at the full significance level of 0.05.”	To substitute the deleted text “There will be 2 primary analysis methods, each tested at the full significance level of 0.05”.
9.3.3. Analysis of Key Secondary Endpoints	Updated text, “Analysis of <u>change in AHI</u> , percent change from baseline in body weight, <del>change in SBP</del> , and CRP at the 52-week visit <u>and change in SBP at the 48-week visit</u> will be conducted in a manner similar to the primary efficacy analyses with baseline AHI stratum added in the model, and baseline of the corresponding variable as a covariate.	To align with changes in Section 3.
	Added text, “Analysis of change in AHI will not include the baseline AHI stratum.”	Clarification.
	Added text, “Analysis aligned to each estimand will be evaluated at the full significance level of 0.05 contingent on reaching statistical significance of the primary objective.”	To substitute the deleted text “There will be 2 primary analysis methods, each tested at the full significance level of 0.05”.
10.3.7.1. Hypoglycemia	Updated text, “Hypoglycemic episodes will be recorded <u>in the eDiary</u> <del>on a specific CRF</del> and should not be recorded as AEs unless the event meets serious criteria.	Correction.
10.3.7.3. Injection-Site Reactions	Modified text, “At the time of AE occurrence <del>in the tirzepatide group</del> , samples will be collected	Correction.

Section # and Name	Description of Change	Brief Rationale
	for measurement of tirzepatide ADAs and tirzepatide concentration.”	
10.8.5. Vital Sign Measurements (Blood Pressure and Heart Rate)	Modified text, “ <ul style="list-style-type: none"> <li>• The participant should sit quietly for <u>at least 5</u> minutes before vital signs measurements are taken”</li> </ul>	Clarification.
10.8.6. Electrocardiogram	Modified text, “ <ul style="list-style-type: none"> <li>• 12-lead ECGs should be obtained after the participant has rested in a supine position for at least <u>5</u> <del>10</del> minutes.</li> <li>• Electrocardiograms should be collected <del>at least 30 minutes</del> prior to collection of blood samples for laboratory testing, including PK samples.</li> <li>• <del>Electrocardiograms should be obtained approximately 1 minute apart, with all 3 tracings to be obtained within approximately 5 minutes. Measurements that deviate substantially from previous readings should be repeated immediately.</del>”</li> </ul>	Clarification. One tracing of ECG is considered adequate.
Throughout	Editorial corrections.	Minor, therefore not described.

**Amendment a: 10-Feb-2022****Overall Rationale for the Amendment:**

This amendment corrects terminology used in the protocol for intervention-specific appendix (ISA), and adds selected clarifications.

<b>Section # and Name</b>	<b>Description of Change</b>	<b>Brief Rationale</b>
Title Page 1.1. Synopsis 1.3. Schedule of Activities 10.10. Abbreviations and Definitions	Changed “indication-specific appendix” to “intervention-specific appendix.”	Correction.
1.3. Schedule of Activities	“Symptom-directed physical assessment” row: Clarified marks for symptom-directed physical assessment during the study, and added a clarifying statement in notes.	Clarification. Original language could be interpreted to mean symptom-directed physical assessment is required at all visits. Symptom-directed physical assessment from visits 2-11 and ED will be conducted at the discretion of the PI, as indicated based on participant status and standard of care.
	“Participant diary dispensed” row: Specified which diaries are electronic versus paper.	Clarification.
	“Review Lifestyle Program instructions” row: comments updated/moved from the “Review diet and exercise goals” row, which has been deleted.	Clarification.
	“Review diet and exercise goals” row: Removed.	Clarification. This review is part of Review Lifestyle Program instructions.
5.2. Exclusion Criteria	Deleted criterion 48.	Duplicate criterion (Criterion 49).

<b>Section # and Name</b>	<b>Description of Change</b>	<b>Brief Rationale</b>
8.1.3.1. Actigraphy	Modified text, “Participants should wear the <u>a</u> Actigraphy device for 7 consecutive days, <u>5 times during the study</u> , per the SoA.”	Clarification.
8.2.1. Physical Examination	Modified text, “A symptom-directed physical assessment will be performed <del>at visits as</del> indicated in the SoA, as clinically indicated.”	Clarification, for consistency with changes in the “Symptom-directed physical assessment” row in the SoA.
10.8.2. Measuring Weight	Modified text, “Body weight will be measured in fasting state <u>at all visits except Visit 1</u> .”	Clarification.
Throughout	Editorial corrections.	Minor, therefore not described.


## 11. References

- [ADA] American Diabetes Association Glycemic targets: Standards of medical care in diabetes-2020. *Diabetes Care*. 2020;43(Suppl 1):S66–S76. <https://doi.org/10.2337/dc20-S006>
- [AHRQ] Agency for Healthcare Research and Quality. Continuous positive airway pressure treatment for obstructive sleep apnea. Accessed July 20, 2021. <https://www.ahrq.gov/sites/default/files/wysiwyg/research/findings/ta/drafts-for-review/sleep-apnea-draftreport.pdf>
- Alosh M, Bretz F, Mohammad H. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693-713. <https://doi.org/10.1002/sim.5974>
- Banks PA, Freeman ML; Practice Parameters Committee of the American College of Gastroenterology. Practice guidelines in acute pancreatitis. *Am J Gastroenterol*. 2006;101(10):2379-2400.
- Blackman A, Foster GD, Zammit G, et al. Effect of liraglutide 3.0 mg in individuals with obesity and moderate or severe obstructive sleep apnea: the SCALE Sleep Apnea randomized clinical trial. *Int J Obes (Lond)*. 2016;40(8):1310-1319.
- Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913. <https://doi.org/10.1002/bimj.201000239>
- Currie AC, Kaur V, Carey I, et al. Obstructive sleep apnea remission following bariatric surgery: a national registry cohort study. *Surgery Obes Relat Dis*. 2021;17(9):1576-1582. <https://doi.org/10.1016/j.soard.2021.05.021>
- Dahl D, Onishi Y, Norwood P, et al. Tirzepatide, a dual GIP/GLP-1 receptor agonist, is effective and safe when added to basal insulin for treatment of type 2 diabetes (SURPASS-5). *Diabetes*. 2021;70(suppl\_1). <https://doi.org/10.2337/db21-80-LB>
- Danne T, Philotheou A, Goldman D, et al. A randomized trial comparing the rate of hypoglycemia – assessed using continuous glucose monitoring – in 125 preschool children with type 1 diabetes treated with insulin glargine or NPH insulin (the PRESCHOOL study). *Pediatr Diabetes*. 2013;14(8):593-601. <https://doi.org/10.1111/pedi.12051>
- EuroQol Research Foundation. EQ-5D-5L user guide, version 3.0. Updated September 2019. <https://euroqol.org/publications/user-guides>
- FAO/WHO/UNU. *Human energy requirements. Report of a joint FAO/WHO/UNU Expert Consultation*. Rome. 2004.
- [FDA] Food and Drug Administration. Guidance for Industry. Drug-Induced Liver Injury: Premarketing Clinical Evaluation. July 2009. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/drug-induced-liver-injury-premarketing-clinical-evaluation>
- Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. *Stat Med*. 1999;18(11):1341-1354. [https://doi.org/10.1002/\(sici\)1097-0258\(19990615\)18:11<1341::aid-sim129>3.0.co;2-7](https://doi.org/10.1002/(sici)1097-0258(19990615)18:11<1341::aid-sim129>3.0.co;2-7)

- Frias JP, Davies MJ, Rosenstock J, et al. Tirzepatide versus semaglutide once weekly in patients with type 2 diabetes. *N Engl J Med*. 2021;385(6):503-515.  
<https://doi.org/10.1056/nejmoa2107519>
- Gottlieb DJ, Punjabi NM. Diagnosis and management of obstructive sleep apnea: a review. *JAMA*. 2020;323(14):1389-1400. <https://doi.org/10.1001/jama.2020.3514>
- Hamilton GS, Gupta R, Vizcarra D, et al; Guidelines Committee Members and Governing Council of the World Sleep Society. Endorsement of: “clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American academy of sleep medicine clinical practice guideline” by the World Sleep Society. *Sleep Med*. 2021;79:152-154.  
<https://doi.org/10.1016/j.sleep.2020.12.044>.
- Hudgel DW, Patel SR, Ahasic AM, et al. The role of weight management in the treatment of adult obstructive sleep apnea. An Official American Thoracic Society Clinical Practice Guideline. *Am J Respir Crit Care Med*. 2018;198(6):e70-e87.  
<https://doi.org/10.1164/rccm.201807-1326ST>
- Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540-545. <https://doi.org/10.1093/sleep/14.6.540>
- Koizumi M, Takada T, Kawarada Y, et al. JPN guidelines for the management of acute pancreatitis: diagnostic criteria for acute pancreatitis. *J Hepatobiliary Pancreat Surg*. 2006;13(1):25-32. <https://doi.org/10.1007/s00534-005-1048-2>
- Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613.
- [Lilly]. Eli Lilly and Company. Lilly’s tirzepatide achieves all primary and key secondary study outcomes against insulin glargine in adults with type 2 diabetes and increased cardiovascular risk in SURPASS-4 trial. Published May 20, 2021. Accessed July 15, 2021.  
<http://lilly.mediaroom.com/2021-05-20-Lillys-tirzepatide-achieves-all-primary-and-key-secondary-study-outcomes-against-insulin-glargine-in-adults-with-type-2-diabetes-and-increased-cardiovascular-risk-in-SURPASS-4-trial>
- Ludvik B, Giorgino F, Jodar E, et al. Efficacy and safety of tirzepatide, a dual GIP/GLP-1 receptor agonist, compared with insulin degludec in patients with type 2 diabetes (SURPASS 3). *Diabetes*. 2021;70(suppl\_1). <https://doi.org/10.2337/db21-78-LB>
- Maruish ME, editor. User’s manual for the SF-36v2 Health Survey. 3rd ed. Lincoln, RI: Quality Metric Incorporated; 2011.
- [Northwestern]. Northwestern University. PROMIS Short Form v1.0 – Sleep-Related Impairment 8a. Published May 09, 2016a. Accessed Dec 01, 2021.  
[https://www.healthmeasures.net/index.php?option=com\\_instruments&view=measure&id=185&Itemid=992](https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=185&Itemid=992)
- [Northwestern]. Northwestern University. PROMIS Short Form v1.0 – Sleep Disturbance 8b. Published May 03, 2016b. Accessed Dec 01, 2021.  
[https://www.healthmeasures.net/index.php?option=com\\_instruments&view=measure&id=184&Itemid=992](https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=184&Itemid=992)

- Opperer M, Cozowicz C, Bugada D, et al. Does obstructive sleep apnea influence perioperative outcome? A qualitative systematic review for the society of anesthesia and sleep medicine task force on preoperative preparation of patients with sleep-disordered breathing. *Anesth Analg*. 2016;122(5):1321-1334. <https://doi.org/10.1213/ane.0000000000001178>
- Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. 2012;33(2):176-182. <https://doi.org/10.1093/eurheartj/her352>
- Rosenstock J, Wysham C, Frias JP, et al. Efficacy and safety of a novel dual GIP and GLP-1 receptor agonist tirzepatide in patients with type 2 diabetes (SURPASS-1): a double-blind, randomised, phase 3 trial. *Lancet*. 2021;398(10295):143-155. [https://doi.org/10.1016/s0140-6736\(21\)01324-6](https://doi.org/10.1016/s0140-6736(21)01324-6)
- Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons Inc.; 1987.
- The Columbia Lighthouse Project. The Columbia Protocol for Research. Columbia-Suicide Severity Rating Scale Scoring and Data Analysis Guide. Version 2.0. Published Feb 2013. Accessed August 03, 2021. <http://cssrs.columbia.edu/wp-content/uploads/ScoringandDataAnalysisGuide-for-Clinical-Trials-1.pdf>
- Tietjens JR, Claman D, Kezirian EJ, et al. Obstructive sleep apnea in cardiovascular disease: a review of the literature and proposed multidisciplinary clinical management strategy. *J Am Heart Assoc*. 2019;8(1):e010440. <https://doi.org/10.1161/JAHA.118.010440>
- Weaver TE, Grunstein RR. Adherence to continuous positive airway pressure therapy: the challenge to effective treatment. *Proc Am Thorac Soc*. 2008;5(2):173-178. <https://doi.org/10.1513/pats.200708-119MG>
- Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*. 1997;20(10):835-843.
- Weaver T, Menno D, Bron M, et al. Determination of thresholds for minimally important difference and clinically important response on the functional outcomes of sleep questionnaire short version in adults with narcolepsy or obstructive sleep apnea. *Sleep Breath*. 2021;25(3):1701-1715.
- Weinberg ME, Bacchetti P, Rushakoff RJ. Frequently repeated glucose measurements overestimate the incidence of inpatient hypoglycemia and severe hyperglycemia. *J Diabetes Sci Technol*. 2010;4(3):577-582. <https://doi.org/10.1177/193229681000400311>
- [WMA] World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310(20):2191-2194. <https://doi.org/10.1001/jama.2013.281053>
- [WHO] World Health Organization. WHO STEPS surveillance manual: the WHO STEPwise approach to noncommunicable disease risk factor surveillance. Updated January 26, 2017. Accessed November 04, 2021. [https://www.who.int/ncds/surveillance/steps/STEPS\\_Manual.pdf](https://www.who.int/ncds/surveillance/steps/STEPS_Manual.pdf)
- Woodcock J, LaVange LM. Master protocols to study multiple therapies, multiple diseases, or both. *N Engl J Med*. 2017;377(1):62-70. <https://doi.org/10.1056/NEJMra1510062>

Signature Page for VV-CLIN-116270 v1.0

Approval	
Approval	

Signature Page for VV-CLIN-116270 v1.0



## Title Page

### Indication-Specific Appendix (ISA1): Participants with OSA Unwilling or Unable to use PAP Therapy

#### Confidential Information

The information contained in this document is confidential and is intended for the use of clinical investigators. It is the property of Eli Lilly and Company or its subsidiaries and should not be copied by or distributed to persons not involved in the clinical investigation of tirzepatide (LY3298176), unless such persons are bound by a confidentiality agreement with Eli Lilly and Company or its subsidiaries.

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

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**Master Protocol Number:** I8F-MC-GPIF

**ISA Protocol Number:** I8F-MC-GPI1

**ISA Amendment Number:** This is the initial version of this ISA.

**ISA Statement:** This ISA is to be performed in addition to all procedures required by protocol I8F-MC-GPIF or any subsequent amendments to that protocol.

**Compound:** Tirzepatide (LY3298176)

**Study Phase:** 3

**Acronym:** SURMOUNT-OSA

**Sponsor Name:** Eli Lilly and Company

**Legal Registered Address:** Indianapolis, Indiana, USA 46285

**Regulatory Agency Identifier Number**

IND: 157090

**Approval Date:** ISA Electronically Signed and Approved by Lilly on date provided below.

Approval Date: 28-Jan-2022 GMT

**Medical Monitor Name and Contact Information will be provided separately.**

**Table of Contents**

<b>1.</b>	<b>Protocol Summary</b> .....	<b>4</b>
1.1.	Synopsis .....	4
1.2.	Schema .....	5
1.3.	Schedule of Activities (SoA) .....	6
<b>2.</b>	<b>Introduction</b> .....	<b>7</b>
<b>3.</b>	<b>Objectives, Endpoints, and Estimands</b> .....	<b>8</b>
<b>4.</b>	<b>Study Design</b> .....	<b>9</b>
<b>5.</b>	<b>Study Population</b> .....	<b>10</b>
5.1.	Inclusion Criteria .....	10
5.2.	Exclusion Criteria .....	10
5.3.	Lifestyle Considerations .....	10
5.4.	Screen Failures .....	10
5.5.	Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant .....	10
<b>6.</b>	<b>Study Intervention(s) and Concomitant Therapy</b> .....	<b>11</b>
6.1.	Concomitant Therapy .....	11
<b>7.</b>	<b>Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal</b> .....	<b>12</b>
<b>8.</b>	<b>Study Assessments and Procedures</b> .....	<b>13</b>
8.1.	GPI1 Efficacy Assessments .....	13
8.1.1.	WatchPAT300 .....	13
<b>9.</b>	<b>Statistical Considerations</b> .....	<b>14</b>
9.1.	Analysis Sets .....	14
<b>10.</b>	<b>Supporting Documentation and Operational Considerations</b> .....	<b>15</b>
10.1.	Appendix 1: Abbreviations and Definitions .....	15
<b>11.</b>	<b>References</b> .....	<b>16</b>

## **1. Protocol Summary**

### **1.1. Synopsis**

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**ISA1 (GPI1):** Participants with OSA Unwilling or Unable to use PAP Therapy

#### **Rationale for ISA:**

This ISA (GPI1) investigates participants who are unwilling or unable to use Positive Airway Pressure (PAP) therapy.

#### **Overall Design:**

See Master Protocol GPIF for Overall Design, Number of Participants, and Intervention Groups and Duration.

## **1.2. Schema**

See Master Protocol GPIF for the Schema.

### **1.3. Schedule of Activities (SoA)**

See Master Protocol GPIF for the SoA, including ISA-specific procedures and assessments.

## **2. Introduction**

For

- Study Rationale, see Master Protocol GPIF (Section 2.1).
- Background, see Master Protocol GPIF (Section 2.2).
- Benefit/Risk Assessment, see Master Protocol GPIF (Section 2.3).

### 3. Objectives, Endpoints, and Estimands

See Master Protocol GPIF for Objectives and Endpoints that apply to all ISAs.

In addition to Master Protocol GPIF objectives and endpoints, the following apply to GPI1.

Objective	Endpoints
<b>Exploratory</b>	
<ul style="list-style-type: none"> <li>• To evaluate the effect of tirzepatide on sleep apnea as measured by WatchPAT300</li> </ul>	<ul style="list-style-type: none"> <li>• Change from baseline to Week 52 in PAT-based device determinations of:               <ul style="list-style-type: none"> <li>○ pAHI</li> <li>○ SASHB</li> </ul> </li> </ul>

Abbreviations: pAHI = peripheral apnea-hypopnea index; PAT = peripheral arterial tone; SASHB = sleep apnea-specific hypoxic burden.



#### **4. Study Design**

For

- Overall Design, see Master Protocol GPIF (Section 4.1).
- Scientific Rationale for Study Design, see Master Protocol GPIF (Section 4.2).
- Justification for Dose, see Master Protocol GPIF (Section 4.3).
- End of Study Definition, see Master Protocol GPIF (Section 4.4).

GPI1 will include participants who are unwilling or are unable to use PAP therapy.

## **5. Study Population**

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

For inclusion and exclusion criteria that apply to all ISAs, see Master Protocol GPIF (Section 5).

### **5.1. Inclusion Criteria**

1000. Participants who are unable or unwilling to use PAP therapy. Participants must not have used PAP for at least 4 weeks prior to Visit 1.

### **5.2. Exclusion Criteria**

No additional exclusion criteria apply to GPI1.

### **5.3. Lifestyle Considerations**

For lifestyle considerations, see Master Protocol GPIF (Section 5.3).

### **5.4. Screen Failures**

For definition of screen failure, see Master Protocol GPIF (Section 5.4).

For GPI1, individuals who do not meet Inclusion Criterion 1000 may be rescreened once.

Rescreened participants should be assigned a new participant number for every screening/rescreening event.

### **5.5. Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant**

Not applicable.

## **6. Study Intervention(s) and Concomitant Therapy**

Study intervention is defined as any investigational intervention(s), marketed product(s), placebo, or medical device(s) intended to be administered to/used by a study participant according to the study protocol.

For

- Study Interventions Administered, see Master Protocol GPIF (Section 6.1).
- Preparation, Handling, Storage, and Accountability, see Master Protocol GPIF (Section 6.2).
- Measures to Minimize Bias: Randomization and Blinding, including stratification factors, see Master Protocol GPIF (Section 6.3).
- Study Intervention Compliance, see Master Protocol GPIF (Section 6.4).
- Dose Modification, see Master Protocol GPIF (Section 6.5).
- Continued Access to Study Intervention after the End of the Study, see Master Protocol GPIF (Section 6.6).
- Treatment of Overdose, see Master Protocol GPIF (Section 6.7).

### **6.1. Concomitant Therapy**

For details related to Concomitant Therapy, see Master Protocol GPIF (Section 6.8).

Positive airway pressure (PAP) therapy may be initiated when urgent compensation for sleep-disordered breathing is needed, based on the opinion of the investigator. The date of PAP therapy initiation and, if applicable, discontinuation must be recorded.

## **7. Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal**

For

- Discontinuation of Study Intervention, see Master Protocol GPIF (Section 7.1).
- Participant Discontinuation/Withdrawal from the Study, see Master Protocol GPIF (Section 7.2).
- Lost to Follow up, see Master Protocol GPIF (Section 7.3).

## **8. Study Assessments and Procedures**

For Study Assessments and Procedures that apply to all ISAs, see Master Protocol GPIF (Section 8).

### **8.1. GPI1 Efficacy Assessments**

#### **8.1.1. WatchPAT300**

Wearable sensor technology (WatchPAT300) will be included per the SoA to evaluate objective assessment of changes in sleep apnea parameters (including AHI, blood oxygen saturation parameters, pulse rate, sleep parameters, etc.) throughout the trial period.

The WatchPAT300 home sleep apnea testing device (Itamar Medical Ltd.) will be utilized to explore treatment response of tirzepatide on sleep apnea-related parameters. The WatchPAT300 is a noninvasive device consisting of multiple biosensors including a proprietary technology known as PAT, pulse oximetry, accelerometer, and an acoustic sensor for detecting chest position and snoring. It is a diagnostic aid for the detection of sleep-related breathing disorders such as respiratory disturbance index and AHI, and sleep staging (REM Sleep, Light Sleep, Deep Sleep, and Wake). The WatchPAT300 has FDA 510(K) clearance (K180775) and meets the CE mark requirements (Itamar Medical Ltd.). Respiratory disturbance index and AHI calculated using PAT-based portable devices, such as the WatchPAT300, are positively correlated with those calculated from the scoring of PSG as reported in a publication by Yalamanchali et al. (2013). Participants should wear the WatchPAT300 for 1 night, 5 different times during the study as described in the SoA.

Lack of participation in WatchPAT300 collections at any time is not considered a protocol deviation.

## **9. Statistical Considerations**

For statistical considerations that apply to all ISAs, see Master Protocol GPIF (Section 9).

### **9.1. Analysis Sets**

In GPI1, for the “efficacy” estimand, the analysis will be conducted on the efficacy analysis set (EAS) as defined in Section 9.2 of the master protocol, but exclude measurements after a participant started PAP therapy during the study as described in Section 6.1. If the participant stops PAP therapy 7 or more days before the scheduled PSG, then it will be included in the analysis.

## 10. Supporting Documentation and Operational Considerations

### 10.1. Appendix 1: Abbreviations and Definitions

<b>Term</b>	<b>Definition</b>
<b>AHI</b>	Apnea-Hypopnea Index
<b>ISA</b>	Indication-Specific Appendix
<b>OSA</b>	obstructive sleep apnea
<b>PAP</b>	positive airway pressure
<b>participant</b>	Equivalent to CDISC term “subject”: an individual who participates in a clinical trial, either as recipient of an investigational medicinal product or as a control.
<b>PAT</b>	peripheral arterial tone
<b>PSG</b>	polysomnography
<b>SASHB</b>	Sleep apnea-specific hypoxic burden. SASHB is measured as the area under the curve (AUC) of baseline oxygen saturation. The unit is % minute/hour.
<b>SoA</b>	Schedule of Activities

## 11. References

Yalamanchali S, Farajian V, Hamilton C, et al. Diagnosis of obstructive sleep apnea by peripheral arterial tonometry: meta-analysis. *JAMA Otolaryngol Head Neck Surg.* 2013;139(12):1343-1350. <https://doi.org/10.1001/jamaoto.2013.5338>





## Title Page

### Intervention-Specific Appendix (ISA1): Participants with OSA Unwilling or Unable to use PAP Therapy

#### Confidential Information

The information contained in this document is confidential and is intended for the use of clinical investigators. It is the property of Eli Lilly and Company or its subsidiaries and should not be copied by or distributed to persons not involved in the clinical investigation of tirzepatide (LY3298176), unless such persons are bound by a confidentiality agreement with Eli Lilly and Company or its subsidiaries.

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

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**Master Protocol Number:** I8F-MC-GPIF

**ISA Protocol Number:** I8F-MC-GPI1

**ISA Amendment Number:** b

**ISA Statement:** This ISA is to be performed in addition to all procedures required by protocol I8F-MC-GPIF or any subsequent amendments to that protocol.

**Compound:** Tirzepatide (LY3298176)

**Study Phase:** 3

**Acronym:** SURMOUNT-OSA

**Sponsor Name:** Eli Lilly and Company

**Legal Registered Address:** Indianapolis, Indiana, USA 46285

**Regulatory Agency Identifier Number**

IND: 157090

**Approval Date:** ISA Amendment (b) Electronically Signed and Approved by Lilly on date provided below.

**Document ID:** VV-CLIN-072199

**Medical Monitor Name and Contact Information will be provided separately.**

## Protocol Amendment Summary of Changes Table

DOCUMENT HISTORY	
Document	Date
Amendment a	10-Feb-2022
Original Protocol	28-Jan-2022

### Amendment b

The master protocol GPIF Amendment b is considered to be substantial. For details related to substantiality, see Master Protocol GPIF (Protocol Amendment Summary of Changes Table).

The changes in this ISA amendment are considered to be nonsubstantial.

### Overall Rationale for the Amendment:

This amendment provides additional instructions for the use of PAP therapy.

Section # and Name	Description of Change	Brief Rationale
6.1 Concomitant Therapy	Added text, "Investigator should notify the sponsor if PAP therapy should be started."	Sponsor wants to track the cases and understand the reasons for starting PAP in GPI1.
	Modified text, "The date of PAP therapy initiation and, if applicable, <u>permanent or temporary</u> discontinuation must be recorded."	Information must be tracked in relation to data analysis.

## Table of Contents

<b>1.</b>	<b>Protocol Summary .....</b>	<b>5</b>
1.1.	Synopsis .....	5
1.2.	Schema.....	6
1.3.	Schedule of Activities (SoA).....	7
<b>2.</b>	<b>Introduction.....</b>	<b>8</b>
<b>3.</b>	<b>Objectives, Endpoints, and Estimands .....</b>	<b>9</b>
<b>4.</b>	<b>Study Design.....</b>	<b>10</b>
<b>5.</b>	<b>Study Population.....</b>	<b>11</b>
5.1.	Inclusion Criteria .....	11
5.2.	Exclusion Criteria .....	11
5.3.	Lifestyle Considerations .....	11
5.4.	Screen Failures.....	11
5.5.	Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant.....	11
<b>6.</b>	<b>Study Intervention(s) and Concomitant Therapy .....</b>	<b>12</b>
6.1.	Concomitant Therapy .....	12
<b>7.</b>	<b>Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal.....</b>	<b>13</b>
<b>8.</b>	<b>Study Assessments and Procedures.....</b>	<b>14</b>
8.1.	GPI1 Efficacy Assessments .....	14
8.1.1.	WatchPAT300.....	14
<b>9.</b>	<b>Statistical Considerations.....</b>	<b>15</b>
9.1.	Analysis Sets.....	15
<b>10.</b>	<b>Supporting Documentation and Operational Considerations .....</b>	<b>16</b>
10.1.	Appendix 1: Abbreviations and Definitions .....	16
10.2.	Appendix 2: Protocol Amendment History .....	17
<b>11.</b>	<b>References.....</b>	<b>18</b>

## **1. Protocol Summary**

### **1.1. Synopsis**

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**ISA1 (GPI1):** Participants with OSA Unwilling or Unable to use PAP Therapy

#### **Rationale for ISA:**

This ISA (GPI1) investigates participants who are unwilling or unable to use Positive Airway Pressure (PAP) therapy.

#### **Overall Design:**

See Master Protocol GPIF for Overall Design, Number of Participants, and Intervention Groups and Duration.

## **1.2. Schema**

See Master Protocol GPIF for the Schema.

### **1.3. Schedule of Activities (SoA)**

See Master Protocol GPIF for the SoA, including ISA-specific procedures and assessments.



## **2. Introduction**

For

- Study Rationale, see Master Protocol GPIF (Section 2.1).
- Background, see Master Protocol GPIF (Section 2.2).
- Benefit/Risk Assessment, see Master Protocol GPIF (Section 2.3).

### 3. Objectives, Endpoints, and Estimands

See Master Protocol GPIF for Objectives and Endpoints that apply to all ISAs.

In addition to Master Protocol GPIF objectives and endpoints, the following apply to GPI1.

Objective	Endpoints
<b>Exploratory</b>	
<ul style="list-style-type: none"> <li>• To evaluate the effect of tirzepatide on sleep apnea as measured by WatchPAT300</li> </ul>	<ul style="list-style-type: none"> <li>• Change from baseline to Week 52 in PAT-based device determinations of:               <ul style="list-style-type: none"> <li>○ pAHI</li> <li>○ SASHB</li> </ul> </li> </ul>

Abbreviations: pAHI = peripheral apnea-hypopnea index; PAT = peripheral arterial tone; SASHB = sleep apnea-specific hypoxic burden.

#### **4. Study Design**

For

- Overall Design, see Master Protocol GPIF (Section 4.1).
- Scientific Rationale for Study Design, see Master Protocol GPIF (Section 4.2).
- Justification for Dose, see Master Protocol GPIF (Section 4.3).
- End of Study Definition, see Master Protocol GPIF (Section 4.4).

GPI1 will include participants who are unwilling or are unable to use PAP therapy.

## **5. Study Population**

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

For inclusion and exclusion criteria that apply to all ISAs, see Master Protocol GPIF (Section 5).

### **5.1. Inclusion Criteria**

1000. Participants who are unable or unwilling to use PAP therapy. Participants must not have used PAP for at least 4 weeks prior to Visit 1.

### **5.2. Exclusion Criteria**

No additional exclusion criteria apply to GPI1.

### **5.3. Lifestyle Considerations**

For lifestyle considerations, see Master Protocol GPIF (Section 5.3).

### **5.4. Screen Failures**

For definition of screen failure, see Master Protocol GPIF (Section 5.4).

For GPI1, individuals who do not meet Inclusion Criterion 1000 may be rescreened once.

Rescreened participants should be assigned a new participant number for every screening/rescreening event.

### **5.5. Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant**

Not applicable.

## **6. Study Intervention(s) and Concomitant Therapy**

Study intervention is defined as any investigational intervention(s), marketed product(s), placebo, or medical device(s) intended to be administered to/used by a study participant according to the study protocol.

For

- Study Interventions Administered, see Master Protocol GPIF (Section 6.1).
- Preparation, Handling, Storage, and Accountability, see Master Protocol GPIF (Section 6.2).
- Measures to Minimize Bias: Randomization and Blinding, including stratification factors, see Master Protocol GPIF (Section 6.3).
- Study Intervention Compliance, see Master Protocol GPIF (Section 6.4).
- Dose Modification, see Master Protocol GPIF (Section 6.5).
- Continued Access to Study Intervention after the End of the Study, see Master Protocol GPIF (Section 6.6).
- Treatment of Overdose, see Master Protocol GPIF (Section 6.7).

### **6.1. Concomitant Therapy**

For details related to Concomitant Therapy, see Master Protocol GPIF (Section 6.8).

Positive airway pressure (PAP) therapy may be initiated when urgent compensation for sleep-disordered breathing is needed, based on the opinion of the investigator. Investigator should notify the sponsor if PAP therapy should be started. The date of PAP therapy initiation and, if applicable, permanent or temporary discontinuation must be recorded.

## **7. Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal**

For

- Discontinuation of Study Intervention, see Master Protocol GPIF (Section 7.1).
- Participant Discontinuation/Withdrawal from the Study, see Master Protocol GPIF (Section 7.2).
- Lost to Follow up, see Master Protocol GPIF (Section 7.3).

## **8. Study Assessments and Procedures**

For Study Assessments and Procedures that apply to all ISAs, see Master Protocol GPIF (Section 8).

### **8.1. GPI1 Efficacy Assessments**

#### **8.1.1. WatchPAT300**

Wearable sensor technology (WatchPAT300) will be included per the SoA to evaluate objective assessment of changes in sleep apnea parameters (including AHI, blood oxygen saturation parameters, pulse rate, sleep parameters, etc.) throughout the trial period.

The WatchPAT300 home sleep apnea testing device (Itamar Medical Ltd.) will be utilized to explore treatment response of tirzepatide on sleep apnea-related parameters. The WatchPAT300 is a noninvasive device consisting of multiple biosensors including a proprietary technology known as PAT, pulse oximetry, accelerometer, and an acoustic sensor for detecting chest position and snoring. It is a diagnostic aid for the detection of sleep-related breathing disorders such as respiratory disturbance index and AHI, and sleep staging (REM Sleep, Light Sleep, Deep Sleep, and Wake). The WatchPAT300 has FDA 510(K) clearance (K180775) and meets the CE mark requirements (Itamar Medical Ltd.). Respiratory disturbance index and AHI calculated using PAT-based portable devices, such as the WatchPAT300, are positively correlated with those calculated from the scoring of PSG as reported in a publication by Yalamanchali et al. (2013). Participants should wear the WatchPAT300 for 1 night, 5 different times during the study as described in the SoA.

Lack of participation in WatchPAT300 collections at any time is not considered a protocol deviation.

## **9. Statistical Considerations**

For statistical considerations that apply to all ISAs, see Master Protocol GPIF (Section 9).

### **9.1. Analysis Sets**

In GPI1, for the “efficacy” estimand, the analysis will be conducted on the efficacy analysis set (EAS) as defined in Section 9.2 of the master protocol, but exclude measurements after a participant started PAP therapy during the study as described in Section 6.1. If the participant stops PAP therapy 7 or more days before the scheduled PSG, then it will be included in the analysis.



## 10. Supporting Documentation and Operational Considerations

### 10.1. Appendix 1: Abbreviations and Definitions

Term	Definition
<b>AHI</b>	Apnea-Hypopnea Index
<b>ISA</b>	intervention-specific appendix
<b>OSA</b>	obstructive sleep apnea
<b>PAP</b>	positive airway pressure
<b>participant</b>	Equivalent to CDISC term “subject”: an individual who participates in a clinical trial, either as recipient of an investigational medicinal product or as a control.
<b>PAT</b>	peripheral arterial tone
<b>PSG</b>	polysomnography
<b>SASHB</b>	Sleep apnea-specific hypoxic burden. SASHB is measured as the area under the curve (AUC) of baseline oxygen saturation. The unit is % minute/hour.
<b>SoA</b>	Schedule of Activities

## 10.2. Appendix 2: Protocol Amendment History

The Protocol Amendment Summary of Changes Table for the current amendment is located directly before the Table of Contents (TOC).

### Amendment a: 10-Feb-2022

#### Overall Rationale for the Amendment

This amendment corrects the terminology used in the protocol for intervention-specific appendix (ISA).

<b>Section # and Name</b>	<b>Description of Change</b>	<b>Brief Rationale</b>
Title Page 10.10. Abbreviations and Definitions	Changed “indication-specific appendix” to “intervention- specific appendix.	Correction.

## 11. References

Yalamanchali S, Farajian V, Hamilton C, et al. Diagnosis of obstructive sleep apnea by peripheral arterial tonometry: meta-analysis. *JAMA Otolaryngol Head Neck Surg.* 2013;139(12):1343-1350. <https://doi.org/10.1001/jamaoto.2013.5338>

Signature Page for VV-CLIN-072199 v1.0

Approval

Approval



PPD

Signature Page for VV-CLIN-072199 v1.0

## Title Page

### Indication-Specific Appendix (ISA2): Participants with OSA on PAP Therapy

#### Confidential Information

The information contained in this document is confidential and is intended for the use of clinical investigators. It is the property of Eli Lilly and Company or its subsidiaries and should not be copied by or distributed to persons not involved in the clinical investigation of tirzepatide (LY3298176), unless such persons are bound by a confidentiality agreement with Eli Lilly and Company or its subsidiaries.

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

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**Master Protocol Number:** I8F-MC-GPIF

**ISA Protocol Number:** I8F-MC-GPI2

**ISA Amendment Number:** This is the initial protocol of this ISA.

**ISA Statement:** This ISA is to be performed in addition to all procedures required by protocol I8F-MC-GPIF or any subsequent amendments to that protocol.

**Compound:** Tirzepatide (LY3298176)

**Study Phase:** 3

**Acronym:** SURMOUNT-OSA

**Sponsor Name:** Eli Lilly and Company

**Legal Registered Address:** Indianapolis, Indiana, USA 46285

**Regulatory Agency Identifier Number**

IND: 157090

**Approval Date:** ISA Electronically Signed and Approved by Lilly on date provided below.

Approval Date: 28-Jan-2022 GMT

**Medical Monitor Name and Contact Information will be provided separately.**

**Table of Contents**

<b>1.</b>	<b>Protocol Summary</b> .....	<b>4</b>
1.1.	Synopsis .....	4
1.2.	Schema .....	5
1.3.	Schedule of Activities (SoA) .....	6
<b>2.</b>	<b>Introduction</b> .....	<b>7</b>
<b>3.</b>	<b>Objectives, Endpoints, and Estimands</b> .....	<b>8</b>
<b>4.</b>	<b>Study Design</b> .....	<b>9</b>
<b>5.</b>	<b>Study Population</b> .....	<b>10</b>
5.1.	Inclusion Criteria .....	10
5.2.	Exclusion Criteria .....	10
5.3.	Lifestyle Considerations .....	10
5.4.	Screen Failures .....	10
5.5.	Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant .....	10
<b>6.</b>	<b>Study Intervention(s) and Concomitant Therapy</b> .....	<b>11</b>
<b>7.</b>	<b>Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal</b> .....	<b>12</b>
<b>8.</b>	<b>Study Assessments and Procedures</b> .....	<b>13</b>
8.1.	Polysomnography .....	13
8.2.	PAP Compliance Assessment .....	13
<b>9.</b>	<b>Statistical Considerations</b> .....	<b>14</b>
9.1.	Statistical Analyses .....	14
9.1.1.	Subgroup Analyses .....	14
<b>10.</b>	<b>Supporting Documentation and Operational Considerations</b> .....	<b>15</b>
10.1.	Appendix 1: Abbreviations and Definitions .....	15
<b>11.</b>	<b>References</b> .....	<b>16</b>

## **1. Protocol Summary**

### **1.1. Synopsis**

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

I8F-MC-GPI2: Participants with OSA on PAP Therapy

#### **Rationale for ISA:**

This ISA (GPI2) investigates participants who are on Positive Airway Pressure (PAP) therapy at time of screening and plan to continue PAP therapy during the study.

#### **Objectives, Endpoints, and Estimands:**

See Master Protocol GPIF for Objectives and Endpoints that apply to all ISAs.

No additional objectives and endpoints apply to GPI2.

#### **Overall Design:**

See Master Protocol GPIF for Overall Design, Number of Participants, and Intervention Groups and Duration.



## **1.2. Schema**

See Master Protocol GPIF for the Schema.

### **1.3. Schedule of Activities (SoA)**

See Master Protocol GPIF for the SoA, including ISA-specific procedures and assessments.

## **2. Introduction**

For

- Study Rationale, see Master Protocol GPIF (Section 2.1).
- Background, see Master Protocol GPIF (Section 2.2).
- Benefit/Risk Assessment, see Master Protocol GPIF (Section 2.3).

### **3. Objectives, Endpoints, and Estimands**

See Master Protocol GPIF for Objectives and Endpoints that apply to all ISAs.

There are no additional objectives and endpoints for GPI2.

#### **4. Study Design**

For

- Overall Design, see Master Protocol GPIF (Section 4.1).
- Scientific Rationale for Study Design, see Master Protocol GPIF (Section 4.2).
- Justification for Dose, see Master Protocol GPIF (Section 4.3).
- End of Study Definition, see Master Protocol GPIF (Section 4.4).

GPI2 will include participants who are on PAP therapy for at least 3 consecutive months at time of screening (Visit 1) and plan to continue PAP therapy during the study.

## **5. Study Population**

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

For inclusion and exclusion criteria that apply to all ISAs, see Master Protocol GPIF (Section 5).

### **5.1. Inclusion Criteria**

2000. Have been on PAP therapy for at least 3 consecutive months prior to Visit 1 and plan to continue PAP therapy during the study.

### **5.2. Exclusion Criteria**

2001. Have personal or job-related responsibilities, or in the opinion of the investigator have any situation, that would make it unsafe to stop PAP therapy for 7 days prior to PSG testing during the course of the study.
2002. Are unwilling to stop PAP therapy for 7 days prior to PSG testing during the course of the study. See GPI2 Section 8.1.

### **5.3. Lifestyle Considerations**

For lifestyle considerations, see Master Protocol GPIF (Section 5.3).

### **5.4. Screen Failures**

For definition of screen failure, see Master Protocol GPIF (Section 5.4).

For GPI2, individuals who do not meet Inclusion Criterion 2000 may be rescreened once.

Rescreened participants should be assigned a new participant number for every screening/rescreening event.

### **5.5. Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant**

Not applicable.

## **6. Study Intervention(s) and Concomitant Therapy**

Study intervention is defined as any investigational intervention(s), marketed product(s), placebo, or medical device(s) intended to be administered to/used by a study participant according to the study protocol.

For

- Study Interventions Administered, see Master Protocol GPIF (Section 6.1).
- Preparation, Handling, Storage, and Accountability, see Master Protocol GPIF (Section 6.2).
- Measures to Minimize Bias: Randomization and Blinding, including stratification factors, see Master Protocol GPIF (Section 6.3).
- Study Intervention Compliance, see Master Protocol GPIF (Section 6.4).
- Dose Modification, see Master Protocol GPIF (Section 6.5).
- Continued Access to Study Intervention after the End of the Study, see Master Protocol GPIF (Section 6.6).
- Treatment of Overdose, see Master Protocol GPIF (Section 6.7).
- Concomitant Therapy, see Master Protocol GPIF (Section 6.8).

## **7. Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal**

For

- Discontinuation of Study Intervention, see Master Protocol GPIF (Section 7.1).
- Participant Discontinuation/Withdrawal from the Study, see Master Protocol GPIF (Section 7.2).
- Lost to Follow up, see Master Protocol GPIF (Section 7.3).



## **8. Study Assessments and Procedures**

For Study Assessments and Procedures that apply to all ISA, see Master Protocol GPIF (Section 8).

### **8.1. Polysomnography**

Participants in ISA2 (on PAP) will suspend their PAP use for 7 days prior to the scheduled PSGs. The wash-out period is to allow for measurement of drug treatment effect unobscured by PAP and is based on literature citing 14 days as a safe discontinuation period (Rossi et al. 2012, Schwarz et al. 2016a, Schwarz et al. 2016b, Schwarz et al. 2016c, Schwarz et al. 2018, Stradling et al. 2015).

Suspension of PAP for up to 9 days will not be considered a protocol deviation.

### **8.2. PAP Compliance Assessment**

Participant adherence to PAP will be collected via the PAP adherence diary (values recorded should be based on device output). Factors of compliance include how many days in a week and how many hours a night the device is used. Data will be analyzed as described in the statistical analysis plan (SAP), and used in subgroup analyses described in Section 9.1.1.

## **9. Statistical Considerations**

For Statistical Considerations that apply to all ISAs, see Master Protocol GPIF (Section 9).

### **9.1. Statistical Analyses**

#### **9.1.1. Subgroup Analyses**

In addition to the subgroup analyses listed in the master protocol 9.3.7.2, the following additional subgroup analyses will be conducted for GPI2. Details on PAP adherence will be described in the SAP.

- Baseline PAP adherence (adhere, not adhere)

## 10. Supporting Documentation and Operational Considerations

### 10.1. Appendix 1: Abbreviations and Definitions

---

<b>Term</b>	<b>Definition</b>
<b>ISA</b>	Indication-Specific Appendix
<b>OSA</b>	obstructive sleep apnea
<b>PAP</b>	positive airway pressure
<b>participant</b>	Equivalent to CDISC term “subject”: an individual who participates in a clinical trial, either as recipient of an investigational medicinal product or as a control
<b>PSG</b>	polysomnography
<b>SAP</b>	statistical analysis plan

## 11. References

- Rossi VA, Stoewhas AC, Camen G, et al. The effects of continuous positive airway pressure therapy withdrawal on cardiac repolarization: data from a randomized controlled trial. *Eur Heart J*. 2012;33(17):2206-2212. <https://doi.org/10.1093/eurheartj/ehs073>
- Schwarz EI, Schlatzer C, Rossi VA, et al. Effect of CPAP withdrawal on BP in OSA data from three randomized controlled trials. *Chest*. 2016a;150(6):1202–1210. <https://doi.org/10.1016/j.chest.2016.07.012>
- Schwarz EI, Schlatzer C, Stehli J, et al. Effect of CPAP withdrawal on myocardial perfusion in OSA: A randomized controlled trial. *Respirology*. 2016b;21(6):1126-1133. <https://doi.org/10.1111/resp.12798>
- Schwarz EI, Stradling JR, Kohler M. Physiological consequences of CPAP therapy withdrawal in patients with obstructive sleep apnoea - an opportunity for an efficient experimental model. *J Thorac Dis*. 2018;10(Suppl 1):S24–S32. <https://doi.org/10.21037/jtd.2017.12.142>
- Schwarz EI, Sinues PML, Bregy L, et al. Effects of CPAP therapy withdrawal on exhaled breath pattern in obstructive sleep apnoea. *Thorax*. 2016c;71(2):110-117. <https://doi.org/10.1136/thoraxjnl-2015-207597>
- Stradling JR, Schwarz, EI, Schlatzer C, et al. Biomarkers of oxidative stress following continuous positive airway pressure withdrawal: data from two randomised trials. *Eur Respir J*. 2015;46(4):1065-1071. <https://doi.org/10.1183/09031936.00023215>



## Title Page

### Intervention-Specific Appendix (ISA2): Participants with OSA on PAP Therapy

#### Confidential Information

The information contained in this document is confidential and is intended for the use of clinical investigators. It is the property of Eli Lilly and Company or its subsidiaries and should not be copied by or distributed to persons not involved in the clinical investigation of tirzepatide (LY3298176), unless such persons are bound by a confidentiality agreement with Eli Lilly and Company or its subsidiaries.

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

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**Master Protocol Number:** I8F-MC-GPIF

**ISA Protocol Number:** I8F-MC-GPI2

**ISA Amendment Number:** a

**ISA Statement:** This ISA is to be performed in addition to all procedures required by protocol I8F-MC-GPIF or any subsequent amendments to that protocol.

**Compound:** Tirzepatide (LY3298176)

**Study Phase:** 3

**Acronym:** SURMOUNT-OSA

**Sponsor Name:** Eli Lilly and Company

**Legal Registered Address:** Indianapolis, Indiana, USA 46285

**Regulatory Agency Identifier Number**

IND: 157090

**Approval Date:** ISA Electronically Signed and Approved by Lilly on date provided below.

Approval Date: 12-Feb-2022 GMT

**Medical Monitor Name and Contact Information will be provided separately.**

## Protocol Amendment Summary of Changes Table

DOCUMENT HISTORY	
Document	Date
Original Protocol	28 Jan 2022

### Amendment a

#### Overall Rationale for the Amendment:

This amendment corrects the terminology used in the protocol for intervention-specific appendix (ISA).

Section # and Name	Description of Change	Brief Rationale
Title Page 10.10. Abbreviations and Definitions	Changed “indication-specific appendix” to “intervention- specific appendix.	Correction.



**Table of Contents**

<b>1.</b>	<b>Protocol Summary .....</b>	<b>5</b>
1.1.	Synopsis .....	5
1.2.	Schema.....	6
1.3.	Schedule of Activities (SoA) .....	7
<b>2.</b>	<b>Introduction .....</b>	<b>8</b>
<b>3.</b>	<b>Objectives, Endpoints, and Estimands .....</b>	<b>9</b>
<b>4.</b>	<b>Study Design .....</b>	<b>10</b>
<b>5.</b>	<b>Study Population .....</b>	<b>11</b>
5.1.	Inclusion Criteria .....	11
5.2.	Exclusion Criteria.....	11
5.3.	Lifestyle Considerations .....	11
5.4.	Screen Failures .....	11
5.5.	Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant .....	11
<b>6.</b>	<b>Study Intervention(s) and Concomitant Therapy .....</b>	<b>12</b>
<b>7.</b>	<b>Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal.....</b>	<b>13</b>
<b>8.</b>	<b>Study Assessments and Procedures .....</b>	<b>14</b>
8.1.	Polysomnography .....	14
8.2.	PAP Compliance Assessment .....	14
<b>9.</b>	<b>Statistical Considerations .....</b>	<b>15</b>
9.1.	Statistical Analyses.....	15
9.1.1.	Subgroup Analyses .....	15
<b>10.</b>	<b>Supporting Documentation and Operational Considerations .....</b>	<b>16</b>
10.1.	Appendix 1: Abbreviations and Definitions .....	16
<b>11.</b>	<b>References .....</b>	<b>17</b>

## 1. Protocol Summary

### 1.1. Synopsis

**Master Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

I8F-MC-GPI2: Participants with OSA on PAP Therapy

#### **Rationale for ISA:**

This ISA (GPI2) investigates participants who are on Positive Airway Pressure (PAP) therapy at time of screening and plan to continue PAP therapy during the study.

#### **Objectives, Endpoints, and Estimands:**

See Master Protocol GPIF for Objectives and Endpoints that apply to all ISAs.

No additional objectives and endpoints apply to GPI2.

#### **Overall Design:**

See Master Protocol GPIF for Overall Design, Number of Participants, and Intervention Groups and Duration.

## **1.2. Schema**

See Master Protocol GPIF for the Schema.

**1.3. Schedule of Activities (SoA)**

See Master Protocol GPIF for the SoA, including ISA-specific procedures and assessments.

## **2. Introduction**

For

- Study Rationale, see Master Protocol GPIF (Section 2.1).
- Background, see Master Protocol GPIF (Section 2.2).
- Benefit/Risk Assessment, see Master Protocol GPIF (Section 2.3).

### **3. Objectives, Endpoints, and Estimands**

See Master Protocol GPIF for Objectives and Endpoints that apply to all ISAs.

There are no additional objectives and endpoints for GPI2.

#### **4. Study Design**

For

- Overall Design, see Master Protocol GPIF (Section 4.1).
- Scientific Rationale for Study Design, see Master Protocol GPIF (Section 4.2).
- Justification for Dose, see Master Protocol GPIF (Section 4.3).
- End of Study Definition, see Master Protocol GPIF (Section 4.4).

GPI2 will include participants who are on PAP therapy for at least 3 consecutive months at time of screening (Visit 1) and plan to continue PAP therapy during the study.

## **5. Study Population**

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

For inclusion and exclusion criteria that apply to all ISAs, see Master Protocol GPIF (Section 5).

### **5.1. Inclusion Criteria**

2000. Have been on PAP therapy for at least 3 consecutive months prior to Visit 1 and plan to continue PAP therapy during the study.

### **5.2. Exclusion Criteria**

2001. Have personal or job-related responsibilities, or in the opinion of the investigator have any situation, that would make it unsafe to stop PAP therapy for 7 days prior to PSG testing during the course of the study.
2002. Are unwilling to stop PAP therapy for 7 days prior to PSG testing during the course of the study. See GPI2 Section 8.1.

### **5.3. Lifestyle Considerations**

For lifestyle considerations, see Master Protocol GPIF (Section 5.3).

### **5.4. Screen Failures**

For definition of screen failure, see Master Protocol GPIF (Section 5.4).

For GPI2, individuals who do not meet Inclusion Criterion 2000 may be rescreened once.

Rescreened participants should be assigned a new participant number for every screening/rescreening event.

### **5.5. Criteria for Temporarily Delaying Enrollment/Randomization/Administration of Study Intervention of a Participant**

Not applicable.



## **6. Study Intervention(s) and Concomitant Therapy**

Study intervention is defined as any investigational intervention(s), marketed product(s), placebo, or medical device(s) intended to be administered to/used by a study participant according to the study protocol.

For

- Study Interventions Administered, see Master Protocol GPIF (Section 6.1).
- Preparation, Handling, Storage, and Accountability, see Master Protocol GPIF (Section 6.2).
- Measures to Minimize Bias: Randomization and Blinding, including stratification factors, see Master Protocol GPIF (Section 6.3).
- Study Intervention Compliance, see Master Protocol GPIF (Section 6.4).
- Dose Modification, see Master Protocol GPIF (Section 6.5).
- Continued Access to Study Intervention after the End of the Study, see Master Protocol GPIF (Section 6.6).
- Treatment of Overdose, see Master Protocol GPIF (Section 6.7).
- Concomitant Therapy, see Master Protocol GPIF (Section 6.8).

## **7. Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal**

For

- Discontinuation of Study Intervention, see Master Protocol GPIF (Section 7.1).
- Participant Discontinuation/Withdrawal from the Study, see Master Protocol GPIF (Section 7.2).
- Lost to Follow up, see Master Protocol GPIF (Section 7.3).

## **8. Study Assessments and Procedures**

For Study Assessments and Procedures that apply to all ISA, see Master Protocol GPIF (Section 8).

### **8.1. Polysomnography**

Participants in ISA2 (on PAP) will suspend their PAP use for 7 days prior to the scheduled PSGs. The wash-out period is to allow for measurement of drug treatment effect unobscured by PAP and is based on literature citing 14 days as a safe discontinuation period (Rossi et al. 2012, Schwarz et al. 2016a, Schwarz et al. 2016b, Schwarz et al. 2016c, Schwarz et al. 2018, Stradling et al. 2015).

Suspension of PAP for up to 9 days will not be considered a protocol deviation.

### **8.2. PAP Compliance Assessment**

Participant adherence to PAP will be collected via the PAP adherence diary (values recorded should be based on device output). Factors of compliance include how many days in a week and how many hours a night the device is used. Data will be analyzed as described in the statistical analysis plan (SAP), and used in subgroup analyses described in Section 9.1.1.

## **9. Statistical Considerations**

For Statistical Considerations that apply to all ISAs, see Master Protocol GPIF (Section 9).

### **9.1. Statistical Analyses**

#### **9.1.1. Subgroup Analyses**

In addition to the subgroup analyses listed in the master protocol 9.3.7.2, the following additional subgroup analyses will be conducted for GPI2. Details on PAP adherence will be described in the SAP.

- Baseline PAP adherence (adhere, not adhere)

## 10. Supporting Documentation and Operational Considerations

### 10.1. Appendix 1: Abbreviations and Definitions

---

<b>Term</b>	<b>Definition</b>
<b>ISA</b>	intervention-specific appendix
<b>OSA</b>	obstructive sleep apnea
<b>PAP</b>	positive airway pressure
<b>participant</b>	Equivalent to CDISC term “subject”: an individual who participates in a clinical trial, either as recipient of an investigational medicinal product or as a control
<b>PSG</b>	polysomnography
<b>SAP</b>	statistical analysis plan

## 11. References

- Rossi VA, Stoewhas AC, Camen G, et al. The effects of continuous positive airway pressure therapy withdrawal on cardiac repolarization: data from a randomized controlled trial. *Eur Heart J*. 2012;33(17):2206-2212. <https://doi.org/10.1093/eurheartj/ehs073>
- Schwarz EI, Schlatzer C, Rossi VA, et al. Effect of CPAP withdrawal on BP in OSA data from three randomized controlled trials. *Chest*. 2016a;150(6):1202–1210. <https://doi.org/10.1016/j.chest.2016.07.012>
- Schwarz EI, Schlatzer C, Stehli J, et al. Effect of CPAP withdrawal on myocardial perfusion in OSA: A randomized controlled trial. *Respirology*. 2016b;21(6):1126-1133. <https://doi.org/10.1111/resp.12798>
- Schwarz EI, Stradling JR, Kohler M. Physiological consequences of CPAP therapy withdrawal in patients with obstructive sleep apnoea - an opportunity for an efficient experimental model. *J Thorac Dis*. 2018;10(Suppl 1):S24–S32. <https://doi.org/10.21037/jtd.2017.12.142>
- Schwarz EI, Sinues PML, Bregy L, et al. Effects of CPAP therapy withdrawal on exhaled breath pattern in obstructive sleep apnoea. *Thorax*. 2016c;71(2):110-117. <https://doi.org/10.1136/thoraxjnl-2015-207597>
- Stradling JR, Schwarz, EI, Schlatzer C, et al. Biomarkers of oxidative stress following continuous positive airway pressure withdrawal: data from two randomised trials. *Eur Respir J*. 2015;46(4):1065-1071. <https://doi.org/10.1183/09031936.00023215>



# **Statistical Analysis Plan I8F-MC-GPIF: Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)**

**Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**Protocol Number:** I8F-MC-GPIF

**Compound Number:** Tirzepatide (LY3298176)

**Short Title:** A Statistical Analysis Plan for Tirzepatide in Participants with Obstructive Sleep Apnea and Obesity

**Sponsor Name:** Eli Lilly and Company

**Legal Registered Address:** Indianapolis, Indiana ISA 46285

**Regulatory Agency Identifier Number(s)**

IND: 157090

## **Confidential Information**

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

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

Document ID: VV-CLIN-069584



## Table of Contents

Statistical Analysis Plan I8F-MC-GPIF: Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA).....	1
Table of Contents .....	2
Version History .....	4
1. Introduction .....	8
1.1. Objectives, Endpoints, and Estimands .....	8
1.1.1. Estimands .....	10
1.2. Study Design .....	12
2. Statistical Hypotheses .....	14
2.1. Multiplicity Adjustment .....	14
3. Analysis Sets .....	15
4. Statistical Analyses.....	16
4.1. General Considerations .....	16
4.1.1. Adjustment for Covariates .....	18
4.1.2. Handling of Dropouts or Missing Data.....	19
4.1.3. Multicenter Studies .....	19
4.1.4. Historical Illnesses and Preexisting Conditions.....	20
4.1.5. Patient Characteristics .....	20
4.1.6. Concomitant Therapy.....	20
4.1.7. Treatment Exposure and Compliance .....	20
4.1.8. Important Protocol Deviations .....	21
4.2. Participant Dispositions.....	21
4.3. Primary Endpoint Analysis.....	21
4.3.1. Analysis Related to the Efficacy Estimand.....	22
4.3.2. Analysis Related to the Treatment Regimen Estimand .....	22
4.3.3. Sensitivity Analyses .....	22
4.4. Secondary Endpoints Analysis .....	23
4.4.1. Key Secondary Endpoints .....	23
4.4.2. Type I Error Rate Control Strategy for Primary and Key Secondary Efficacy Analyses.....	25
4.4.3. Supportive Secondary Endpoints .....	25
4.5. Exploratory Endpoint Analyses.....	27
4.6. Safety Analyses .....	28
4.6.1. Analysis of Adverse Events .....	28

4.6.2.	Patient Narratives .....	30
4.6.3.	Special Safety Topics.....	30
4.6.4.	Vital Signs.....	44
4.6.5.	Clinical Laboratory Evaluation.....	45
4.7.	Other Analyses .....	46
4.7.1.	PAP Adherence.....	46
4.7.2.	Health Outcomes.....	46
4.7.3.	Subgroup Analyses .....	48
4.8.	Interim Analyses.....	49
4.8.1.	Unblinding Plan .....	49
4.9.	Changes to Protocol-Planned Analyses.....	49
5.	Sample Size Determination .....	50
6.	Supporting Documentation.....	51
6.1.	Appendix 1: Clinical Trial Registry Analyses.....	51
6.2.	Appendix 2: Search Criteria for Special Safety Topics.....	51
6.3.	Appendix 3: Magnetic Resonance Imaging Addendum.....	53
7.	References .....	55

## Version History

This statistical analysis plan (SAP) is the first version and is based on amendment (b) of the protocol for I8B-MC-GPIF (GPIF) approved on 30 September 2022. This SAP was approved prior to the first unblinding of the treatment assignments for an Independent Data Monitoring Committee review during the study.

**Table GPIF.1.1. SAP Version History Summary**

<b>SAP Version</b>	<b>Approval Date</b>	<b>Change</b>	<b>Rationale</b>
1	See date on Page 1	Not Applicable	Original version

## Abbreviations and Definitions

<b>Term</b>	<b>Definition</b>
<b>ADA</b>	anti-drug antibody
<b>AE</b>	adverse event
<b>AESI</b>	adverse event of special interest
<b>AHI</b>	Apnea-Hypopnea Index
<b>ANCOVA</b>	analysis of covariance
<b>BG</b>	blood glucose
<b>BMI</b>	body mass index
<b>CSR</b>	clinical study report
<b>C-SSRS</b>	Columbia-Suicide Severity Rating Scale
<b>EAS</b>	efficacy analysis set
<b>eCRF</b>	electronic case report form
<b>EQ-5D-5L</b>	EuroQol - 5 Dimension - 5 Level
<b>ESS</b>	Epworth Sleepiness Scale
<b>FAS</b>	Full Analysis Set
<b>FOSQ</b>	Functional Outcomes of Sleep Questionnaire
<b>FOSQ-10</b>	Functional Outcomes of Sleep Questionnaire, 10 items
<b>GI</b>	gastrointestinal
<b>GIP</b>	glucose-dependent insulintropic polypeptide
<b>GIPR</b>	glucose-dependent insulintropic polypeptide receptor
<b>GLP-1</b>	glucagon-like peptide-1
<b>GLP-1R</b>	glucagon-like peptide-1 receptor
<b>HLT</b>	High Level Term
<b>ISA</b>	intervention-specific appendix
<b>ISR</b>	injection-site reaction
<b>LLT</b>	Lowest Level Term

---

<b>Term</b>	<b>Definition</b>
<b>MACE</b>	major adverse cardiovascular event(s)
<b>MedDRA</b>	Medical Dictionary for Regulatory Activities
<b>mITT</b>	modified intent-to-treat
<b>MMRM</b>	mixed model repeated measures
<b>MRD</b>	minimum required dilution
<b>MRI</b>	magnetic resonance imaging
<b>MTD</b>	maximum tolerated dose
<b>NAb</b>	neutralizing antibodies
<b>NAb-</b>	neutralizing antibody negative
<b>NAb+</b>	neutralizing antibody positive
<b>OSA</b>	obstructive sleep apnea
<b>PAP</b>	positive airway pressure
<b>PD</b>	pharmacodynamic
<b>PG</b>	plasma glucose
<b>PGIS</b>	Patient Global Impression of Status
<b>PHQ-9</b>	Patient Health Questionnaire
<b>PK</b>	pharmacokinetic
<b>PRO</b>	patient-reported outcome
<b>PROMIS</b>	Patient-Reported Outcomes Measurement Information System
<b>PSG</b>	polysomnography
<b>PT</b>	Preferred Term
<b>REML</b>	restricted maximum likelihood
<b>SAE</b>	serious adverse event
<b>SAP</b>	statistical analysis plan
<b>SBP</b>	systolic blood pressure
<b>SF36v2</b>	Short-Form 36 version 2

---

---

<b>Term</b>	<b>Definition</b>
<b>SMQ</b>	Standardized MedDRA Query
<b>SOC</b>	System Organ Class
<b>TE ADA</b>	treatment-emergent anti-drug antibody
<b>TE ADA-</b>	treatment-emergent anti-drug antibody negative
<b>TE ADA+</b>	treatment-emergent anti-drug antibody positive
<b>TEAE</b>	treatment-emergent adverse event
<b>TZP</b>	tirzepatide
<b>UACR</b>	urine albumin-to-creatinine ratio
<b>ULN</b>	upper limit of normal

---

# 1. Introduction

## 1.1. Objectives, Endpoints, and Estimands

Primary Objective	Endpoints
To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for percent decrease in AHI.	Percent change in AHI from baseline to Week 52.
Key Secondary Objectives (controlled for type I error)	Endpoints
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in AHI</li> <li>• A hierarchical assessment of PROs</li>   <li>• Clinically meaningful change in AHI</li> <li>• Achieving OSA remission or mild nonsymptomatic OSA</li> <li>• Change in body weight</li> <li>• Change in inflammatory status</li>   <li>• Change in SBP</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in AHI</li> <li>• A hierarchical combination of the following:                             <ul style="list-style-type: none"> <li>○ Change in FOSQ-10 total score</li> <li>○ Change in FOSQ (30 items) Vigilance domain score</li> <li>○ Change in FOSQ (30 items) Activity Level domain score</li> </ul> </li> <li>• Percent of participants with <math>\geq 50\%</math> AHI reduction</li> <li>• Percent of participants with                             <ul style="list-style-type: none"> <li>○ AHI <math>&lt; 5</math> or</li> <li>○ AHI 5-14 with ESS <math>\leq 10</math></li> </ul> </li> <li>• Percent change in body weight</li> <li>• Change in hsCRP concentration</li> </ul> <p>From baseline to Week 48<sup>a</sup></p> <ul style="list-style-type: none"> <li>• Change in SBP</li> </ul>

Other Secondary Objectives	Endpoints
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in excessive daytime sleepiness</li> <li>• Change in patient-reported functional status as assessed by FOSQ (30 items)</li>   <li>• Change in body weight</li>   <li>• Change in lipid parameters</li>   <li>• Change in PROs</li>   <li>• Insulin</li> <li>• Hypoxic burden</li> <li>• Change in DBP</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in ESS score</li> <li>• Change in all FOSQ domain scores, specifically               <ul style="list-style-type: none"> <li>○ General Productivity</li> <li>○ Activity level</li> <li>○ Vigilance</li> <li>○ Social outcomes</li> <li>○ Intimate and sexual relationships</li> </ul> </li> <li>• Percent of participants who achieve               <ul style="list-style-type: none"> <li>○ ≥10% body weight reduction</li> <li>○ ≥15% body weight reduction</li> <li>○ ≥20% body weight reduction</li> </ul> </li> <li>• Change in               <ul style="list-style-type: none"> <li>○ HDL-cholesterol</li> <li>○ non-HDL-cholesterol</li> <li>○ triglycerides</li> </ul> </li> <li>• Change in:               <ul style="list-style-type: none"> <li>○ PROMIS Sleep-related impairment short form 8a</li> <li>○ PROMIS Sleep disturbance short form 8b</li> <li>○ SF-36v2 acute form domain and summary scores</li> </ul> </li> <li>• Percent of participants with improved categorical shift in:               <ul style="list-style-type: none"> <li>○ PGIS-OSA Sleepiness</li> <li>○ PGIS-OSA Fatigue</li> <li>○ PGIS-OSA Snoring</li> </ul> </li> <li>• Change in fasting insulin</li> <li>• Change in SASHB (% min/hour)</li> </ul> <p>From baseline to Week 48<sup>a</sup></p> <ul style="list-style-type: none"> <li>• Change in DBP</li> </ul>



Exploratory Objectives	Endpoints
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in PROs</li> </ul> <ul style="list-style-type: none"> <li>• To evaluate the effect of tirzepatide on sleep parameters as measured by Actigraphy (AX6)</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in                             <ul style="list-style-type: none"> <li>○ EQ-5D-5L utility index</li> <li>○ EQ-VAS scores</li> </ul> </li> <li>• Percent of participants with improved categorical shift in:                             <ul style="list-style-type: none"> <li>○ PGIC-OSA Sleepiness</li> <li>○ PGIC-OSA Fatigue</li> <li>○ PGIC-OSA Sleep quality</li> <li>○ PGIC-OSA Snoring</li> </ul> </li> <li>• Mean change from baseline to endpoint assessment in                             <ul style="list-style-type: none"> <li>○ Daytime sleep duration</li> <li>○ Daily step counts</li> <li>○ Average acceleration</li> </ul> </li> </ul>

Abbreviations: AHI = Apnea-Hypopnea Index; AX6 = Axivity 6; DBP = diastolic blood pressure; BP = blood pressure; ESS = Epworth Sleepiness Scale; EQ-5D-5L = EuroQol — 5 Dimension — 5 Level; EQ-VAS = EuroQol Visual Analogue Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; HDL = high-density lipoprotein; hsCRP = high-sensitivity C reactive protein; MTD = maximum-tolerated dose; OSA = obstructive sleep apnea; PAP = positive airway pressure; PGIC-OSA = Patient Global Impression of Change – Obstructive Sleep Apnea; PGIS-OSA = Patient Global Impression of Status – Obstructive Sleep Apnea; PROs = patient-reported outcomes; PROMIS = Patient-Reported Outcomes Measurement Information System; QW = once weekly; SBP = systolic blood pressure; SF-36v2 = Short-Form 36 version 2; SASHB = sleep apnea-specific hypoxic burden.

<sup>A</sup> BP will be assessed at Week 48 because PAP withdrawal at Week 52 may confound BP assessment.

### 1.1.1. Estimands

#### Primary estimands

The primary and key secondary efficacy analysis will be guided by 2 estimands, the “*treatment regimen*” estimand and the “*efficacy*” estimand to support global regulatory submissions and publications. The “*efficacy*” estimand provides an on-treatment assessment of efficacy without confounding the treatment effect from the data collected following treatment discontinuation. It represents on-treatment efficacy. The “*treatment-regimen*” estimand estimates the treatment effect, including the effect of intervention discontinuation to reflect clinical practice. It represents the efficacy irrespective of adherence to study intervention.

#### Efficacy estimand

The clinical question of interest for the efficacy estimand is the treatment difference between tirzepatide and placebo after 52 weeks of intervention in patients with obesity and OSA, prior to intervention discontinuation for any reason, excluding rare cases of patients who have discontinued study intervention due to inadvertent enrollment.

**Efficacy estimand attributes**

- *Population:* Adult patients with obesity and OSA who qualify per the inclusion and exclusion criteria in the protocol and received at least 1 dose of study treatment.
- *Treatment condition:* On randomized treatment.
- *Endpoints:* The primary and key secondary endpoints will be studied. Further details on the endpoints can be found in the Objectives and Endpoints table.
- *Population level summary:* The difference in mean change from baseline to 52 weeks will be used for continuous endpoints; the difference in proportion (absolute or relative, as appropriate) will be used for dichotomous endpoints. The population level summary will be conducted using the EAS described in Section 3.
- *Intercurrent events:* The intercurrent event of treatment discontinuation or missing data due to a technical issue (for example, sensor malfunction during PSG) or scheduling problem is addressed by the treatment condition attribute.
- *Rationale:* The efficacy estimand provides an on-treatment assessment without confounding the treatment effect from off-treatment data.

**Treatment-regimen estimand**

The clinical question of interest for the treatment-regimen estimand is the treatment difference between tirzepatide and placebo after 52 weeks of intervention in patients with obesity and OSA, regardless of intervention discontinuation for any reason (excluding rare cases of patients who have discontinued study intervention due to inadvertent enrollment).

**Treatment-regimen estimand attributes**

- *Population:* Adult patients with obesity and OSA who meet the inclusion and exclusion criteria and received at least 1 dose of study treatment.
- *Treatment condition:* On- or off-randomized-treatment.
- *Endpoints:* The primary and key secondary endpoints will be studied. Further details on the endpoints are in the Objectives and Endpoints table (Section 1.1).
- *Population level summary:* The difference in mean change from baseline to 52 weeks will be used for continuous endpoints and the difference in proportion (absolute or relative, as appropriate) will be used for dichotomous endpoints. The population level summary will be conducted using the FAS described in Section 3.
- *Intercurrent events:* Approaches for handling intercurrent events “missing data due to a technical issue (for example, sensor malfunction during PSG) or scheduling problem,” “study discontinuation due to the COVID-19 pandemic,” and “all other study discontinuations” will be described in Section 4.1.2.
- *Rationale:* The treatment-regimen estimand estimates treatment effect, including the effect of intervention discontinuation to reflect clinical practice. It is used for submission and registration purpose with regulatory agencies.

Efficacy and treatment-regimen estimands will be evaluated for key secondary objectives similarly to the primary objectives.

### **Safety estimand**

The clinical interest for safety estimands is the safety assessment of individual treatment arms up to Week 52 and up to the end of the study in patients with obesity and OSA, from all randomly assigned participants who are exposed to at least 1 dose of study intervention, regardless of adherence to study intervention.

### **Safety estimand attributes**

- *Population*: Adult patients with obesity and OSA who meet the inclusion and exclusion criteria and received at least 1 dose of study treatment.
- *Treatment condition*: On- or off-randomized-treatment.
- *Endpoints*: Endpoints corresponding to the safety analyses described in Section 4.6.
- *Population level summary*: Population level summaries will be conducted using the Safety Analysis Set described in Section 3.
- *Intercurrent events*: Potential intercurrent events include study discontinuation or data missing due to a technical or scheduling issue but there are no planned approaches for accommodating intercurrent events.

## **1.2. Study Design**

Study I8F-MC-GPIF (GPIF) is a multicenter, randomized, parallel-arm, double-blind, placebo-controlled Phase 3 study to evaluate the efficacy and safety of tirzepatide at the MTD (10 mg or 15 mg) once weekly versus placebo in participants who have obesity and moderate-to-severe OSA.

This basket-type master protocol will investigate 2 participant populations, described in 2 ISAs:

- GPI1 will include participants who are unwilling or are unable to use PAP therapy.
- GPI2 will include participants who have been on PAP therapy for at least 3 consecutive months prior to Visit 1 and plan to continue PAP therapy during the study.

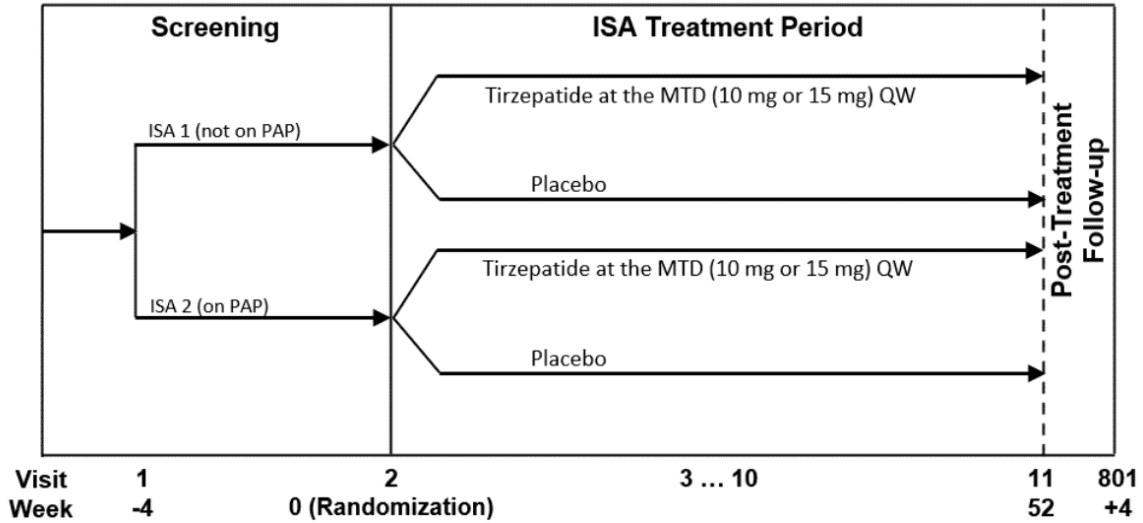
Participants to be assigned to whichever ISA they qualify for. Participants will then be randomly assigned to:

- tirzepatide at the MTD (10 or 15 mg) subcutaneous once weekly, or
- placebo.

The expected total duration of study participation for each participant, including screening and the posttreatment follow-up periods, is 60 weeks across the following study periods:

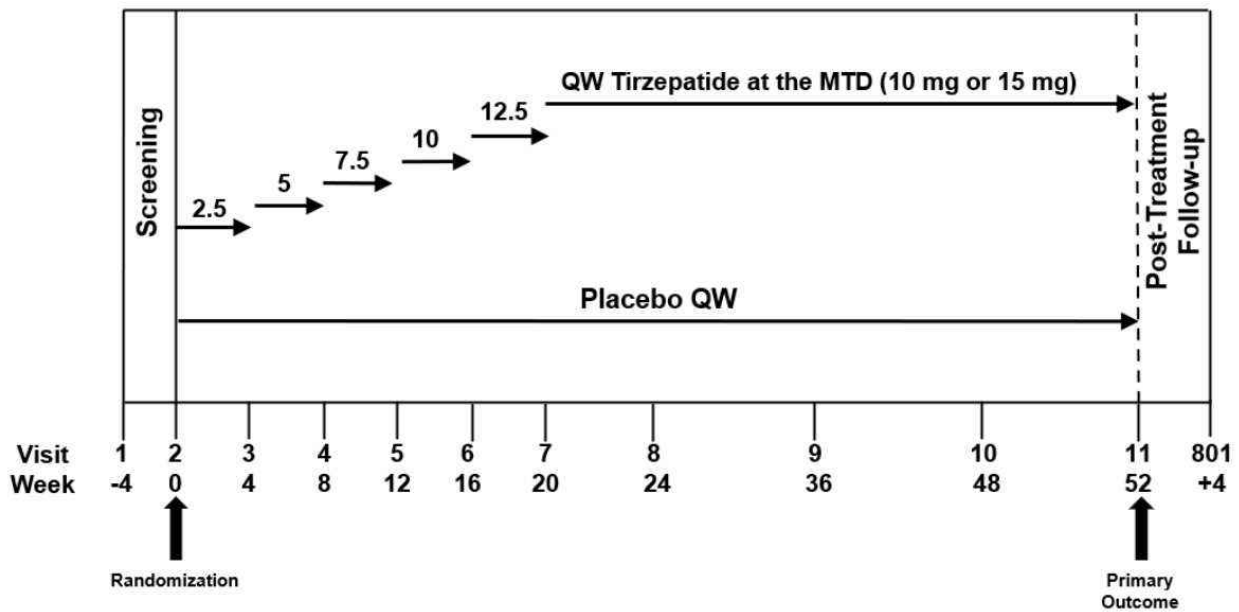
- Screening: 4 weeks
- ISA Treatment Period: 52 weeks
- Post-Treatment Follow-up Period: 4 weeks

The maximum duration of treatment is 52 weeks.



Abbreviations: ISA = intervention-specific appendix; MTD = maximum tolerated dose; PAP = positive airway pressure; QW = once weekly.

**Figure GPIF.1.1. Illustration of master protocol design for clinical protocol I8F-MC-GPIF.**



Abbreviations: MTD = maximum tolerated dose; QW = weekly.

**Figure GPIF.1.2. Illustration of dose escalation and visit schema for clinical protocol I8F-MC-GPIF.**

## 2. Statistical Hypotheses

For each ISA, the primary objective is to demonstrate that tirzepatide at the MTD (10 mg or 15 mg) is superior to placebo in treating participants with OSA with respect to the percent change in AHI. Thus, the null and alternative hypotheses will be defined as below.

Null hypothesis: tirzepatide at the MTD (10 mg or 15 mg) is not different from the placebo with respect to the mean percent change from baseline in AHI at 52 weeks.

Alternative hypothesis: tirzepatide at the MTD (10 mg or 15 mg) is superior to the placebo with respect to the mean percent change from baseline in AHI at 52 weeks.

The treatment effect will be defined as the difference between the estimates of the mean percent change from baseline at 52 weeks for tirzepatide at the MTD (10 mg or 15 mg) and placebo. Operationally, the hypotheses will be evaluated by 2-sided tests.

### 2.1. Multiplicity Adjustment

The type 1 error rate control strategy for primary and key secondary objectives is discussed in Section 4.4.2. There will be no multiplicity adjustments for evaluating exploratory objectives and safety assessments. Multiplicity adjustment to maintain the type 1 error rate of 0.05 will be used for the primary and key secondary objectives within each ISA and within each estimand (treatment regimen and efficacy estimands).

### 3. Analysis Sets

Table GPIF.3.1 describes the populations that will be used for statistical analyses within each ISA of the master protocol. Additional intervention-specific populations for analyses may be described in the respective ISA.

**Table GPIF.3.1. Description of Analysis Datasets**

<b>Analysis Set or Population</b>	<b>Description</b>
Modified Intent-to-Treat (mITT)	All randomly assigned participants who are exposed to at least 1 dose of study intervention.
Full Analysis Set (FAS)	Data obtained during treatment period from the mITT population excluding those discontinuing study intervention due to inadvertent enrollment, regardless of adherence to study intervention
Efficacy Analysis Set (EAS)	Data obtained during treatment period from the mITT population excluding those discontinuing study intervention due to inadvertent enrollment, excluding data after discontinuation of study intervention (last dose + 7 days)
Safety Analysis Set (SS)	Data obtained during treatment period plus safety follow-up period from mITT population, regardless of adherence to study intervention

## 4. Statistical Analyses

### 4.1. General Considerations

Statistical analysis will be the responsibility of Eli Lilly and Company (Lilly) or its designee. Statistical analysis for each ISA will be conducted individually and a combined analysis with both ISAs is not planned. All analyses specified will apply to both ISAs unless the analysis is specified as ISA-specific.

The SAP will be finalized prior to the unblinding of the first ISA.

Some analyses and summaries described in this analysis plan may not be conducted if not warranted by data (for example, too few events to justify conducting an analysis). Additional exploratory analyses of the data will be conducted as deemed appropriate.

Efficacy analyses will be conducted on all participants randomly assigned to study intervention according to the treatment to which the participants are assigned and were exposed to at least one dose, excluding those discontinuing study intervention due to inadvertent enrollment. For the “treatment regimen” estimand, the analysis will be conducted using FAS. To minimize missing data, participants randomly assigned to study intervention who prematurely discontinue study treatment will be encouraged to remain in the study. However, some participants may choose to permanently discontinue from the study which will lead to missing endpoints. Details on handling missing values can be found in Section 4.1.2. For the “efficacy” estimand, the analysis will be conducted using the EAS.

Safety analysis will be conducted using the Safety Analysis Set. Selected safety analyses may be conducted after excluding the data after permanent discontinuation of the study intervention. For the safety related parameters, the definition of baseline and postbaseline are specified in [Table GPIF.4.1](#).

**Table GPIF.4.1. Baseline and Postbaseline Definition for Safety Analyses**

Analysis Set	Analysis Type	Baseline	Postbaseline
SS	1.1) Treatment-Emergent Adverse Events	The baseline period is defined as the start of screening and ends prior to the first dose of study treatment (typically at Week 0).	Starts after the first dose of study treatment and ends at the end of the study period (including off-drug follow up visit).
SS	1.2) Treatment-Emergent Abnormal Labs <sup>a</sup> and Vital Signs.	For labs, baseline period is defined as prior to the first dose time and will include all scheduled and unscheduled measurements. For vital signs, baseline period is defined as measurements collected prior to the first dose.	Postbaseline will be defined as after the baseline period through the end of the study participation. All scheduled and unscheduled measurements will be included.

Analysis Set	Analysis Type	Baseline	Postbaseline
SS	1.3) Change from Baseline for Labs <sup>a</sup> , and Vital Signs.	The last scheduled and unscheduled non-missing assessment recorded during the baseline period defined above (1.2).	Postbaseline will be defined as above (1.2). Only scheduled visits will be included. The ED visits are considered scheduled visits.

Abbreviations: ED = early discontinuation; SS = Safety Analysis Set.

<sup>a</sup> Immunogenicity related analysis is specified in Section 4.6.3.5.

The following paragraphs define selection of the PRO response which will be used for analysis at baseline and postbaseline visits. In order to select the baseline observation for PROs which are planned to be completed on the same day as the PSG (ESS, FOSQ, PROMIS, PGIS, SF-36v2 acute form, and EQ-5D-5L), we will select the observation completed within 24 hours of the start of the baseline PSG. If multiple responses are completed within 24 hours of the start of the PSG, we will select the last response given within 24 hours of the start of the PSG. If no response was provided within 24 hours of the start of the PSG, we will select the latest observation completed prior to the first dose.

For postbaseline visits with a planned PSG measurement, we will select the response for the PROs which are planned to be completed on the same day as the PSG (ESS, FOSQ, PROMIS, PGIS, PGIC, SF-36v2 acute form, and EQ-5D-5L) in the following way. If available, select the response completed within 24 hours of the start of the PSG. If multiple responses are completed within 24 hours of the start of the PSG, we will select the most recently completed response. If no response was provided within 24 hours of the start of the PSG, we will select the latest observation completed within the visit window.

For postbaseline visits without a planned PSG measurement, we will select the response for the PROs which are planned to be completed on the same day as the PSG (ESS, FOSQ, PROMIS, PGIS, PGIC, SF-36v2 acute form, and EQ-5D-5L) by selecting the latest observation completed within the visit window.

In order to select the baseline observation for the PHQ-9 and C-SSRS, if multiple responses are completed prior to the first dose and there are no differences in these responses, we will select the observation completed most recently prior to the first dose. If multiple responses are completed prior to the first dose and there are differences in these responses, the approach differs based on the questionnaire. For the PHQ-9, select the response with the worst total score, and for the C-SSRS, select the worst response for each question and combine each of the worst responses into a single response which will be used for analysis. For each postbaseline visit, we will carry out the same approach to select a response in the case of multiple responses within a visit window.

For all other analyses, baseline is defined as the last nonmissing measurement prior to the first dose unless otherwise specified.



Statistical treatment comparisons will be performed between tirzepatide MTD and placebo. Unless otherwise noted, all tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, and the confidence interval will be calculated at a 2-sided 95%, level. In statistical summaries and analyses, participants will be analyzed as randomized. Countries in similar geographic regions with fewer than 10 participants, based on the all-randomized population within an ISA, will be pooled to achieve a pooled country of at least 10 participants. All analyses using country in the model will use a pooled country, unless otherwise specified. The final pooling by country and geographic region will be finalized prior to data lock.

Analysis of covariance will be used to analyze continuous variables collected only at baseline and endpoint. Unless otherwise specified, the model will include treatment and strata (pooled country/geographic region, AHI stratum [moderate ( $AHI \geq 15$  and  $AHI < 30$ ), severe ( $AHI \geq 30$ )] and gender) as fixed effects and baseline as a covariate.

MMRMs will be used to analyze continuous variables collected at baseline and more than one postbaseline visit. For the MMRM analysis, restricted-maximum-likelihood will be used to obtain model parameter estimates for continuous longitudinal variables. All the longitudinal observations at each scheduled postbaseline visit will be included in the analysis. The model will include the fixed class effects of treatment, strata (pooled country/geographic region and gender), visit, and treatment-by-visit interaction, as well as the continuous, fixed covariate of baseline value. For analyses of variables other than AHI, the AHI stratum will also be included in the model. Significance tests will be based on least-squares means and Type III tests.

For continuous measures, summary statistics may include sample size, mean, standard deviation, median, minimum, and maximum for both the actual and the change from baseline measurements. Least-square means and standard errors derived from the analysis models will also be displayed for the change from baseline measurements. Treatment comparisons will be displayed showing the treatment difference least-square means and the 95% confidence intervals for the treatment differences, along with the p-values for the treatment comparisons.

For categorical measures, summary statistics may include sample size, frequency, and percentages. Fisher's exact test or Pearson's chi-square test will be used for treatment comparisons unless otherwise specified.

Not all analyses described in this SAP will necessarily be included in the CSRs. Any analysis described in this SAP and not provided in the CSR would be available upon request.

#### **4.1.1. Adjustment for Covariates**

The study is stratified by country, OSA severity (moderate [ $AHI \geq 15$  and  $AHI < 30$ ], severe [ $AHI \geq 30$ ]), and gender. Unless otherwise specified, the stratification factors will be adjusted for in efficacy analyses. The value for stratification factors will be obtained from the data collected or derived from the eCRF or PSG results. In addition, the baseline value of the endpoint will be used as a covariate when appropriate.

### 4.1.2. Handling of Dropouts or Missing Data

For the primary and key secondary efficacy endpoint analyses aligned to the treatment regimen estimand and subject to type 1 error rate control, missing data will be imputed based on the reason for the intercurrent event, as described in [Table GPIF.4.2](#) below. For analyses aligned to the “efficacy” estimand, missing data will be considered missing at random and hence no explicit imputation will be performed.

For exploratory endpoints and safety analyses, missing values will not be explicitly imputed unless specified otherwise.

For analyses aligned to the treatment regimen estimand, the statistical inference over multiple imputations will be guided by the method proposed by Rubin (1987). The intercurrent events and the resulting missing values will be handled as follows:

**Table GPIF.4.2. Imputation Approaches to Handle Intercurrent Events for Treatment Regimen Estimand**

Intercurrent events	Strategy to handle intercurrent events	Assumptions for missing values	Methods to handle missing values
Study DC due to the COVID-19 pandemic (after other reasons for missing data are ruled out), or missing data due to a technical issue (that is, sensor error on PSG) or scheduling problem	Hypothetical	MAR	Use observed data without any imputation
All other study discontinuations	Treatment policy	MNAR	Retrieved dropout imputation <sup>a</sup> . If there are not enough retrieved dropouts to provide a reliable imputation model, placebo-based multiple imputation will be used.

Abbreviations: COVID-19 = coronavirus disease-2019; DC = discontinuation; MAR = missing at random; MNAR = missing not at random; PSG = polysomnography; SBP = systolic blood pressure.

<sup>a</sup> Retrieved dropout imputation utilizes observed data from participants in the same treatment group who had outcome measures at Week 52 (or Week 48 for SBP) after early discontinuation of study drug to impute the missing value.

### 4.1.3. Multicenter Studies

Randomization will be stratified by country, and country/pooled country will be used as a covariate.

#### **4.1.4. Historical Illnesses and Preexisting Conditions**

The count and percentages of participants with historical illnesses and preexisting conditions will be summarized by treatment group using the MedDRA PTs nested within SOC. The SOC will be in alphabetical order. Conditions (that is, PTs) will be ordered by decreasing frequency in the TZP MTD arm within the SOC. This will be summarized for all randomized participants.

#### **4.1.5. Patient Characteristics**

A listing of participant demographics for all randomized participants will be provided. The demographic and baseline clinical characteristics will also be summarized by study treatment for all randomized participants. Baseline demographic and clinical characteristics of special interest include but are not limited to: age (years), sex (female, male), race, ethnicity, height (cm), weight (kg), BMI (kg/m<sup>2</sup>), waist circumference (cm), age group (<50, ≥50), BMI group (<35, ≥35 and <40, ≥40 kg/m<sup>2</sup>), OSA severity (moderate [AHI ≥15 and AHI < 30] versus severe [AHI ≥30]), region, and country.

#### **4.1.6. Concomitant Therapy**

Concomitant medication will be summarized by treatment groups and displayed by decreasing frequency of WHODrug PTs in TZP MTD arm.

In addition, medications of interest (as defined below) will be summarized by treatment groups:

- baseline use of:
  - lipid lowering therapy, by type/class and
  - antihypertensive therapy, by type/class
- changes to baseline medication in postrandomization (in term of type/class):
  - lipid lowering therapy and
  - Antihypertensive therapy.
- Utilization after randomization of
  - antihyperglycemic medication for the treatment of diabetes for participants who develop T2DM during the study
  - antidiarrheal medication, and
  - antiemetic medication

In addition, for GPI2 patients only, a summary of PAP machine use at baseline and postbaseline PAP machine adherence will be provided. Further details are provided in Section 4.7.1. For GPI1 patients, a listing will be provided summarizing any patients who use a PAP machine during the course of the trial.

#### **4.1.7. Treatment Exposure and Compliance**

##### **4.1.7.1. Study and Study Treatment Exposure**

Summary of duration of follow-up (defined as time in days from date of randomization to the date of the last study visit) will be provided by treatment group in the mITT population.

Summary of duration on study treatment (defined as time in days from date of first dose of study

treatment to date of last dose of study treatment plus 7 days) will be provided by treatment group in the safety analysis set.

For the summary of duration on study treatment, the frequency and percentage of participants falling into the following categorical ranges will also be summarized by planned treatment group as well: >0 week,  $\geq 4$  weeks,  $\geq 8$  weeks,  $\geq 12$  weeks,  $\geq 16$  weeks,  $\geq 20$  weeks,  $\geq 24$  weeks,  $\geq 36$  weeks,  $\geq 48$  weeks, and  $\geq 52$  weeks.

No p-values will be reported in these summaries as they are intended to describe the study populations rather than test hypotheses.

#### **4.1.7.2. Adherence to Study Treatment**

Summary of prematurely discontinuing study treatment (including reason for discontinuation) will be provided by study treatment. A time-to-event analysis of premature study treatment discontinuation will also be conducted.

If data warrants, the counts and percentages of participants who follow the planned escalation scheme, have dose interruption, or have dose de-escalation will be summarized for the tirzepatide treatment group. This will include the percentage of patients who have 10 mg or 15 mg tirzepatide as their MTD. In addition, the proportion of participants receiving 2.5, 5, 7.5, 10, 12.5, or 15 mg may be presented by randomized tirzepatide treatment and visit during the dose escalation period.

Treatment adherence will be defined as taking at least 75% of the scheduled tirzepatide doses. Adherence (%) at each treatment visit and over the whole treatment period will be calculated using the number of doses administered (regardless of the actual dose in mg administered) divided by the total number of doses expected to be administered  $\times 100$  at the specific visit or over the treatment period, respectively. Treatment adherence will be summarized descriptively at each treatment visit and over the treatment period by treatment using mITT population.

#### **4.1.8. Important Protocol Deviations**

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

### **4.2. Participant Dispositions**

Summaries and a listing of study disposition and study drug disposition will be provided for all randomized participants, separately for each ISA. Comparison between treatment arms will be performed using Fisher's exact test.

### **4.3. Primary Endpoint Analysis**

The primary objective of this study is to test the hypothesis that tirzepatide at the MTD (10 or 15 mg) is superior to placebo for participants with moderate to severe OSA on the mean percent AHI reduction from baseline to Week 52. The primary and key secondary efficacy analyses will be guided by 2 estimands, the "treatment regimen" estimand and the "efficacy" estimand to support global regulatory submissions and publications.

### **4.3.1. Analysis Related to the Efficacy Estimand**

The primary analysis guided by the “efficacy” estimand will be conducted using the EAS. This analysis will be based on the contrast between tirzepatide at the MTD (10 or 15 mg) and placebo at Week 52 (Visit 11) from the MMRM analysis of mean percent change from baseline in AHI. All the longitudinal observations at each scheduled postbaseline visit will be included in the analysis. REML will be used to obtain model parameter estimates and the Kenward-Roger option will be used to estimate the denominator degrees of freedom. The response variable of the MMRM will be the percent change in AHI from baseline values obtained at each scheduled postbaseline AHI measurement.

The model will include the fixed class effects of treatment, strata (pooled country/geographic region and gender), visit, and treatment-by-visit interaction, as well as the continuous, fixed covariate of baseline AHI. An unstructured covariance structure will be used to model the within-patient errors. Significance tests will be based on least squares means and Type III tests. If this analysis fails to converge, the following covariance structures will be tested in order until convergence is achieved:

- Toeplitz with heterogeneity
- autoregressive with heterogeneity
- compound symmetry with heterogeneous variances
- Toeplitz
- autoregressive, and
- compound symmetry without heterogeneous variances.

Analysis aligned to each estimand will be evaluated at the full significance level of 0.05.

### **4.3.2. Analysis Related to the Treatment Regimen Estimand**

For the primary analysis guided by the “treatment regimen” estimand, the analysis will be conducted using the FAS. Missing values will be imputed based on the strategy to handle intercurrent events described in Section 4.1.2. Following imputation, the primary efficacy comparison will be based on the contrast between tirzepatide at the MTD (10 or 15 mg) and placebo from the ANCOVA analysis of mean percent change from baseline to Week 52 in AHI using FAS. The ANCOVA model will include treatment and strata (pooled country/geographical region and gender) as fixed effects and baseline AHI as a fixed covariate. Statistical inference over multiple imputed data sets will be guided by Rubin (1987).

### **4.3.3. Sensitivity Analyses**

For patients in GPI2, we will carry out a sensitivity analysis for the primary endpoint. When carrying out this sensitivity analysis using a treatment regimen estimand, patients with PAP withdrawal less than 5 days before the PSG at baseline or at Week 52 will have their data censored. Censored postbaseline data will be imputed using the approach outlined in Section 4.1.2. An ANCOVA model will be fit using the approaches outlined in Section 4.3.2. Carrying out this sensitivity analysis using the efficacy estimand, we will censor observations made with PAP withdrawal less than 5 days prior to PSG from the MMRM. We may consider

additional sensitivity analyses for GPI2 patients to accommodate patients with PAP withdrawal <5 days prior to the PSG.

Additional sensitivity analyses for both ISAs may be included as needed.

## 4.4. Secondary Endpoints Analysis

### 4.4.1. Key Secondary Endpoints

A graphical approach for multiple comparisons will be used to strongly control the overall Type I error (2-sided alpha level of 0.05) for testing the superior treatment effect of tirzepatide MTD over placebo including the key secondary endpoints as listed below.

- a hierarchical composite of the following (additional details provided in Section 4.4.1.1):
  - change from baseline to Week 52 in FOSQ-10 score
  - change from baseline to Week 52 in FOSQ (30 items) Vigilance domain score
  - change from baseline to Week 52 in FOSQ (30 items) Activity Level domain score
- percent of participants with  $\geq 50\%$  AHI reduction at Week 52
- percent of participants at Week 52 with
  - AHI <5 or
  - (AHI 5 through 14 and ESS  $\leq 10$ )
- percent change from baseline to Week 52 in body weight
- change from baseline to Week 48 in SBP
- change from baseline to Week 52 in C-reactive protein (high-sensitivity C reactive protein)

Analytical approaches for the hierarchical assessment of PROs are described in Section 4.4.1.1 and a summary of the analysis approach for all other key secondary endpoints is provided in Section 4.4.1.2.

#### 4.4.1.1. Hierarchical Assessment of PROs

The analysis of the hierarchical composite endpoint will be performed with the Finkelstein-Schoenfeld method, and the win ratio (Pocock et al. 2012) will be reported as the measure of treatment effect. The population-level summary of win ratio will be calculated as number of pairs of tirzepatide-treated participant “wins” divided by number of pairs of placebo treated participant “wins.”

The Finkelstein-Schoenfeld method is based on the principle that each tirzepatide-treated participant is compared with every other placebo-treated participant in a pairwise manner that proceeds in a hierarchical fashion. Differences will be calculated as tirzepatide patient value minus placebo patient value. Each pairwise comparison will proceed in the following order and a winner has:

- Stage 1: For the change from baseline at Week 52 FOSQ 10 score:
  - A comparison is a win when the treatment difference is  $\geq 2.2$

- A comparison is a loss when the treatment difference is  $\leq -2.2$
- All other cases are a tie and the comparison of FOSQ vigilance score will be conducted (that is, proceed to Stage 2).
- Stage 2: For the change from baseline at Week 52 FOSQ vigilance score:
  - A comparison is a win when the treatment difference is  $\geq 0.44$ .
  - A comparison is a loss when the treatment difference is  $\leq -0.44$ .
  - All other cases are a tie and the comparison of FOSQ vigilance score will be conducted (that is, proceed to Stage 3).
- Stage 3: For the change from baseline at week 52 FOSQ activity level score:
  - A comparison is a win when the treatment difference is  $\geq 0.44$ .
  - A comparison is a loss when the treatment difference is  $\leq -0.44$ .
  - In all other cases, the pair will be recorded as a tie.

Based on Weaver and colleagues (2021), a clinically important response of FOSQ 10 for OSA patients is 2.2. FOSQ 10 is calculated by adding the average of the responses within each domain from 5 domains. Because a clinically meaningful response is unclear for individual domains, equal weights were given to define a criterion for a win for the individual domains resulting in 0.44 ( $=2.2/5$ ). The scoring definition for domain specific and total score is identical between FOSQ-10 questionnaire and FOSQ 30 questionnaire, so domain specific clinically important response from FOSQ 10 will be carried over for FOSQ 30.

For treatment policy estimand, missing values at Week 52 will be imputed through multiple imputations based on the reason of missingness with details described in Section 4.1.2.

#### 4.4.1.2. Main Analytical Approaches

Analysis of change in AHI, percent change from baseline to Week 52 in body weight, change from baseline to Week 52 in high-sensitivity C reactive protein, and change from baseline to Week 48 in SBP will be conducted in a manner similar to the primary efficacy analyses using ANCOVA model with treatment, strata (pooled country/geographic region, AHI stratum [moderate ( $AHI \geq 15$  and  $AHI < 30$ ), severe ( $AHI \geq 30$ )] and gender) and baseline of the corresponding variable as a covariate for the treatment regimen estimand. For the efficacy estimand, the MMRM analyses will be conducted as described in Section 4.1. For both estimands, analysis of change in AHI will adjust for the continuous, fixed baseline value of AHI instead of the baseline AHI stratum (moderate, severe).

Comparisons at the 52-week visit between the treatments relative to the proportion of patients achieving  $\geq 50\%$  AHI reduction and  $AHI < 5$  or ( $AHI 5$  through  $14$  and  $ESS \leq 10$ ) will be conducted using logistic regression analysis including terms: treatment, pooled country, baseline AHI, and gender as a covariate. For the treatment regimen estimand, missing value at the endpoint will be imputed as not achieving the target.

Analysis aligned to each estimand will be evaluated at the full significance level of 0.05 contingent on reaching statistical significance of the primary objective.



#### **4.4.1.3. Sensitivity Analyses**

For patients in GPI2, we will carry out sensitivity analyses for the key secondary endpoints absolute change in AHI, the hierarchical PRO assessment, clinically meaningful change in AHI, and achieving OSA remission/mild nonsymptomatic OSA.

For absolute change in AHI and the hierarchical PRO assessment, we will carry out the sensitivity analysis using both the treatment regimen and efficacy estimand. When using a treatment regimen estimand approach, patients with PAP withdrawal less than 5 days before the PSG or FOSQ at baseline or at Week 52 will have their data censored. Censored postbaseline data will be imputed using the approach outlined in Section 4.1.2. An ANCOVA model will be fit using the approaches outlined in Section 4.3.2. When using the efficacy estimand, we will censor observations made with a PAP withdrawal less than 5 days prior to the PSG or FOSQ from the MMRM. Of note, when using the efficacy estimand, the Week 12 FOSQ data will have no corresponding PAP withdrawal and will not be censored but some FOSQ data may be censored at Week 20 under the efficacy estimand for this sensitivity analysis.

For the binary endpoints of clinically meaningful change in AHI and achievement of OSA remission/mild nonsymptomatic OSA, we will also carry out the sensitivity analysis using both estimands. PSG measurements taken after <5 days of PAP withdrawal will be censored. Following censoring, we will carry out analysis as described in Section 4.4.1.2.

We may consider additional sensitivity analyses for GPI2 patients to accommodate patients with PAP withdrawal <5 days prior to the PSG or FOSQ or ESS.

Additional sensitivity analyses for both ISAs may be included as needed.

#### **4.4.2. Type I Error Rate Control Strategy for Primary and Key Secondary Efficacy Analyses**

All primary and key secondary hypotheses will be tested with the overall family-wise type I error rate at a 2-sided alpha level of 0.05 through the multiplicity control approach based on the graphical multiple testing procedure. The primary endpoint hypothesis will be tested at a 2-sided alpha level of 0.05 for statistical significance. If the primary efficacy endpoint is significant, the alpha of 0.05 will be propagated to the key secondary efficacy endpoints. The detailed graphical testing scheme will be outlined in a later version of the SAP.

The analyses will be performed for both the treatment regimen and efficacy estimands described in Section 4.3 using the same graphical testing scheme. We will use an overall 2-sided alpha of 0.05 to control type 1 error rate separately for the treatment regimen estimand and the efficacy estimand.

#### **4.4.3. Supportive Secondary Endpoints**

Unless otherwise specified, all supportive/other secondary efficacy analyses will be guided by the “efficacy” estimand and will be conducted using the EAS. Missing data will be handled using an MMRM and multiple imputation will not be used.



**Table GPIF.4.3. Secondary Measures Not Controlled for Type 1 Error**

<b>Objective – Demonstrate superiority of tirzepatide MTD to placebo for:</b>	<b>Endpoint</b>	<b>Analytical Approaches</b>
Change in excessive daytime sleepiness	Change in ESS score from baseline to Week 52	MMRM analysis described in Section 4.4.1.2 will be conducted.
Change in patient-reported functional status as assessed by FOSQ (30 items)	Change in all functional domain scores from baseline to Week 52	MMRM analysis described in Section 4.4.1.2 will be conducted.
Change in Body Weight	Percent of participants who achieve $\geq 10\%$ , $\geq 15\%$ , and $\geq 20\%$ body weight reduction.	Logistic models described in Section 4.4.1.2 with the following covariates: treatment, pooled country, baseline AHI (moderate/severe), gender, and baseline bodyweight as a covariate.
Change in Lipid Parameters	Change in HDL-cholesterol non-HDL-cholesterol triglycerides	MMRM analysis described in Section 4.4.1.2 will be conducted.
Change in PROs	Change in: PROMIS Sleep-related impairment short form 8a score PROMIS Sleep disturbance short form 8b score SF-36v2 acute form domain scores From baseline to Week 52	MMRM analysis of T-score described in Section 4.4.1.2 will be conducted. Description of T-score calculation provided in Sections 4.7.2.4 and 4.7.2.5.
	Percent of participants with improved categorical shift in: PGIS-OSA Sleepiness PGIS-OSA Fatigue PGIS-OSA Snoring From baseline to Week 52	For each question, the proportion of participants with improvements from baseline will be summarized. Shift analysis from baseline to Week 52 will also be performed.
Change in Insulin	Change in fasting insulin from baseline to Week 52	MMRM analysis described in Section 4.4.1.2 will be conducted.
Change in Hypoxic Burden	Change in SASHB (% min/hour) from baseline to Week 52	MMRM analysis described in Section 4.4.1.2 will be conducted.
Change in DBP	Change in DBP from baseline to Week 48	MMRM analysis described in Section 4.4.1.2 will be conducted.

Abbreviations: AHI = Apnea-Hypopnea Index; DBP = diastolic blood pressure; ESS = Epworth Sleepiness Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; HDL = high-density lipoprotein; MMRM = mixed model repeated measures; MTD = maximum tolerated dose; PGIS-OSA = Patient Global Impression of Status Obstructive Sleep Apnea; PRO = Patient-Reported Outcomes; PROMIS = Patient-Reported Outcomes Measurement Information System; SASHB = sleep apnea specific hypoxic burden; SF-36v2 = Short-Form 36 version 2.

### 4.5. Exploratory Endpoint Analyses

Unless otherwise specified, all exploratory efficacy analyses will be guided by the “efficacy” estimand and will be conducted using the EAS. Missing data will be handled using an MMRM and multiple imputation will not be used.

The following efficacy analyses apply to both ISAs and will be carried out separately for each ISA.

**Table GPIF.4.4. Exploratory Efficacy Analysis for Both ISAs**

Objective – Demonstrate superiority of tirzepatide MTD to placebo for:	Endpoint	Analytical Approaches
Change in PROs	Change from baseline to Week 52 in EQ-5D-5L utility index EQ-VAS scores	MMRM analysis described in Section 4.4.1.2 will be conducted.
	Percent of participants with improved categorical shift from baseline to Week 52 in: PGIC-OSA Sleepiness PGIC-OSA Fatigue PGIC-OSA Sleep quality PGIC-OSA Snoring	For each question, the proportion of participants with improvements from baseline will be summarized. Shift analysis from baseline to Week 52 will also be performed.
Change in parameters measured by Actigraphy (AX6)	Mean change from baseline to week 52 in: Daytime sleep duration Daily step counts Average acceleration	MMRM analysis described in Section 4.4.1.2 will be conducted.

Abbreviations: EQ-5D-5L = EuroQoL-5 Dimension-5 Level; EQ-VAS = EuroQol Visual Analogue Scale; ISA = intervention-specific appendix; MMRM = Mixed model repeated measures; MTD = Maximum tolerated dose; PRO = Patient reported outcomes; PGIC-OSA = Patient Global Impression of Change Obstructive Sleep Apnea.

The efficacy analyses summarized in [Table GPIF.4.5](#) only apply to patients in GPI1.

**Table GPIF.4.5. Exploratory Efficacy Analysis Conducted only for GPI1 Patients**

Objective – Demonstrate superiority of tirzepatide MTD to placebo for:	Endpoint	Analytical Approaches
Change in parameters measured by WatchPAT300	Change from baseline to Week 52 in PAT-based device determinations of: <ul style="list-style-type: none"> <li>o pAHI</li> <li>o SASHB</li> </ul>	MMRM analysis described in Section 4.4.1.2 will be conducted.

Abbreviations: MMRM = mixed model repeated measures; MTD = maximum tolerated dose; pAHI = peripheral tone apnea-hypopnea index; PAT = peripheral arterial tonometry; SASHB = sleep apnea specific hypoxic burden.

## 4.6. Safety Analyses

Unless specified otherwise, safety assessments will be guided by an estimand comparing safety of tirzepatide MTD with placebo irrespective of adherence to study drug or initiation of rescue therapy. Thus, unless specified otherwise, safety analyses will be conducted in the safety analysis set ([Table GPIF.3.1](#)); all events that occur between the first dose date of study drug and the end date of study participation will be included, regardless of the adherence to study drug. Selected safety analyses may be conducted after excluding data on rescue therapy or data after starting another antihyperglycemic medication.

The statistical assessment of homogeneity of the distribution of categorical safety responses between tirzepatide MTD and placebo will be conducted using Fisher’s exact test, unless specified otherwise.

The mean change from baseline differences among treatments at all scheduled visits will be assessed via an MMRM using REML. The model will include treatment group, stratification factors, visit and treatment-by-visit interaction as fixed effects, and baseline value of the safety parameter as a covariate. To model the covariance structure within participants, the unstructured covariance matrix will be used. If this model fails to converge, the covariance structures specified in Section 4.3.1 will be tested in order until convergence is met. If the data does not warrant the MMRM model, then an ANCOVA model will be used.

For selected safety parameters, time-to-first-event analysis via the Cox-proportional hazards model may be conducted. Participants without the event will be censored at the end of study participation. For participants experiencing the event, the “time-to-first-event” will be the time (in days) from first dose to first occurrence of the event.

### 4.6.1. Analysis of Adverse Events

#### 4.6.1.1. Treatment Emergent Adverse Events

A TEAE is defined as an event that first occurred or worsened in severity after the first dose of study treatment. The MedDRA LLT will be used in the treatment-emergent derivation. The

maximum severity for each LLT during the baseline period including ongoing medical history will be used as baseline severity. For events with a missing severity during the baseline period, it will be treated as “mild” in severity for determining treatment-emergence. Events with a missing severity during the postbaseline period will be treated as “severe” and treatment-emergence will be determined by comparing to baseline severity.

For events occurring on the day of taking study medication for the first time, the case report form collected information (for example, treatment-emergent flag, start time of study treatment and event) will be used to determine whether the event was pre versus posttreatment if available. If the relevant information is not available, then the events will be counted as posttreatment.

Unless otherwise specified, the counts and percentages of participants with TEAEs will be summarized by treatment using MedDRA PT nested within SOC. Statistical comparisons will be applied at both the SOC and PT levels. Events will be ordered by decreasing frequency in the TZP arm within the SOC. The SOC will be in alphabetical order.

An overview of the number and percentage of participants who experienced a TEAE, SAE, death, discontinued from study treatment or study due to an AE, or with a TEAE related to study treatment will be summarized by treatment.

The counts and percentages of patients with TEAEs by maximum severity will be summarized by treatment using MedDRA PT. For each participant and TEAE, the maximum severity for the MedDRA PT is the maximum postbaseline severity observed from all associated LLTs mapping to the MedDRA PT. The maximum severity will be determined based on the nonmissing severities. If all severities are missing for the defined postbaseline period of interest, it will show as missing in the table.

For events that are gender specific, the denominator and computation of the percentage will only include participants of the given gender.

#### **4.6.1.2. Common Adverse Events**

The counts and percentages of participants with TEAEs, overall and common (common TEAEs occurred in  $\geq 5\%$  of participants before rounding), will be summarized by treatment using MedDRA PT. Events will be ordered by decreasing frequency in the TZP arm.

#### **4.6.1.3. Deaths**

A listing of all deaths during the study will be provided. The listing will include participant identification including the treatment, site number, date of death, age at the time of enrollment, sex, associated AE group identification, time from last dose of study drug to death (if participant had discontinued study drug), and primary cause of death.

#### **4.6.1.4. Other Serious Adverse Events**

The counts and percentages of participants who experienced an SAE (including deaths and SAEs temporally associated or preceding deaths) during the postbaseline period will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency in the TZP arm within the SOC. The SOC will be in alphabetical order.

A listing of all SAEs will be provided. The listing will include treatment, participant identification including the site number, date of event, age at the time of enrollment, sex, AE group identification, MedDRA SOC and PT, severity, outcome, relationship to study drug, time from first dose of study drug to the event, and time from most recent dose to event (if participant discontinued study drug prior to the event).

#### **4.6.1.5. Other Significant Adverse Events**

The counts and percentages of participants who discontinued from study treatment or study due to an AE during the postbaseline period may be summarized by treatment group using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency in the TZP arm within the SOC.

#### **4.6.2. Patient Narratives**

Patient narratives will be provided for all participants who experience any of the following “notable” events:

- death
- SAE
- pregnancy, or
- permanent discontinuation of study treatment due to AEs.

Patient narratives (patient level data and summary paragraph) will be provided for participants in the randomized population with at least 1 notable event.

#### **4.6.3. Special Safety Topics**

For AESI or special safety topics, the counts and percentages of participants will be summarized by treatment and PT with decreasing frequency in the TZP arm if the overall count is  $\geq 10$ . Individual participant level data may be presented. Displays with individual participant level data may be created using various formats, such as a customized listing and/or a customized graphical participant profile. AESI are defined in each section of special safety topics, where applicable.

##### **4.6.3.1. Exocrine Pancreas Safety**

###### **4.6.3.1.1. Pancreatic Enzyme**

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

The counts and percentages of participants with maximum postbaseline pancreatic enzyme value exceeding the following thresholds will be provided by baseline pancreatic enzyme value ( $\leq$ ULN,  $>$ ULN), and postbaseline:  $\leq 1 \times$  ULN,  $(>1$  to  $\leq 3) \times$  ULN,  $(>3$  to  $\leq 5) \times$  ULN,  $(>5$  to  $\leq 10) \times$  ULN,  $>10 \times$  ULN.

An MMRM analysis will be used to analyze each pancreatic enzyme with a log transformed (postbaseline measure/baseline measure) response variable and treatment, nominal visit, treatment-by-nominal visit interaction as fixed effects.

#### **4.6.3.1.2. Pancreatitis Events**

Summaries of adjudicated and investigator-reported pancreatic events will be provided by treatment. Detailed searching criteria can be found in Appendix 6.2.

Treatment-emergent adjudication confirmed pancreatitis will be considered as an AESI. Listing of participants with adjudicated pancreatitis may be provided if deemed necessary.

#### **4.6.3.2. Gastrointestinal Adverse Events**

##### **4.6.3.2.1. Nausea, Vomiting, and Diarrhea**

Summaries and analyses for incidence and severity of nausea, vomiting (including “vomiting” and “vomiting projectile”), diarrhea (including “diarrhea” and “diarrhoea”), and 3 events combined, will be provided by each treatment group.

Summary of the prevalence over time for nausea, vomiting, and diarrhea will also be presented. Time to the onset of nausea, vomiting, and diarrhea will be plotted.

##### **4.6.3.2.2. Severe Gastrointestinal Events**

The PTs under the *Gastrointestinal disorders* SOC in MedDRA will be used to identify GI AEs, and only the PTs with serious/severe treatment-emergent cases will be considered as AESIs.

The counts and percentages of participants with severe/serious treatment-emergent GI events may be summarized by treatment, or a listing may be provided.

#### **4.6.3.3. Hepatobiliary Disorders**

##### **4.6.3.3.1. Hepatobiliary Events**

Severe/serious treatment-emergent hepatobiliary disorders will be considered as AESIs. The counts and percentages of participants with treatment-emergent potentially drug-related hepatobiliary events may be summarized by treatment using the MedDRA PTs. The detailed search criteria can be found in Appendix 6.2.

Events related to acute gallbladder disease may also be summarized or a listing may be provided. The search criteria can be found in Appendix 6.2.

##### **4.6.3.3.2. Liver Enzymes**

Common analyses for laboratory analyte measurements described in Section 4.6.5 are applicable for the liver enzyme related measurements. This section describes additional analyses for liver enzymes.

The counts and percentages of participants with the following elevations in hepatic laboratory tests at any time during the treatment and follow-up period will be summarized between treatment groups:

- The counts and percentages of participants with an alanine aminotransferase measurement  $\geq 3\times$ ,  $5\times$ , and  $10\times$  ULN will be summarized for all participants with a postbaseline value and for subsets based on the following levels of baseline value:

- participants whose nonmissing maximum baseline value is  $\leq 1 \times \text{ULN}$
- participants whose maximum baseline is  $> 1 \times \text{ULN}$ , and
- participants whose baseline values are missing.
- The counts and percentages of participants with an aspartate aminotransferase measurement  $\geq 3 \times$ ,  $5 \times$ , and  $10 \times \text{ULN}$  during the study will be summarized for all patients with a postbaseline value and for subsets based on various levels of baseline, as described above for alanine aminotransferase.
- The counts and percentages of participants with a total bilirubin measurement  $\geq 2 \times \text{ULN}$  during the study will be summarized for all participants with a postbaseline value, and for the following subsets based on the baseline values:
  - participants whose nonmissing maximum baseline value is  $\leq 1 \times \text{ULN}$
  - participants whose maximum baseline is  $> 1 \times \text{ULN}$ , but  $< 2 \times \text{ULN}$
  - participants whose maximum baseline value is  $\geq 2 \times \text{ULN}$ , and
  - participants whose baseline values are missing.
- The counts and percentages of participants with a serum alkaline phosphatase measurement  $\geq 2 \times \text{ULN}$  during the study will be summarized for all participants with a postbaseline value and for the same subsets described above for total bilirubin.

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

#### 4.6.3.4. Severe Hypoglycemia

The following categories in accordance with the 2020 American Diabetes Association position statement on glycemic targets (ADA 2020) will be defined in the database.

##### **Glucose Alert Value (Level 1):**

- **Documented symptomatic hypoglycemia** is defined as any time a patient feels that he/she is experiencing symptoms and/or signs associated with hypoglycemia, and has a BG level of  $< 70 \text{ mg/dL}$  ( $< 3.9 \text{ mmol/L}$ ).
- **Documented asymptomatic hypoglycemia** is defined as any event not accompanied by typical symptoms of hypoglycemia, but with a measured BG  $< 70 \text{ mg/dL}$  ( $< 3.9 \text{ mmol/L}$ ).
- **Documented unspecified hypoglycemia** is defined as any event with no information about symptoms of hypoglycemia available, but with a measured BG  $< 70 \text{ mg/dL}$  ( $< 3.9 \text{ mmol/L}$ ).

##### **Documented Clinically Significant Hypoglycemia (Level 2):**

- **Documented symptomatic hypoglycemia** is defined as any time a participant feels that he/she is experiencing symptoms and/or signs associated with hypoglycemia and has a BG level of  $< 54 \text{ mg/dL}$  ( $< 3.0 \text{ mmol/L}$ ).
- **Documented asymptomatic hypoglycemia** is defined as any event not accompanied by typical symptoms of hypoglycemia but with a measured BG  $< 54 \text{ mg/dL}$  ( $< 3.0 \text{ mmol/L}$ ).

- **Documented unspecified hypoglycemia** is defined as any event with no information about symptoms of hypoglycemia available but with a measured BG <54 mg/dL (<3.0 mmol/L).

### **Severe Hypoglycemia (Level 3):**

Severe hypoglycemia is defined as an episode with severe cognitive impairment requiring the assistance of another person to actively administer carbohydrate, glucagon, or other resuscitative actions. These episodes may be associated with sufficient neuroglycopenia to induce seizure or coma. PG measurements may not be available during such an event, but neurological recovery attributable to the restoration of PG to normal is considered sufficient evidence that the event was induced by a low PG concentration.

### **Other hypoglycemia categories:**

Nocturnal hypoglycemia is defined as any hypoglycemic event that occurs between bedtime and waking. If a hypoglycemic event meets the criteria of severe, the event would specifically be collected as an SAE. Serious hypoglycemia is defined by pharmacovigilance criteria and will also be captured with a SAE form.

To avoid duplicate reporting, all consecutive hypoglycemic events occurring within a 1-hour period will be considered a single hypoglycemic event.

Incidence of documented symptomatic hypoglycemia events and severe hypoglycemia will be summarized and compared between tirzepatide MTD and placebo. Rate of hypoglycemic episodes will be analyzed using a generalized linear mixed-effects model assuming negative binomial distribution for hypoglycemic episodes if data warrant. Some analyses may be conducted excluding data after introducing another antihyperglycemic therapy

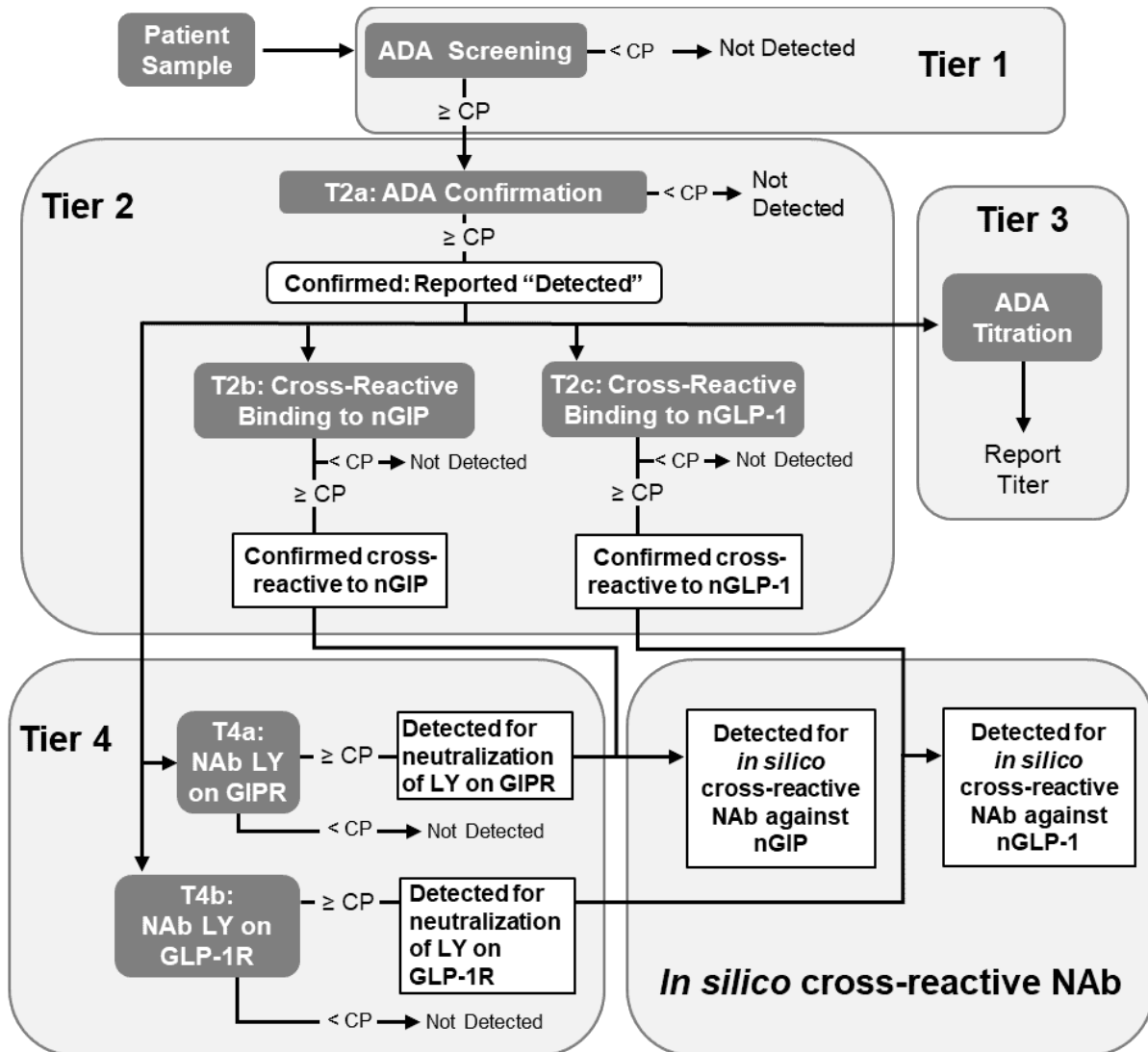
## **4.6.3.5. Immunogenicity**

### **4.6.3.5.1. Definitions of Sample ADA Status**

At a high level, an individual sample is potentially examined multiple times, in a hierarchical procedure, to produce a sample ADA assay result and potentially multiple cross-reactive antibodies assay results and multiple Nab assay results.

The cut points used, the drug tolerance of each assay, and the possible values of titers are operating characteristics of the assay. [Figure GPIF.4.1](#) details a flow chart that reflects the multitiered testing approach.





Abbreviations: ADA = anti-drug antibodies; CP = cut point; GIP = glucose-dependent insulinotropic polypeptide; GIPR = glucose-dependent insulinotropic polypeptide receptor; GLP-1 = glucagon-like peptide-1; GLP-1R = glucagon-like peptide-1 receptor; nGIP= native GIP; nGLP-1 = native GLP-1; LY = LY3298176; Nab = neutralizing antibodies.

Figure GPIF.4.1. Flowchart of immunogenicity multitiered testing approach.

[Table GPIF.4.6](#) outlines results as reported from Tier 2a of the multitiered testing approach. Tier 4 results are reported similarly.

**Table GPIF.4.6. Sample ADA Assay Results**

Sample Laboratory Result	Explanation
Detected	ADA are detected and confirmed.
Not Detected	The raw result as reported from the laboratory indicates not detected. The clinical interpretation of such results depends on other factors (see <a href="#">Table GPIF.4.7</a> ).
NO TEST, QNS, etc.	Sample exists but was unevaluable by the assay.

Abbreviations: ADA = anti-drug antibodies; QNS = quantity not sufficient.

It can be the case that the presence of high concentrations of tirzepatide will affect ADA immunoassays, and conversely high levels of ADA may affect the measurement of tirzepatide concentration. Thus, a tirzepatide drug concentration, assessed from a sample drawn at the same time as the ADA sample, plays a key role in clinical interpretation of a sample when the laboratory result is Not Detected (see [Table GPIF.4.7](#)).

**Table GPIF.4.7. Sample Clinical ADA Interpretation Results**

Sample Clinical Interpretation	Explanation
ADA Present	ADA assay result is Detected
ADA Not Present	ADA assay result is Not Detected and simultaneous drug concentration is at a level that has been demonstrated to not interfere in the ADA detection method (i.e., drug concentration is below the assay's drug tolerance level). For participants receiving placebo, drug concentration is not assessed and is assumed to be below the assay's drug tolerance level. If drug concentration was planned but is not available for a treatment-period sample, a Not Detected sample will be declared ADA Not present.
ADA Inconclusive	ADA assay result is Not Detected but drug concentration in the sample is at a level that can cause interference in the ADA detection method.
ADA Missing	ADA sample not drawn, QNS, not tested, etc., causing there to be no laboratory result reported or the result is reported as "no test."

Abbreviations: ADA = antidrug antibodies; QNS = quantity not sufficient.

All ADA present samples will be evaluated for cross-reactive GIP (Tier 2b), cross-reactive GLP-1 (Tier 2c), Nab LY (tirzepatide) on GIPR (Tier 4a), and Nab LY (tirzepatide) on GLP-1R (Tier 4b).

Similar terminology to [Table GPIF.4.7](#) applies for each type of cross-reactive and Nab assay. Importantly, each of these are distinct assays and, in general, have different assay operating characteristics. The following are considered inconclusive for the Nab result:

- Nab LY on GIPR: if Nab result is not detected, and PK concentration is greater than or equal to drug tolerance limit of the Nab LY on GIPR assay

- Nab LY on GLP-1R: if Nab result is not detected, and PK concentration is greater than or equal to drug tolerance limit of the Nab LY on GLP-1R assay

For cross-reactive Nab interpretations against native GIP and GLP-1, an in silico method utilizing results from Tiers 2b and 2c, Tiers 4a and 4b, and tirzepatide concentrations is introduced. The in silico method is outlined in [Table GPIF.4.8](#).

**Table GPIF.4.8. In Silico Classification for Cross-Reactive Nab**

<i>In Silico</i> Classification	Cross-Reactive ADA Result	Nab Result	Circulating Tirzepatide Level (ng/mL)	<i>In Silico</i> Cross- Reactive Nab Interpretation
Cross-reactive Nab to nGIP	Tier 2b: “Not Detected	Tier 4a: “Not Detected” Or Tier 4a: “Detected” or N/A or Missing	Any Value or Missing	Not Present
	Tier 2b: “Detected”	Tier 4a: “Not Detected”	< drug tolerance limit of Tier 4a assay	Not Present
	Tier 2b: “Detected”	Tier 4a: “Not Detected”	≥ drug tolerance limit of Tier 4a assay	Inconclusive
	Tier 2b: “Detected”	Tier 4a: “Detected”	< drug tolerance limit of Tier 4a assay	Present
	Tier 2b: “Detected”	Tier 4a: “Detected”	≥ drug tolerance limit of Tier 4a assay	Present
Cross-reactive Nab to nGLP- 1	Tier 2c: “Not Detected	Tier 4b: “Not Detected” Or Tier 4b: “Detected” or N/A or Missing	Any Value or Missing	Not Present
	Tier 2c: “Detected”	Tier 4b: “Not Detected”	< drug tolerance limit of Tier 4b assay	Not Present
	Tier 2c: “Detected”	Tier 4b: “Not Detected”	≥ drug tolerance limit of Tier 4b assay	Inconclusive
	Tier 2c: “Detected”	Tier 4b: “Detected”	< drug tolerance limit of Tier 4b assay	Present
	Tier 2c: “Detected”	Tier 4b: “Detected”	≥ drug tolerance limit of Tier 4b assay	Present

Abbreviations: ADA = antidrug antibodies; Nab = neutralizing antibody; nGIP = native glucose-dependent insulintropic polypeptide; nGLP-1 = native glucagon-like peptide-1; Tier 2b = cross-reactive ADA to nGIP; Tier 2c = cross-reactive ADA to nGLP-1; Tier 4a = Nab LY (tirzepatide) on GIPR; Tier 4b = Nab LY (tirzepatide) on GLP-1R.

Note: Only the drug tolerance limits of the Tier 4a and 4b assays are used for in silico classifications as they are lower than the drug tolerance limits of the Tier 2b and 2c assays, respectively.

#### **4.6.3.5.2. Definitions of Immunogenicity Assessment Periods**

**Immunogenicity baseline observations:** Baseline period for immunogenicity assessment for each participant includes all observations prior to first dose of study treatment. In instances where multiple baseline observations are collected, to determine participant ADA status the last

nonmissing immunogenicity assessment prior to first administration of study drug is used to determine treatment-emergent status (see below).

Immunogenicity postbaseline period observations: Postbaseline period observations for each participant includes all observations after the first administration of study drug.

#### **4.6.3.5.3. Definitions of Participant ADA Status**

TE ADA-evaluable participants: A participant with a nonmissing baseline ADA result and at least 1 nonmissing postbaseline ADA result.

TE ADA-unevaluable participant: any participant who does not meet the evaluable criteria.

Baseline ADA Present (preexisting antibody): ADA detected in a sample collected up to the first dose date and time.

Baseline ADA Not Present: ADA is not detected, and the corresponding PK concentration is missing or below the drug tolerance limit in a sample collected up to the first dose date and time.

TE ADA+ participant: An evaluable participant who had a:

- baseline status of ADA Not Present and at least 1 postbaseline status of ADA Present with titer  $\geq 2 \times$  MRD, where the MRD is the minimum required dilution of the ADA assay or
- baseline and postbaseline status of ADA Present, with the postbaseline titer being 2 dilutions (4-fold) greater than the baseline titer. That is, the participant has baseline (B) status of ADA Present, with titer 1:B, and at least 1 postbaseline (P) status of ADA Present, with titer 1:P, with  $P/B \geq 4$ .

As shown in [Figure GPIF.4.1](#), a titer is expected when ADA assay result is Detected. On occasion, the corresponding assay cannot be performed, in which case a titer value will be imputed for the purpose of TE ADA determination. A baseline sample with detected ADA and no titer is imputed to be the MRD (1:10), and a postbaseline sample with ADA detected and no titer is imputed to be one dilution above the MRD (1:20).

TE ADA-Inconclusive participant: A TE ADA-evaluable participant is TE ADA Inconclusive if  $\geq 20\%$  of the participant's postbaseline samples, drawn predose, are ADA Inconclusive and all remaining postbaseline samples are ADA Not Present.

TE ADA- participant: A TE ADA-evaluable participant is TE ADA- when the participant is not TE ADA+ and not TE ADA Inconclusive.

For each Nab assay, the following are defined:

Nab+ participant: A participant who is TE ADA+ and has a Nab positive sample in the postbaseline period.

Nab Inconclusive participant: A participant who is TE ADA+, is not Nab+, and all samples that have TE ADA+ titer have a Nab Inconclusive sample result.

Nab- participant: A participant is neither Nab+ nor Nab Inconclusive.

Unless specified otherwise, the above-mentioned definitions of Nab are applicable to all Nab analyses, including cross-reactive Nab analyses, and cross-reactive antibodies.

#### 4.6.3.5.4. Analyses to be Performed

The count and proportion of participants who are TE ADA+ will be tabulated by treatment group, where the proportions are relative to the number of TE ADA-evaluable participants, as defined above. The tabulation will include the count and proportion of participants with ADA Present at baseline, and the count and proportion of TE ADA+ participants exhibiting each type of cross-reactive antibodies and Nab. This analysis will be performed for the planned treatment period. The cross-reactive Nab will include the in silico classification as cross-reactive Nab for summary. A summary will be provided of the count and percentage of tirzepatide-treated participants experiencing specific TEAEs (see [Table GPIF.4.9](#)) by participant TE ADA status (TE ADA+, TE ADA-, TE ADA Inconclusive). The PT will be ordered by decreasing incidence in TE ADA+ status group.

**Table GPIF.4.9. Adverse Events for Analysis with Immunogenicity Results**

TEAE Category	Criteria
Hypersensitivity reactions	Anaphylaxis SMQ (narrow or algorithm)
	Hypersensitivity SMQ (narrow)
	Angioedema SMQ (narrow)
	Severe Cutaneous Adverse Reaction SMQ (narrow)
Injection site reactions	Injection site reaction HLT
	Infusion site reaction HLT
	Administration site reaction HLT

Abbreviations: HLT = High Level Term; MedDRA = Medical Dictionary for Regulatory Activity; SMQ = Standardized MedDRA Query; TEAE = treatment-emergent adverse event.

Additional immunogenicity analyses as determined later may be presented. The relationship between the presence of antibodies and tirzepatide PK and PD response including safety and efficacy to tirzepatide may be assessed.

Cases of TE ADA that are associated with TEAEs of either severe/serious hypersensitivity or severe/serious ISR will be classified as AESIs.

#### 4.6.3.6. Hypersensitivity Reactions

Two main analyses are performed in support of assessment of potential immediate hypersensitivity, including anaphylaxis as well as potential nonimmediate hypersensitivity.

**Time Period A**, of potential immediate hypersensitivity includes all TEAEs occurring from start of study drug administration up to 24 hours after end of study drug administration. For events without the hypersensitivity eCRF, only date (no time) information is collected. Among these events without time information, the event occurred on the same date as the study drug injection date will be included in Time Period A.

**Time Period B**, of potential non-immediate hypersensitivity, includes all TEAEs occurring more than 24 hours after the end of study drug administration, but prior to subsequent drug administration.

Analyses for both time periods are based on the following:

- narrow and algorithm terms in Anaphylactic reaction SMQ (20000021) (analysis for algorithm term only applicable for Time Period A)
- narrow terms in Angioedema SMQ (20000024)
- narrow terms in Severe cutaneous adverse reactions SMQ (20000020), and
- narrow terms in Hypersensitivity SMQ (20000214)

For the Anaphylactic reaction SMQ, each term is classified by scope (narrow, broad) and by category (A, B, C, D). All narrow terms are category A, and all broad terms are category B, C, or D. In addition to the usual narrow and broad searches, the SMQ defines an algorithm to further refine the cases of interest. For Time Period A analysis, the Anaphylactic reaction SMQ algorithm will be included. The algorithm is based upon events that occur within Time Period A. The counts and percentages of participants who experienced a TEAE for the following will be analyzed for each of the 2 time periods:

- any narrow term from any one of the 4 SMQs indicated above (that is, combined search across narrow of all 4 SMQs), and
- any narrow scope term within each SMQ, separately (that is, narrow SMQ search). For Time Period A analysis, any term from Anaphylactic reaction SMQ algorithm.

Within each query, individual PTs that satisfied the queries will be summarized. For Time Period A analysis, the Anaphylactic reaction SMQ algorithm will be summarized. Also, a single event may satisfy multiple SMQs, in which case the event contributes to every applicable SMQ.

#### **4.6.3.6.1. Severe/Serious Hypersensitivity Reactions**

The severe/serious cases of hypersensitivity will be considered as AESIs. Summary of severe/serious hypersensitivity reactions or listing may be provided.

#### **4.6.3.7. Injection Site Reaction**

Injection site reaction, incidence and rates, and related information reported via “Injection Site Reactions” eCRF will be summarized by treatment. Information to be summarized include location of the reaction, timing of reaction relative to study drug administration, and characteristics of the injection site reaction: erythema, induration, pain, pruritus, and edema.

Patient-based analysis and event-based analysis may be provided if necessary. The patient-based analysis summarizes all ISR questionnaire forms for an individual patient with a single statistic, typically an extreme value. This analysis allows each patient to contribute only once for each parameter, at the expense of a focus on the most extreme events. By contrast, the event-based analysis summarizes all ISR questionnaire forms received, without regard to individual patients. This provides characteristics of ISR events as a proportion of all events for which questionnaire

responses were provided, at the expense of some potential bias due to differential contribution of individual patients to the analysis.

The counts and percentages of participants with treatment-emergent injection site reaction will be summarized by treatment using the MedDRA PTs. Detailed searching criteria can be found in Appendix 6.2.

The PTs will be summarized in decreasing order of incidence for tirzepatide-treated participants.

#### **4.6.3.7.1. Severe/Serious Injection Site Reactions**

Severe/serious injection site reactions (for example, abscess, cellulitis, erythema, hematomas/hemorrhage, exfoliation/necrosis, pain, subcutaneous nodules, swelling, induration, inflammation) will be considered as AESI.

The counts and percentage of participants with severe/serious ISRs may be summarized by treatment, or a listing of participants with treatment-emergent severe/serious ISRs may be provided.

#### **4.6.3.8. Major Adverse Cardiovascular Events**

MACE reported by investigators are adjudicated by an independent clinical endpoint committee in a blinded fashion. Unreported events may also be independently identified by the clinical endpoint committee.

The following positively adjudicated MACE will be considered as AESIs:

- death due to cardiovascular AEs
- myocardial infarction
- hospitalization for unstable angina
- hospitalization for heart failure
- coronary interventions (such as coronary artery bypass graft or percutaneous coronary intervention, and
- cerebrovascular events, including cerebrovascular accident (stroke) and transient ischemic attack.

The counts and percentages of participants with adjudicated MACE may be summarized by treatment. In addition, MACE reported by investigator may also be summarized although a MACE reported by investigator that is not positively adjudicated is not considered an AESI.

A listing of participants reporting MACE events, either reported by investigator or identified by the clinical endpoint committee, may be provided.

#### **4.6.3.9. Major Depressive Disorder/Suicidal Ideation or Behavior**

The severe/serious treatment-emergent major depressive disorder/suicidal ideation or behavior will be captured as AESI. Adverse events will be searched using MedDRA PT terms. Detailed searching criteria can be found in Appendix 6.2.

The counts and percentages of participants with TEAEs will be summarized by treatment group using MedDRA PT nested within SMQ. Events will be ordered by decreasing frequency in the

total tirzepatide group nested within SMQ. A listing of participants with major depressive disorder/suicidal ideation or behavior may be provided if deemed necessary.

Additionally, suicidal ideation and behavior, and depression will be assessed by the investigator via spontaneously reported AEs and through the use of the Columbia-Suicide Severity Rating Scale (C-SSRS) and the PHQ-9.

#### **4.6.3.9.1. Patient Health Questionnaire**

Total scores for the PHQ-9 range from 0 to 27 with total scores categorized as

- none (not depressed): 0 through 4
- mild: 5 through 9
- moderate: 10 through 14
- moderately severe: 15 through 19, and
- severe: 20 through 27.

Shift tables will be provided showing the counts and percentages of participants within each baseline category (maximum value) versus each postbaseline category (maximum value) by treatment.

Additionally, the following 3 outcomes of interest will be compared between treatments (based on the maximum value during baseline and postbaseline):

- any increase in depression category (that is, worsening of depression): includes participants in the none, mild, moderate, or moderately severe category during baseline and with at least 1 postbaseline measurement
- increase from No or Mild Depression to Moderate, Moderately Severe or Severe Depression: includes participants in the none or mild depression category during baseline and with at least 1 postbaseline measurement; and
- increase from Mild or Moderate Depression to Moderately Severe or Severe Depression: includes participants in the mild or moderate depression category during baseline and with at least 1 postbaseline measurement

#### **4.6.3.9.2. Suicidal Ideation and Behavior Solicited Through C-SSRS**

Suicide-related thoughts and behaviors occurring during the entire study period, based on the C-SSRS, will be summarized by treatment group. In particular, for each of the following suicide related events, the counts and percentages of participants with the event will be summarized by treatment group:

- died by suicide
- nonfatal suicide attempt
- interrupted attempt
- aborted attempt
- preparatory acts or behavior
- active suicidal ideation with specific plan and intent
- active suicidal ideation with some intent to act without specific plan



- active suicidal ideation with any methods (no plan) without intent to act
- nonspecific active suicidal thoughts
- wish to be dead, and
- nonsuicidal, self-injurious behavior.

In addition, the counts and percentages of participants who experienced at least 1 of the composite measures will be presented. The participants with at least 1 post-baseline C-SSRS assessment are included. The composite measure is determined at each assessment by the “yes” or “no” responses in C-SSRS categories by the study participant:

- Category 1 – Wish to be Dead
- Category 2 – Non-specific Active Suicidal Thoughts
- Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
- Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
- Category 5 – Active Suicidal Ideation with Specific Plan and Intent
- Category 6 – Preparatory Acts or Behavior
- Category 7 – Aborted Attempt
- Category 8 – Interrupted Attempt
- Category 9 – Actual Attempt (non-fatal), and
- Category 10 – Completed Suicide.

Composite endpoints of suicidal ideation and suicidal behavior based on the above categories are defined below:

- **Suicidal ideation:** A “yes” answer at any time during study to any 1 of the 5 suicidal ideation questions (Categories 1 through 5) on the C-SSRS.
- **Suicidal behavior:** A “yes” answer at any time during study to any 1 of the 5 suicidal behavior questions (Categories 6 through 10) on the C-SSRS.
- **Suicidal ideation or behavior:** A “yes” answer at any time during study to any 1 of the 10 suicidal ideation and behavior questions (Categories 1 through 10) on the C-SSRS.

A listing contains data for each participant with suicidal ideation, suicidal behavior, or nonsuicidal self-injurious behavior during the study by treatment and visit. Data from all visits are displayed, regardless of a “yes” or “no” answer, for participants with any “yes” answer for suicidal ideation, suicidal behavior, or self-injurious behavior without suicidal intent.

#### 4.6.3.10. Renal Safety

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 4.6.5.

In addition, 2 shift tables examining renal function will be created. A min-to-min shift table of estimated glomerular filtration rate estimated by the Chronic Kidney Disease Epidemiology Collaboration equation with units mL/min/1.73 m<sup>2</sup>, using categories (<30, ≥30 to <45, ≥45 to <60, ≥60 to <90, and ≥90 mL/min/1.73m<sup>2</sup>). A max-to-max shift table of UACR, using the categories UACR <30 mg/g, ≥30 mg/g UACR to ≤300 mg/g, UACR >300 mg/g (respectively, these represent normal, microalbuminuria, and macroalbuminuria).

MMRM analyses as described in Section 4.6 for estimated glomerular filtration rate and log-transformed UACR will be provided. Log transformation will be performed for UACR.

#### **4.6.3.10.1. Acute Renal Events**

Because severe GI events may lead to dehydration, which could cause a deterioration in renal function including acute renal failure, dehydration events will be analyzed as described in the next section. Acute renal events associated with chronic renal failure exacerbation will also be captured.

Severe/serious renal events from the following SMQ search will be considered as AESI.

The counts and percentages of participants with acute renal events may be summarized by treatment if overall count >10 by using the MedDRA PTs contained in any of the following SMQs:

- Acute renal failure: narrow terms in Acute renal failure SMQ (20000003) and
- Chronic kidney disease: narrow terms in Chronic kidney disease SMQ (20000213).

In addition, a listing of participants with treatment-emergent acute renal events may be provided, if deemed necessary.

#### **4.6.3.10.2. Dehydration**

Dehydration events will be captured in the narrow terms in Dehydration SMQ (20000232). Severe/serious dehydration events will be considered as AESI. A listing of participants with treatment-emergent dehydration events may be provided.

#### **4.6.3.11. Thyroid Safety Monitoring**

##### **4.6.3.11.1. Calcitonin**

The purpose of calcitonin measurements is to assess the potential of tirzepatide to affect thyroid C-cell function, which may indicate development of C-cell hyperplasia and neoplasms.

Observed calcitonin data (a thyroid-specific laboratory assessment) will be summarized by treatment and nominal visit.

The counts and percentages of participants with a maximum postbaseline calcitonin value in the following thresholds will be provided by treatment and baseline calcitonin value ( $\leq 20$  ng/L,  $>20$  ng/L to  $\leq 35$  ng/L,  $>35$  ng/L). Postbaseline categories are:  $\leq 20$  ng/L,  $>20$  ng/L to  $\leq 35$  ng/L,  $>35$  ng/L to  $\leq 50$  ng/L,  $>50$  ng/L to  $\leq 100$  ng/L, and  $>100$  ng/L.

##### **4.6.3.11.2. C-Cell Hyperplasia and Thyroid Malignancies**

Treatment-emergent thyroid malignancies and C-cell hyperplasia will be considered as AESI. Thyroid malignancies and C-cell hyperplasia will be identified using MedDRA HLT for thyroid neoplasms and PT for thyroid C-cell hyperplasia.

The counts and percentages of participants with treatment-emergent thyroid C-cell hyperplasia and malignancies may be summarized or a listing of participants with treatment-emergent thyroid C-cell hyperplasia and neoplasms may be provided if deemed necessary.

#### **4.6.3.12. Treatment-Emergent Arrhythmias and Cardiac Conduction Disorders**

Severe/serious treatment-emergent arrhythmias and cardiac conduction disorders will be considered as AESI. The cardiovascular events will include clinically relevant rhythm and conduction disorders.

The treatment-emergent arrhythmias and cardiac conduction disorder events will be identified using the MedDRA PTs. Detailed searching criteria can be found in Appendix 6.2.

The counts and percentages of participants with treatment emergent arrhythmias and cardiac conduction disorders may be summarized by treatment and PT nested within SMQ. The PT will be ordered with decreasing frequency in TZP arm within SMQ. A listing of participants with treatment-emergent arrhythmias and cardiac conduction disorders may be provided if deemed necessary.

#### **4.6.4. Vital Signs**

In the case that multiple records of an individual vital sign are collected at the same visit, they will be averaged prior to being used for data summaries and analyses.

Descriptive summaries by treatment and by nominal visit will be provided for baseline and postbaseline values as well as change from baseline values.

An MMRM and/or an ANCOVA model as in Section 4.6 using data up to 52 weeks only might be conducted if necessary.

Counts and percentages of participants with treatment-emergent abnormal sitting SBP, sitting diastolic blood pressure, and pulse at any time during the entire study (including the off-drug follow up time period) will be presented by treatment for participants who have both baseline and at least 1 postbaseline result. A treatment-emergent high result is defined as a change from a value less than or equal to the high limit at baseline to a value greater than the high limit at any time that meets the specified change criteria during the postbaseline period. A treatment-emergent low result is defined as a change from a value greater than or equal to the low limit at baseline to a value less than the low limit at any time that meets the specified change criteria during the postbaseline period. To assess decreases, change from the minimum value during the baseline period to the minimum value during the postbaseline period will be used. To assess increases, changes from the maximum value during the baseline period to the maximum value during the postbaseline period will be used. Both planned and unplanned measurements will be included in the analysis. The criteria for identifying participants with treatment-emergent vital sign abnormalities are stated in [Table GPIF.4.10](#).

**Table GPIF.4.10. Categorical Criteria for Abnormal Treatment-Emergent Blood Pressure and Pulse Measurements**

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

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

#### 4.6.5. Clinical Laboratory Evaluation

Limits from the performing lab will be used to define low (L) and high (H). Descriptive summaries by treatment and by nominal visit will be provided for the baseline and postbaseline values as well as the change from baseline values. The associated descriptive will be presented in International System of Units and in conventional units.

For selected laboratory analyte measurements collected quantitatively, observed, and change from baseline values for each visit may be displayed in plots for participants who have both a baseline and at least 1 postbaseline planned measurement. Unplanned measurements will be excluded from plots.

A shift table will be provided including unplanned measurements. The shift table will include the number and percentage of participants within each baseline category (low, normal, high, or missing) versus each postbaseline category (low, normal, high, or missing) by treatment. The proportion of participants shifted will be compared between treatments using Fisher's exact test.

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

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

The MMRM model or ANCOVA (if MMRM model is not applicable) as described in Section 4.6 will be used for the analysis during the treatment period (excluding the safety follow-up period) for the continuous measurements for selected lab tests with or without log-transformed (postbaseline measure/baseline measure) response variables. For measures analyzed using log-transformed values, the results will be presented with the scale back transforming to the original, related scale.

The summary of treatment-emergent abnormal, high, or low laboratory results any time will be provided.

## 4.7. Other Analyses

### 4.7.1. PAP Adherence

For patients in GPI2 only, we will summarize the adherence to use of the PAP machine over the course of the study. Specifically, we will provide summary statistics for PAP adherence at baseline and at each postbaseline week, stratified by treatment arm. Additionally, we will assess the categorical shift in PAP adherence between baseline and Week 52, stratified by treatment arm using a shift table for increased, decreased, or stable PAP use. Finally, we will also summarize the number and percentage of patients in GPI2 who withdrawal from regular PAP use.

### 4.7.2. Health Outcomes

The PRO questionnaires will be analyzed using the mITT population on the EAS, unless specified otherwise.

Item-level missingness will be dealt with per the instrument developers' instruction.

Additional psychometric analyses will be performed by Value, Evidence, and Outcomes at Lilly and documented in a separate analysis plan.

Analyses of actual and change from baseline in PRO scores will be conducted using linear models with baseline PRO scores, treatment, stratification factors, and other factors that may be considered relevant.

If an administrative error occurs where more than one PRO questionnaire is completed within the same visit window by the same participant with different responses on at least 1 item, the questionnaire with the worst response will be used (for example, the questionnaire with the highest PHQ-9 score will be used). If more than 1 PRO questionnaire is completed within the same visit window with the same response to each item, the most recent response will be used.

#### 4.7.2.1. Functional Outcomes of Sleep Questionnaire

The FOSQ will be included to assess change in FOSQ domains and total score from baseline to Week 52. The FOSQ is a 30-item sleep-specific, participant-completed questionnaire used to assess the effect of disorders associated with excessive daytime sleepiness on daily functioning in adults. It assesses the following 5 domains of

- General productivity (8 items)
- Activity level (9 items)
- Vigilance (7 items)
- Social outcomes (2 items), and
- Intimate and sexual relationships (4 items).

The FOSQ items assess participant's current status with each item rated on a scale of 1 (extreme difficulty) to 4 (no difficulty), with an additional not applicable (0 = "I don't do this activity for other reasons") also available. Individual domain scores are calculated by taking the mean of answered, non-zero items within each domain and ranges from 1 to 4 with higher scores

indicating better outcomes. A total score can be calculated by first computing the mean score for each domain, then multiplying the mean of the domain scores by 5 (Weaver et al. 1997). The change from baseline in all 5 FOSQ domain scores will be assessed.

#### **4.7.2.2. Functional Outcomes of Sleep Questionnaire, 10 items**

The FOSQ-10 will be included to assess change in FOSQ total score domains from baseline to Week 52. The 10-item sleep-specific, participant-completed questionnaire is a shortened version of the FOSQ with the same number of domains as the parent version. Of note, the FOSQ-10 has the same 5 domains as the FOSQ, but with fewer items per domain.

Calculation of the individual domain scores and the total score for the FOSQ-10 is carried out in a similar manner to FOSQ. The domain scores are first calculated by taking the mean of the answered, non-zero items within each domain. The total score is calculated by multiplying the mean of the domain scores by 5 (for each domain which has at least 1 response).

#### **4.7.2.3. Epworth Sleepiness Scale**

The ESS is an 8-item participant-completed measure that asks the participant to rate on a scale of 0 (would never doze) to 3 (high chance of dozing), their usual chances of dozing in 8 different daytime situations, with a recall period of “in recent times.” The ESS total score is the sum of the 8-item scores and ranges from 0 to 24, with higher scores indicating greater daytime sleepiness (Johns 1991). Of note, if one or more item scores are missing, the ESS total score is not valid and will not be included in the analysis.

#### **4.7.2.4. PROMIS Short Form v1.0 Sleep-Related Impairment 8a**

The PROMIS Short Form v1.0 Sleep-Related Impairment 8a consists of 8 items each rated on a 5-point scale ranging from “not at all” to “very much.” Items have a recall period of “in the past 7 days.” Individual item scores are totaled to obtain a raw score, with higher scores indicating more sleep-related impairment. The raw score can be calculated by summing the responses for the items. Raw scores can be converted to a T-score for each participant using a reference table (Northwestern 2016a) as long as the participant has provided a response for all questions. The T-score standardizes the raw score to a distribution with a mean of 50 and standard deviation of 10.

#### **4.7.2.5. PROMIS Short Form v1.0 Sleep Disturbance 8b**

The PROMIS Short Form v1.0 Sleep Disturbance 8b consists of 8 items each rated on a 5-point scale ranging from “not at all” to “very much,” “never” to “always,” or “very poor” to “very good.” Items have a recall period of “in the past 7 days.” Individual item scores are totaled to obtain a raw score, with higher scores indicating more sleep disturbance. The raw score can be calculated by summing the responses for the items. Raw scores can be converted to a T-score for each participant using a reference table (Northwestern 2016b) as long as the participant has provided a response for all questions. The T-score standardizes the raw score to a distribution with a mean of 50 and standard deviation of 10.

For item 8 of this scale (which is a measure of sleep quality), counts and percentages of participants at each time point will be summarized by nominal visit and by treatment. A shift table from baseline to postbaseline will be created at each postbaseline visit.

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

Per copyright owner, the QualityMetric Health Outcomes™ Scoring (PRO\_CoRe V2.0) Software will be used to derive the following domain and component scores:

- Mental Component Score (MCS)
- Physical Component Score (PCS)
- Physical Functioning domain (PF)
- Role-Physical domain (RP)
- Bodily Pain domain (BP)
- General Health domain (GH)
- Vitality domain (VT)
- Social Functioning domain (SF)
- Role-Emotional domain (RE), and
- Mental Health domain (MH).

The following analyses for the actual value and change from baseline value for each domain and component score will be conducted:

- descriptive summaries by treatment group, and
- analysis described in [Table GPIF.4.3](#).

#### **4.7.2.7. Patient Global Impression of Status/Change for OSA Outcomes**

The counts and percentages of participants for Patient Global Impression of Status for Physical Activity and Patient Global Impression of Change response categories at each time point will be summarized by nominal visit and by treatment. A shift table from baseline to postbaseline of 3 Patient Global Impression of Status (OSA Sleepiness, Fatigue, Snoring) response categories and 4 Patient Global Impression of Change (OSA Sleepiness, Fatigue, Snoring, Sleep Quality) response categories will be created at each postbaseline visit.

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

For the utility score and the Visual Analog Scale scores, following analyses of the actual value and change from baseline value will be conducted:

- descriptive summaries by treatment group and
- analysis of covariance described in [Table GPIF.4.4](#).

### **4.7.3. Subgroup Analyses**

The following subgroups will be analyzed using the efficacy estimand on percent change in AHI values from baseline to 52-week visit if there is an adequate number of patients in each treatment by subgroup (for example, 10%):

- age (<50 years, ≥50 years)

- baseline OSA severity (moderate, severe)
- race
- ethnicity
- region of enrollment (US, outside of US)
- gender (male or female)
- baseline BMI (<35, ≥35 and <40, ≥40 kg/m<sup>2</sup>)
- baseline ESS (ESS ≤10, ESS >10)

Analyses for percent change from baseline in AHI will be performed using an MMRM model that includes the same fixed effects given for the primary analysis model plus factors of subgroup, 2-way interaction of subgroup and treatment, 2-way interaction of subgroup and visit, and 3-way interaction of treatment, visit and subgroup. The possible interaction effect of subgroup and treatment at the primary endpoint (Week 52) will be evaluated. When analyzing OSA severity (moderate, severe) as a subgroup, the baseline AHI will be not be included as a covariate to avoid confounding.

#### **4.8. Interim Analyses**

The details for the interim analyses and Data Monitoring Committee will be provided in the Data Monitoring Committee Charter.

##### **4.8.1. Unblinding Plan**

Details of the blinding and unblinding are provided in the Blinding and Unblinding Plan document for Master Protocol GPIF.

#### **4.9. Changes to Protocol-Planned Analyses**

Not applicable.



## **5. Sample Size Determination**

Details of sample size determination are provided in the protocol.

## 6. Supporting Documentation

### 6.1. Appendix 1: Clinical Trial Registry Analyses

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

Analyses provided for the Clinical Trial Registry requirements include the following:

- Summary of AEs, provided as a dataset which will be converted to an XML file. Both SAEs and ‘Other’ non-SAEs are summarized: by treatment group, by MedDRA PT.
  - An AE is considered ‘Serious’ whether or not it is a TEAE.
  - An AE is considered in the ‘Other’ category if it is both a TEAE and is not serious. For each SAE and ‘Other’ AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event
    - the number of participants who experienced each event term, and
    - the number of events experienced.
- For each SAE, these additional terms are provided for EudraCT:
  - the total number of occurrences causally related to treatment
  - the total number of deaths, and
  - the total number of deaths causally related to treatment.
- Consistent with [www.ClinicalTrials.gov](http://www.ClinicalTrials.gov) requirements, ‘Other’ AEs that occur in fewer than 5% of patients/subjects in every treatment group may be excluded.
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

Demographic table including the following age ranges required by EudraCT: adults (18 to 64 years), (65 to 85 years), and (85 years and over).

### 6.2. Appendix 2: Search Criteria for Special Safety Topics

#### Arrhythmias and cardiac conduction disorders

Treatment-emergent arrhythmias, arrhythmias and cardiac conduction disorders will be considered an AESI. The cardiovascular events will include clinically relevant rhythm and conduction disorders. The treatment-emergent arrhythmias and cardiac conduction disorders events will be included using the MedDRA PT contained in any of the following SMQs:

- 1) Arrhythmias:
  - a. For symptoms: Arrhythmia related investigations, signs and symptoms SMQ (20000051), narrow and broad terms
  - b. For supraventricular arrhythmias: In Cardiac arrhythmia SMQ, under tachyarrhythmia sub SMQ
    - i. Supraventricular tachyarrhythmia SMQ (20000057), broad and narrow terms

- ii. Tachyarrhythmia terms, nonspecific SMQ (20000164), narrow terms only; and
  - iii. Ventricular tachyarrhythmia SMQ (20000058), narrow terms only.
- 2) Cardiac conduction disorders
  - a. Conduction defects SMQ (20000056), narrow terms only; and
  - b. Cardiac conduction disorders HLT (10000032), all PTs.

### **Hepatic treatment-emergent adverse events**

Treatment-emergent potentially drug-related hepatic disorders will be summarized by treatment using the MedDRA PTs contained in any of the following SMQs:

- broad and narrow terms in the Liver related investigations, signs and symptoms SMQ (20000008)
- broad and narrow terms in the Cholestasis and jaundice of hepatic origin SMQ (20000009)
- broad and narrow terms in the Hepatitis non-infections SMQ (20000010)
- broad and narrow terms in the Hepatic failure, fibrosis and cirrhosis and other liver damage SMQ (20000013)
- narrow terms in the Liver-related coagulation and bleeding disturbances SMQ (20000015)
- narrow PTs in Gallbladder related disorders SMQ (20000124)
- narrow PTs in Biliary tract disorders SMQ (20000125); and
- narrow PTs in Gallstone related disorders SMQ (20000127).

### **Acute gallbladder disease**

All events of TEAE biliary colic, cholecystitis, or other suspected events related to gallbladder disease will be summarized by treatment groups by PT with decreasing frequency under following SMQs:

- narrow PTs in Gallbladder related disorders SMQ (20000124)
- narrow PTs in Biliary tract disorders SMQ (20000125), and
- narrow PTs in Gallstone related disorders SMQ (20000127).

### **Major depressive disorder/suicidal ideation**

AEs will be searched using MedDRA PTs from SMQs narrow scope: 20000037 [Suicide/self-injury] and 20000167 [Depression (excl suicide and self injury)].

### **C-cell hyperplasia and thyroid malignancies**

Thyroid malignancies and C-Cell hyperplasia will be identified using MedDRA HLT for Thyroid neoplasms and PT for thyroid C-cell hyperplasia.

## Hypersensitivity reactions

Analyses are based on the following:

- narrow and algorithm terms in Anaphylactic reaction SMQ (20000021)
- narrow terms in Angioedema SMQ (20000024)
- narrow terms in Severe cutaneous adverse reactions SMQ (20000020), and
- narrow terms in Hypersensitivity SMQ (20000214).

For the Anaphylactic reaction SMQ, each term is classified by scope (narrow, broad) and by category (A, B, C, D). All narrow terms are category A, and all broad terms are category B, C, or D. In addition to the usual narrow and broad searches, the SMQ defines an algorithm to further refine the cases of interest. For Time Period A analysis, the Anaphylactic reaction SMQ algorithm will be included. The algorithm is based upon events that occur within Time Period A. The counts and percentages of participants who experienced a TEAE for the following will be analyzed for each of the 2 time periods:

- any narrow term from any one of the 4 SMQs indicated above (that is, combined search across narrow of all 4 SMQs), and
- any narrow scope term within each SMQ, separately (that is, narrow SMQ search). For Time Period A analysis, any term from Anaphylactic reaction SMQ algorithm.

## Injection site reactions

The ISR AE will be identified using the MedDRA PT in any of the following:

- HLT of Injection site reaction
- HLT of Administration site reaction, and
- HLT of Infusion site reactions.

## Pancreatitis events

Determination of investigator-reported events will be through the “Acute pancreatitis” MedDRA SMQ (20000022, narrow scope) and a “Chronic pancreatitis” PT search of the AE database, while adjudication-confirmed pancreatitis is found from adjudication forms.

## 6.3. Appendix 3: Magnetic Resonance Imaging Addendum

This section is applicable to the participants who are enrolled in the MRI addendum.

This addendum applies to a subset of participants (approximately 58 patients) enrolled in GPII. MRIs for the assessment of upper airway fat disposition will be collected at baseline and Week 52. The MRI at baseline needs to be completed prior to Visit 2 or within 7 days after Visit 2. The MRI at Week 52 may be scheduled for any day  $\pm$  14 days.

MRI analyses will be guided by the treatment policy strategy and conducted among all patients who are enrolled in the addendum, received at least 1 dose of study drug, and have at least 1 MRI measurement. The patient’s demographics and baseline characteristics for the MRI addendum will be summarized.

Objectives	Endpoints	Analytical Approaches
<p>Compare the effect of once weekly tirzepatide at MTD versus placebo on the changes of soft tissues volumes, fat volumes and fat content (%) in upper airway structures and in the abdomen in patients with OSA and obesity.</p>	<p>Changes of absolute soft tissue volume, fat volume and fat content (%) of the following:</p> <ul style="list-style-type: none"> <li>• Tongue</li> <li>• Soft palate</li> <li>• Pterygoid muscle</li> <li>• Lateral pharyngeal walls</li> <li>• Neck and submandibular area</li> <li>• Total, visceral and subcutaneous abdominal fat</li> </ul> <p>Change of total, subcutaneous and visceral abdominal adipose tissue volume</p>	<p>Change from baseline to Week 52 for each parameter will be compared between treatment arms using an analysis of covariance (ANCOVA) approach. The model will include treatment, the stratification factors of country, gender, and baseline AHI (moderate/severe), and the baseline value for the parameter. Summary statistics for MRI parameters at baseline and at Week 52 will be provided. The treatment comparison at baseline will be performed using an ANOVA model.</p>
<p>Explore correlation of changes of soft tissue volumes, fat volumes, and fat content (%) in upper airway structures with changes of AHI.</p>	<p>Correlations between the change in absolute soft tissue volume, fat volume, and fat content (%) for the structures listed above and the % change in AHI.</p>	<p>Spearman correlations between the change from baseline for each of the MRI endpoints and the % change in AHI will be calculated.</p>

Abbreviations: AHI = Apnea-Hypopnea Index; ANCOVA = analysis of covariance; ANOVA = analysis of variance; MRI = magnetic resonance imaging; MTD = maximum tolerated dose; OSA = obstructive sleep apnea.

## 7. References

- [ADA] American Diabetes Association Glycemic targets: Standards of medical care in diabetes2020. *Diabetes Care*. 2020;43(Suppl 1):S66–S76. <https://doi.org/10.2337/dc20-S006>
- Error! Hyperlink reference not valid.** Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540-545. <https://doi.org/10.1093/sleep/14.6.540>
- [Northwestern]. Northwestern University. PROMIS Short Form v1.0 – Sleep-Related Impairment 8a. Published May 09, 2016a. Accessed December 14, 2022. [https://www.healthmeasures.net/index.php?option=com\\_instruments&view=measure&id=185&Itemid=992](https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=185&Itemid=992)
- [Northwestern]. Northwestern University. PROMIS Short Form v1.0 – Sleep Disturbance 8b. Published May 03, 2016b. Accessed December 14, 2022. [https://www.healthmeasures.net/index.php?option=com\\_instruments&view=measure&id=184&Itemid=992](https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=184&Itemid=992)
- Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. 2012;33(2):176-182. <https://doi.org/10.1093/eurheartj/her352>
- Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons Inc.; 1987.
- Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*. 1997;20(10):835-843.
- Weaver T, Monno D, Bron M, et al. Determination of thresholds for minimally important difference and clinically important response on the functional outcomes of sleep questionnaire short version in adults with narcolepsy or obstructive sleep apnea. *Sleep Breath*. 2021;25:1707-1715. <https://doi.org/10.1007/s11325-020-02270-3>

Signature Page for VV-CLIN-069584 v1.0

Approval

**PPD**

Signature Page for VV-CLIN-069584 v1.0

# **Statistical Analysis Plan I8F-MC-GPIF: Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)**

**Protocol Title:** A Master Protocol to Investigate the Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA)

**Protocol Number:** I8F-MC-GPIF

**Compound Number:** Tirzepatide (LY3298176)

**Short Title:** A Statistical Analysis Plan for Tirzepatide in Participants with Obstructive Sleep Apnea and Obesity

**Sponsor Name:** Eli Lilly and Company

**Legal Registered Address:** Indianapolis, Indiana USA 46285

**Regulatory Agency Identifier Number(s)**

IND 157090

## **Confidential Information**

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

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

**Document ID:** VV-CLIN-117998



## Table of Contents

Statistical Analysis Plan I8F-MC-GPIF: Efficacy and Safety of Tirzepatide Once Weekly in Participants who have Obstructive Sleep Apnea and Obesity: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-OSA).....1

Table of Contents .....2

Version History .....4

1. Introduction .....16

    1.1. Objectives, Endpoints, and Estimands .....16

        1.1.1. Estimands .....18

    1.2. Study Design .....20

2. Statistical Hypotheses .....22

    2.1. Multiplicity Adjustment .....22

3. Analysis Sets .....24

4. Statistical Analyses .....25

    4.1. General Considerations .....25

        4.1.1. Adjustment for Covariates .....28

        4.1.2. Handling of Dropouts or Missing Data .....28

        4.1.3. Multicenter Studies .....29

        4.1.4. Historical Illnesses and Preexisting Conditions .....29

        4.1.5. Participant Characteristics .....30

        4.1.6. Concomitant Therapy .....30

        4.1.7. Treatment Exposure and Compliance .....31

        4.1.8. Important Protocol Deviations .....31

    4.2. Participant Dispositions .....31

    4.3. Primary Endpoint Analysis .....32

        4.3.1. Analysis Related to the Efficacy Estimand .....32

        4.3.2. Analysis Related to the Treatment Regimen Estimand .....32

        4.3.3. Sensitivity Analyses .....33

    4.4. Secondary Endpoints Analysis .....33

        4.4.1. Key Secondary Endpoints .....33

        4.4.2. Type 1 Error Rate Control Strategy for Primary and Key Secondary Efficacy Analyses .....35

        4.4.3. Supportive Secondary Endpoints .....35

    4.5. Exploratory Endpoint Analyses .....38

    4.6. Safety Analyses .....39

        4.6.1. Analysis of Adverse Events .....39

- 4.6.2. Patient Narratives .....41
- 4.6.3. Special Safety Topics .....41
- 4.6.4. Vital Signs .....56
- 4.6.5. Clinical Laboratory Evaluation .....57
- 4.6.6. Product Complaints .....57
- 4.7. Other Analyses .....58
  - 4.7.1. PAP Adherence .....58
  - 4.7.2. Health Outcomes .....58
  - 4.7.3. Subgroup Analyses .....60
- 4.8. Interim Analyses .....61
  - 4.8.1. Unblinding Plan .....61
- 4.9. Changes to Protocol-Planned Analyses .....61
- 5. Sample Size Determination .....62
- 6. Supporting Documentation .....63
  - 6.1. Appendix 1: Clinical Trial Registry Analyses .....63
  - 6.2. Appendix 2: Search Criteria for Special Safety Topics .....63
  - 6.3. Appendix 3: Magnetic Resonance Imaging Addendum .....65
  - 6.4. Appendix 4: Statistical Analysis for China .....66
  - 6.5. Appendix 5: Statistical Analysis for Japan .....66
- 7. References .....68

## Version History

This statistical analysis plan (SAP) is the second version and is based on amendment (c) of the protocol for I8F-MC-GPIF (GPIF) approved on 02 June 2023. This SAP was approved prior to the first unblinding of the treatment assignments for the primary outcome lock.

**Table GPIF.1.1. SAP Version History Summary**

SAP Version	Approval Date	Change	Rationale
1	12 January 2023	Not Applicable	Original version
2	17 October 2023	<p><b>Section 1.1:</b></p> <ul style="list-style-type: none"> <li>• Revised the primary endpoint from percent change in AHI to change in AHI and added percent change in AHI to key secondaries.</li> <li>• Moved hypoxic burden from secondary to key secondary endpoint.</li> <li>• Moved FOSQ to secondary from key secondary endpoints and added PROMIS score related endpoints to key secondary.</li> </ul>	<ul style="list-style-type: none"> <li>• Changed to align with regulatory recommendation.</li> <li>• Change made due to the increasing importance of hypoxic burden in OSA disease state.</li> <li>• Changed to reflect regulatory recommendation.</li> </ul>
		<p><b>Section 1.1.1:</b></p> <ul style="list-style-type: none"> <li>• Added language to clarify the population and intercurrent events for the estimands.</li> </ul>	<ul style="list-style-type: none"> <li>• Added for clarification in alignment with regulatory feedback.</li> </ul>
		<p><b>Section 2.1:</b></p> <ul style="list-style-type: none"> <li>• Added detailed multiplicity control scheme for controlling Type 1 error.</li> </ul>	<ul style="list-style-type: none"> <li>• Details of Type 1 error control provided as planned in the protocol.</li> </ul>

SAP Version	Approval Date	Change	Rationale
		<p><b>Section 3:</b></p> <ul style="list-style-type: none"> <li>• Modified language on analysis sets to clarify the definition of the analysis sets and population.</li> </ul>	<ul style="list-style-type: none"> <li>• Analysis set definitions updated for clarity.</li> </ul>
		<p><b>Section 4.1:</b></p> <ul style="list-style-type: none"> <li>• Updated the definition of baseline and postbaseline measures for safety analyses.</li> <li>• Updated baseline and postbaseline definition for PRO measures.</li> </ul>	<ul style="list-style-type: none"> <li>• Changed to minimize missing baseline data relevant to dosing.</li> <li>• Updated to reflect the proper collection time of PROs associated with a PSG measurement. Added a 7-day window for each PRO visit to minimize missing data.</li> </ul>
		<p><b>Section 4.1.2:</b></p> <ul style="list-style-type: none"> <li>• Updated the intercurrent events in Table GPIF.4.2</li> </ul>	<ul style="list-style-type: none"> <li>• Change made to clarify the definition of intercurrent events.</li> </ul>
		<p><b>Section 4.1.6:</b></p> <ul style="list-style-type: none"> <li>• Removed analysis of changes to baseline medication in postrandomization (in term of type/class):                             <ul style="list-style-type: none"> <li>○ lipid lowering therapy, and</li> <li>○ antihypertensive therapy.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Data not collected.</li> </ul>
		<p><b>Section 4.3.2:</b></p> <ul style="list-style-type: none"> <li>• Added option to include interaction term between treatment and covariates in ANCOVA model.</li> </ul>	<ul style="list-style-type: none"> <li>• Change made in alignment with industry guidance for handling covariates.</li> </ul>

SAP Version	Approval Date	Change	Rationale
		<b>Section 4.4.1:</b> <ul style="list-style-type: none"> <li>• Updated primary endpoint.</li> </ul>	<ul style="list-style-type: none"> <li>• Changed to align with the endpoint change in Section 1.1.</li> </ul>
		<b>Section 4.4.1:</b> <ul style="list-style-type: none"> <li>• Updated key secondary endpoints.</li> </ul>	<ul style="list-style-type: none"> <li>• Changed to align with the endpoint change in Section 1.1.</li> </ul>
		<b>Section 4.4.1.1:</b> <ul style="list-style-type: none"> <li>• Removed hierarchical endpoint for FOSQ and added hierarchical endpoint for PROMIS score.</li> <li>• Added an option to impute missing baseline PRO with multiple imputation.</li> </ul>	<ul style="list-style-type: none"> <li>• Changed to align with the endpoint change in Section 1.1.</li> <li>• Added to mitigate the effect of missing data on PRO measurements.</li> </ul>
		<b>Section 4.4.1.3:</b> <ul style="list-style-type: none"> <li>• Added an option to use tipping point analysis as a sensitivity analysis.</li> </ul>	<ul style="list-style-type: none"> <li>• Changed to align general regulatory recommendations for handling missing data.</li> </ul>
		<b>Section 4.4.3:</b> <ul style="list-style-type: none"> <li>• Added analysis methods for secondary endpoints not controlled for Type 1 error.</li> </ul>	<ul style="list-style-type: none"> <li>• Changed to align with the endpoint change in Section 1.1.</li> </ul>
		<b>Section 4.6.3.3.2:</b> <ul style="list-style-type: none"> <li>• Updated analysis for hepatic safety.</li> </ul>	<ul style="list-style-type: none"> <li>• Changed to align to the standardized analysis approach across tirzepatide indications.</li> </ul>
		<b>Section 4.6.3.4:</b> <ul style="list-style-type: none"> <li>• Updated severe hypoglycemia definition.</li> </ul>	<ul style="list-style-type: none"> <li>• Updated to align with the protocol.</li> </ul>

SAP Version	Approval Date	Change	Rationale
		<p><b>Section 4.6.3.5.4:</b></p> <ul style="list-style-type: none"> <li>Removed specific immunogenicity analyses.</li> </ul>	<ul style="list-style-type: none"> <li>Some immunogenicity analyses are planned as an integrated summary instead of at individual study level and are removed.</li> </ul>
		<p><b>Section 6.4:</b></p> <ul style="list-style-type: none"> <li>Added Appendix 4.</li> </ul>	<ul style="list-style-type: none"> <li>Added to prespecify statistical analyses for China subpopulation.</li> </ul>
		<p><b>Section 6.5:</b></p> <ul style="list-style-type: none"> <li>Added Appendix 5.</li> </ul>	<ul style="list-style-type: none"> <li>Added to prespecify statistical analyses for Japan subpopulation.</li> </ul>
3	28 February 2024	<p><b>Section 1.1:</b></p> <ul style="list-style-type: none"> <li>Moved the hierarchical combination of PROMIS endpoints from key secondary to other secondary.</li> <li>Added change in FOSQ (30-item) total score in other secondary endpoints.</li> </ul>	<ul style="list-style-type: none"> <li>Changed per FDA recommendation.</li> <li>Added to clarify that the overall as well as by domain analysis will be performed.</li> </ul>
		<p><b>Section 1.1.1:</b></p> <ul style="list-style-type: none"> <li>Identified the treatment regimen estimand as the primary estimand for marketing application.</li> <li>Data from inadvertent enrollees are to be included in primary and key secondary analysis.</li> </ul>	<ul style="list-style-type: none"> <li>Clarified per FDA suggestion.</li> <li>Changed per FDA recommendation.</li> </ul>

SAP Version	Approval Date	Change	Rationale
		<ul style="list-style-type: none"> <li>Updated the subsections titled “Handling of intercurrent events” under estimand definition.</li> </ul>	<ul style="list-style-type: none"> <li>Clarified language per FDA feedback.</li> </ul>
		<p><b>Section 2.1:</b></p> <ul style="list-style-type: none"> <li>Updated graphical testing strategy and Figure GPIF.2.1.</li> </ul>	<ul style="list-style-type: none"> <li>Graphical testing strategy revised per FDA recommendation. PROMIS related endpoints removed from individual study graph and included in the integrated efficacy analysis subject to submission wide type 1 error rate control.</li> </ul>
		<p><b>Section 3:</b></p> <ul style="list-style-type: none"> <li>Changed the definitions for data point sets.</li> </ul>	<ul style="list-style-type: none"> <li>Changed to align with changes in Section 1.1.1.</li> </ul>
		<p><b>Section 4.1:</b></p> <ul style="list-style-type: none"> <li>Added language on baseline AHI.</li> <li>Modified the postbaseline definition for PRO measures.</li> <li>Updated that geographic region will be used in lieu of pooled country as a covariate in analysis models.</li> <li>Changed AHI groups to be included in the model as covariates.</li> </ul>	<ul style="list-style-type: none"> <li>Language added for clarification.</li> <li>Updated to clarify the definition.</li> <li>Changed to reduce the number of strata in the model and to keep enough participants in each covariate strata.</li> <li>Changed the group category from “moderate” to “not severe” to incorporate inadvertently enrolled participants.</li> </ul>

SAP Version	Approval Date	Change	Rationale
		<p><b>Section 4.1.2:</b></p> <ul style="list-style-type: none"> <li>Updated missing data and imputation method in Table GPIF.4.2.</li> </ul>	<ul style="list-style-type: none"> <li>Missing data imputation algorithm revised, in part to include inadvertent enrollees in analysis, and language on missing data clarified per FDA feedback.</li> </ul>
		<p><b>Section 4.1.2:</b></p> <ul style="list-style-type: none"> <li>Added no OSA and mild categories to baseline OSA category</li> </ul>	<ul style="list-style-type: none"> <li>Added to account for inadvertently enrolled patients</li> </ul>
		<p><b>Section 4.1.6:</b></p> <ul style="list-style-type: none"> <li>Added definition for baseline and postbaseline concomitant medication use.</li> </ul>	<ul style="list-style-type: none"> <li>Clarified definition for related tables and listings.</li> </ul>
		<p><b>Section 4.2:</b></p> <ul style="list-style-type: none"> <li>Added definition for participant study disposition.</li> </ul>	<ul style="list-style-type: none"> <li>Clarified the definition of participant study disposition based on collected CRF data.</li> </ul>
		<p><b>Section 4.4.1.2:</b></p> <ul style="list-style-type: none"> <li>Added analysis for change in log hypoxic burden.</li> <li>Added reporting of unconditional risk difference from logistic regression.</li> </ul>	<ul style="list-style-type: none"> <li>Log scale is deemed appropriate for analysis as the measure is an area under the curve.</li> <li>Added to provide additional measure of treatment effect in accordance to FDA guidance.</li> </ul>



SAP Version	Approval Date	Change	Rationale
		<p><b>Section 4.6:</b></p> <ul style="list-style-type: none"> <li>Added hepatobiliary events, malignancies, and abuse potential as special safety topics.</li> </ul>	<ul style="list-style-type: none"> <li>Updated to provide information on special safety topics.</li> </ul>
		<p><b>Section 4.6.1.1:</b></p> <ul style="list-style-type: none"> <li>Clarified definition for TEAEs.</li> </ul>	<ul style="list-style-type: none"> <li>Clarification</li> </ul>
		<p><b>Section 4.6.6:</b></p> <ul style="list-style-type: none"> <li>Added information on product complaints</li> </ul>	<ul style="list-style-type: none"> <li>Added in accordance with regulatory requirements for devices or combination products</li> </ul>
		<p><b>Section 4.7.2.4:</b></p> <ul style="list-style-type: none"> <li>Updated method of calculation for PROMIS T-scores</li> </ul>	<ul style="list-style-type: none"> <li>Updated to calculate T-scores using a response pattern scoring approach in accordance with FDA feedback.</li> </ul>
		<p><b>Section 4.7.2.5:</b></p> <ul style="list-style-type: none"> <li>Updated method of calculation for PROMIS T-scores</li> </ul>	<ul style="list-style-type: none"> <li>Updated to calculate T-scores using a response pattern scoring approach in accordance with FDA feedback.</li> </ul>
		<p><b>Section 5:</b></p> <ul style="list-style-type: none"> <li>Added sample size determination from protocol</li> </ul>	<ul style="list-style-type: none"> <li>Added in accordance with FDA feedback</li> </ul>
		<p><b>Section 6.3:</b></p> <ul style="list-style-type: none"> <li>Added language that no imputation will be performed for MRI data.</li> </ul>	<ul style="list-style-type: none"> <li>Clarification</li> </ul>

SAP Version	Approval Date	Change	Rationale
		<p><b>Throughout the document</b></p> <ul style="list-style-type: none"> <li>• Changed GPI1 to ISA1 and GPI2 to ISA2</li> </ul>	<ul style="list-style-type: none"> <li>• Changed for consistency</li> </ul>
4	See Date on Page 1	<p><b>Section 3:</b></p> <p>Modified the definition of the data points set for AHI and PRO data associated with AHI.</p>	<ul style="list-style-type: none"> <li>• To include data taken outside the anticipated window for data capture or delayed data entry. Approach justified due to sustained pharmacodynamic effect.</li> </ul>
		<p><b>Section 4.1:</b></p> <p>Updated the postbaseline definition for Visit 11 (Week 52) for PRO measures associated with PSG (ESS, FOSQ, PROMIS, PGIS, SF-36v2 acute form, and EQ-5D-5L).</p>	<ul style="list-style-type: none"> <li>• To maximize the utilization of collected data while accounting for the effect of PAP withdrawal prior to Week 52 PSG visits for patients in ISA2.</li> </ul>
		<p><b>Section 4.1.2</b></p> <ul style="list-style-type: none"> <li>• Added a method of imputation for missing data from participants completing treatment period on study drug intervention.</li> <li>• Added multiple imputation assuming MAR for analysis of categorical data aligned to efficacy estimand.</li> </ul>	<ul style="list-style-type: none"> <li>• To address an unanticipated missing data pattern.</li> <li>• Changed to align with the hypothetical strategy for handling intercurrent events for the efficacy estimand.</li> </ul>
		<p><b>Section 4.1.6</b></p> <ul style="list-style-type: none"> <li>• Changed “utilization after randomization” to “utilization after baseline”</li> </ul>	<ul style="list-style-type: none"> <li>• Changed to be consistent with reporting other concomitant medications</li> </ul>

SAP Version	Approval Date	Change	Rationale
		<p>for antihyperglycemic medication for the treatment of diabetes for participants who develop type 2 diabetes mellitus during the study, for the use of antidiarrheal medication, and for the use of antiemetic medication.</p>	
		<p><b>Section 4.4.1.1:</b>                      Changed log SASHB value for analysis when SASHB = 0 from log(1) to log (0.01)</p>	<ul style="list-style-type: none"> <li>• Change made to ensure consistency with observed minimum SASHB values.</li> </ul>
		<p><b>Section 4.4.3:</b>                      In Table GPIF.4.3., populated the threshold values for calculating proportion of Proportion of participants achieving clinically meaningful within-patient change in PROMIS scores.</p>	<ul style="list-style-type: none"> <li>• Threshold values updated as they are available now after blinded interim analysis of psychometric data.</li> </ul>

## Abbreviations and Definitions

<b>Term</b>	<b>Definition</b>
<b>ADA</b>	anti-drug antibody
<b>AE</b>	adverse event
<b>AESI</b>	adverse event of special interest
<b>AHI</b>	Apnea-Hypopnea Index
<b>ANCOVA</b>	analysis of covariance
<b>BG</b>	blood glucose
<b>BMI</b>	body mass index
<b>CRF</b>	case report form
<b>CSR</b>	clinical study report
<b>C-SSRS</b>	Columbia-Suicide Severity Rating Scale
<b>EAS</b>	efficacy analysis set
<b>eCRF</b>	electronic case report form
<b>EQ-5D-5L</b>	EuroQoL-5 Dimension-5 Level
<b>ESS</b>	Epworth Sleepiness Scale
<b>EudraCT</b>	European Union Drug Regulating Authorities Clinical Trials
<b>FAS</b>	Full Analysis Set
<b>FDA</b>	United States Food and Drug Administration
<b>FOSQ</b>	Functional Outcomes of Sleep Questionnaire
<b>FOSQ-10</b>	Functional Outcomes of Sleep Questionnaire, 10 items
<b>GI</b>	gastrointestinal
<b>GIP</b>	glucose-dependent insulintropic polypeptide
<b>GIPR</b>	glucose-dependent insulintropic polypeptide receptor
<b>GLP-1</b>	glucagon-like peptide-1
<b>GLP-1R</b>	glucagon-like peptide-1 receptor
<b>HLT</b>	High Level Term

<b>Term</b>	<b>Definition</b>
<b>ISA</b>	intervention-specific appendix
<b>ISR</b>	Injection site reaction
<b>JASSO</b>	Japan Society for the Study of Obesity
<b>Lilly</b>	Eli Lilly and Company
<b>LLT</b>	Lowest Level Term
<b>MACE</b>	major adverse cardiovascular event(s)
<b>MedDRA</b>	Medical Dictionary for Regulatory Activities
<b>mITT</b>	modified intent-to-treat
<b>MMRM</b>	mixed model repeated measures
<b>MRD</b>	minimum required dilution
<b>MRI</b>	magnetic resonance imaging
<b>MTD</b>	maximum tolerated dose
<b>Nab</b>	neutralizing antibodies
<b>Nab-</b>	neutralizing antibody negative
<b>Nab+</b>	neutralizing antibody positive
<b>Nab LY</b>	tirzepatide
<b>OSA</b>	obstructive sleep apnea
<b>OSAS</b>	obstructive sleep apnea syndrome
<b>OUS</b>	outside of the United States
<b>PAP</b>	positive airway pressure
<b>PGIC</b>	Patient Global Impression of Change
<b>PGIS</b>	Patient Global Impression of Status
<b>PHQ-9</b>	Patient Health Questionnaire
<b>PK</b>	pharmacokinetic
<b>PRO</b>	patient-reported outcome
<b>PROMIS</b>	Patient-Reported Outcomes Measurement Information System

---

<b>Term</b>	<b>Definition</b>
<b>PSG</b>	polysomnography
<b>PT</b>	Preferred Term
<b>REML</b>	restricted maximum likelihood
<b>SAE</b>	serious adverse event
<b>SAP</b>	statistical analysis plan
<b>SBP</b>	systolic blood pressure
<b>SF-36v2</b>	Short-Form 36 version 2
<b>SMQ</b>	Standardized MedDRA Query
<b>SOC</b>	System Organ Class
<b>TE ADA</b>	treatment-emergent anti-drug antibody
<b>TE ADA-</b>	treatment-emergent anti-drug antibody negative
<b>TE ADA+</b>	treatment-emergent anti-drug antibody positive
<b>TEAE</b>	treatment-emergent adverse event
<b>UACR</b>	urine albumin-to-creatinine ratio
<b>ULN</b>	upper limit of normal

---

# 1. Introduction

## 1.1. Objectives, Endpoints, and Estimands

Objective	Endpoints
<b>Primary</b>	
To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for decrease in AHI.	Change in AHI from baseline to Week 52.
<b>Key Secondary (controlled for Type 1 error)</b>	
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Percent change in AHI</li> <li>• Clinically meaningful change in AHI</li> <li>• Achieving OSA remission or mild non-symptomatic OSA</li> <li>• Change in body weight</li> <li>• Change in inflammatory status</li> <li>• Hypoxic burden</li> <li>• Change in PROs</li> <li>• Change in SBP</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Percent change in AHI</li> <li>• Percent of participants with <math>\geq 50\%</math> AHI reduction</li> <li>• Percent of participants with                             <ul style="list-style-type: none"> <li>○ AHI <math>&lt; 5</math> or</li> <li>○ AHI 5-14 with ESS <math>\leq 10</math></li> </ul> </li> <li>• Percent change in body weight</li> <li>• Change in hsCRP concentration</li> <li>• Change in SASHB (% min/hour)</li> <li>• Change in<sup>a</sup>:                             <ul style="list-style-type: none"> <li>○ PROMIS Sleep-related impairment short form 8a</li> <li>○ PROMIS Sleep disturbance short form 8b</li> </ul> </li> </ul> <p>From baseline to Week 48<sup>b</sup></p> <ul style="list-style-type: none"> <li>• Change in SBP</li> </ul>
<b>Other Secondary</b>	
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in excessive daytime sleepiness</li> <li>• Change in patient-reported functional status as assessed by FOSQ (30 items)</li> <li>• Change in body weight</li> <li>• Change in lipid parameters</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in ESS score</li> <li>• Change in FOSQ-10 score</li> <li>• Change in FOSQ (30 items) Score</li> <li>• Change in all FOSQ domain scores, specifically                             <ul style="list-style-type: none"> <li>○ General Productivity</li> <li>○ Activity level</li> <li>○ Vigilance</li> <li>○ Social outcomes</li> <li>○ Intimate and sexual relationships</li> </ul> </li> <li>• Percent of participants who achieve                             <ul style="list-style-type: none"> <li>○ <math>\geq 10\%</math> body weight reduction</li> <li>○ <math>\geq 15\%</math> body weight reduction</li> <li>○ <math>\geq 20\%</math> body weight reduction</li> </ul> </li> <li>• Change in                             <ul style="list-style-type: none"> <li>○ HDL-cholesterol</li> <li>○ non-HDL-cholesterol</li> <li>○ triglycerides</li> </ul> </li> </ul>

Objective	Endpoints
<ul style="list-style-type: none"> <li>• A hierarchical assessment of PRO change</li> <li>• Change in supportive secondary PROs</li> <li>• Insulin</li> <li>• Change in DBP</li> </ul>	<ul style="list-style-type: none"> <li>• A hierarchical combination of the following:                             <ul style="list-style-type: none"> <li>○ Change in PROMIS Sleep-related impairment short form 8a</li> <li>○ Change in PROMIS Sleep disturbance short form 8b</li> </ul> </li> <li>• Change in:                             <ul style="list-style-type: none"> <li>○ SF-36v2 acute form domain and summary scores</li> </ul> </li> <li>• Percent of participants with improved categorical shift in:                             <ul style="list-style-type: none"> <li>○ PGIS-OSA Sleepiness</li> <li>○ PGIS-OSA Fatigue</li> <li>○ PGIS-OSA Snoring</li> </ul> </li> <li>• Proportion of participants achieving clinically meaningful within-patient change in:                             <ul style="list-style-type: none"> <li>○ PROMIS Sleep-related impairment</li> <li>○ PROMIS Sleep disturbance</li> </ul> </li> <li>• Change in fasting insulin</li> </ul> <p>From baseline to Week 48<sup>a</sup></p> <ul style="list-style-type: none"> <li>• Change in DBP</li> </ul>
<b>Exploratory</b>	
<p>To demonstrate that tirzepatide at the MTD (10 mg or 15 mg) QW is superior to placebo for</p> <ul style="list-style-type: none"> <li>• Change in exploratory PROs</li> <li>• To evaluate the effect of tirzepatide on sleep parameters as measured by Actigraphy (AX6)</li> </ul>	<p>From baseline to Week 52</p> <ul style="list-style-type: none"> <li>• Change in                             <ul style="list-style-type: none"> <li>○ EQ-5D-5L utility index</li> <li>○ EQ-VAS scores</li> </ul> </li> <li>• Percent of participants with improved categorical shift in:                             <ul style="list-style-type: none"> <li>○ PGIC-OSA Sleepiness</li> <li>○ PGIC-OSA Fatigue</li> <li>○ PGIC-OSA Sleep quality</li> <li>○ PGIC-OSA Snoring</li> </ul> </li> <li>• Change from baseline to endpoint assessment in                             <ul style="list-style-type: none"> <li>○ Daytime sleep duration</li> <li>○ Daily step counts</li> <li>○ Average acceleration</li> </ul> </li> </ul>

Abbreviations: AHI = Apnea-Hypopnea Index; AX6 = Axivity 6; BP = blood pressure; DBP = diastolic blood pressure; ESS = Epworth Sleepiness Scale; EQ-5D-5L = EuroQol-5 Dimension-5-Level; EQ-VAS = EuroQol Visual Analogue Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; HDL = high-density lipoprotein; hsCRP = high-sensitivity C-reactive protein; MTD = maximum tolerated dose; OSA = obstructive sleep apnea; PAP = positive airway pressure; PGIC-OSA = Patient Global Impression of Change – Obstructive Sleep Apnea; PGIS-OSA = Patient Global Impression of Status – Obstructive Sleep Apnea; PRO = patient-reported outcome; PROMIS = Patient-Reported Outcomes Measurement Information System; QW = once weekly; SASHB = sleep apnea-specific hypoxic burden; SBP = systolic blood pressure; SF-36v2 = Short-Form 36 version 2.

<sup>a</sup> Subject to submission wide type 1 error rate control (Vandemeulebroecke et al. 2024).

<sup>b</sup> BP will be assessed at Week 48 because PAP withdrawal at Week 52 may confound BP assessment.



### 1.1.1. Estimands

#### Primary estimands

The primary and each key secondary efficacy analysis will be guided by the “*treatment regimen*” estimand and the “*efficacy*” estimand to support global regulatory submissions and publications. The “*efficacy*” estimand provides an on-treatment assessment of efficacy without confounding the treatment effect from the data collected after treatment discontinuation. It represents on-treatment efficacy. The “*treatment regimen*” estimand estimates the treatment effect, including the effect of intervention discontinuation to reflect clinical practice. It represents the efficacy irrespective of adherence to study intervention. The “*treatment regimen*” estimand will be used as the primary estimand to support a marketing application for the FDA.

#### Efficacy estimand

The clinical question of interest for the efficacy estimand is the treatment difference between tirzepatide and placebo after 52 weeks of intervention in treated participants with obesity and OSA, prior to study intervention discontinuation for any reason.

#### Efficacy estimand attributes

- *Population*: Adult participants with obesity and OSA who received at least 1 dose of study treatment.
- *Treatment condition*: On randomized treatment.
- *Endpoints*: The primary and key secondary endpoints will be studied. Further details on the endpoints can be found in the Objectives and Endpoints table (Section 1.1).
- *Population level summary*: The difference in mean change from baseline to 52 weeks will be used for continuous endpoints; the difference in proportion (absolute or relative, as appropriate) will be used for dichotomous endpoints. The population level summary will be conducted using the EAS described in Section 3.
- *Handling of intercurrent events*: The intercurrent events of treatment discontinuation and use of PAP therapy for participants in ISA1 is addressed by the hypothetical strategy. The potential outcome of interest is the response in the efficacy measurement if participants would remain on their randomly assigned treatment for 52 weeks and would not initiate PAP therapy during the study.
- *Rationale*: The efficacy estimand provides an on-treatment assessment without confounding the treatment effect from off-treatment data.

#### Treatment regimen estimand

The clinical question of interest for the treatment regimen estimand is the treatment difference between tirzepatide and placebo after 52 weeks of intervention in treated participants with obesity and OSA, regardless of intervention discontinuation for any reason.

**Treatment regimen estimand attributes**

- *Population:* Adult participants with obesity and OSA who received at least 1 dose of study treatment.
- *Treatment condition:* On- or off-randomized-treatment.
- *Endpoints:* The primary and key secondary endpoints will be studied. Further details on the endpoints are in the Objectives and Endpoints table (Section 1.1).
- *Population level summary:* The difference in mean change from baseline to 52 weeks will be used for continuous endpoints and the difference in proportion (absolute or relative, as appropriate) will be used for dichotomous endpoints. The population level summary will be conducted using the FAS described in Section 3.
- *Handling of intercurrent events:* No intercurrent events since treatment adherence and the initiation of PAP therapy are part of the treatment condition. Methods to handle missing data are described in detail in Section 4.1.2.
- *Rationale:* The treatment regimen estimand estimates treatment effect, including the effect of intervention discontinuation to reflect clinical practice. It is used for submission and registration purpose with regulatory agencies.

Efficacy and treatment regimen estimands will be evaluated for key secondary objectives similarly to the primary objectives.

**Safety estimand**

The clinical interest for safety estimands is the safety assessment of individual treatment arms up to the end of safety follow-up or study discontinuation in participants with obesity and OSA, from all randomly assigned participants who are exposed to at least 1 dose of study intervention, regardless of adherence to study intervention.

**Safety estimand attributes**

- *Population:* Adult participants with obesity and OSA who received at least 1 dose of study treatment.
- *Treatment condition:* On- or off-randomized-treatment.
- *Endpoints:* Endpoints corresponding to the safety analyses described in Section 4.6.
- *Population level summary:* Population level summaries will be conducted using the safety analysis set described in Section 3.
- *Intercurrent events:* Potential intercurrent events may lead to study discontinuation or missing data due to a technical or scheduling issue, but there are no planned approaches for accommodating intercurrent events.

## 1.2. Study Design

Study I8F-MC-GPIF (GPIF) is a multicenter, randomized, parallel-arm, double-blind, placebo-controlled Phase 3 study to evaluate the efficacy and safety of tirzepatide at the MTD (10 mg or 15 mg) once weekly versus placebo in participants who have obesity and moderate to severe OSA.

This basket-type master protocol will investigate 2 participant populations, described in 2 ISAs:

- ISA1 will include participants who are unwilling or are unable to use PAP therapy.
- ISA2 will include participants who have been on PAP therapy for at least 3 consecutive months prior to Visit 1 and plan to continue PAP therapy during the study.

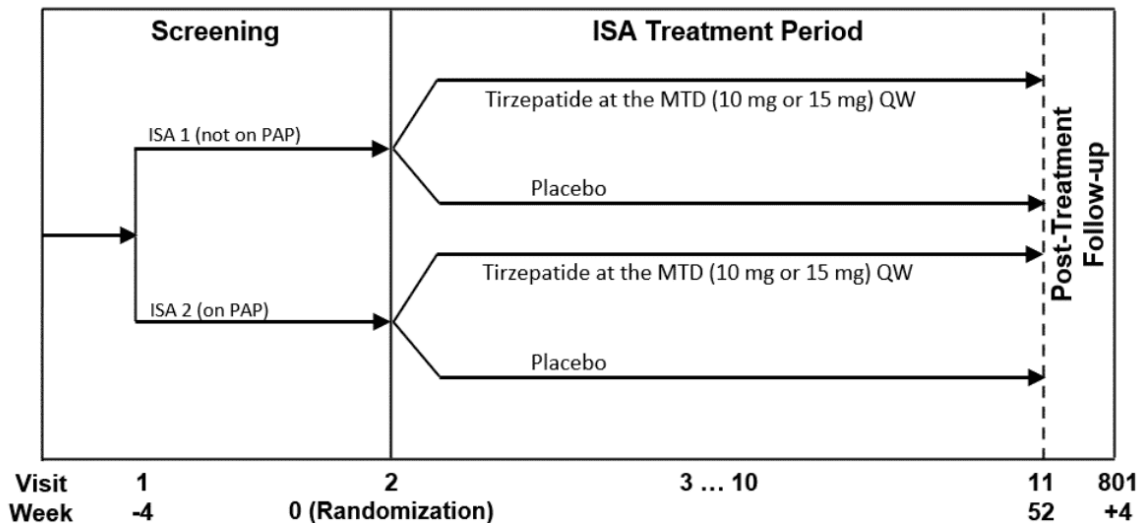
Participants to be assigned to whichever ISA they qualify for. Participants will then be randomly assigned to:

- tirzepatide at the MTD (10 mg or 15 mg) subcutaneous once weekly, or
- placebo.

The expected total duration of study participation for each participant, including screening and the posttreatment follow-up periods, is 60 weeks across the following study periods:

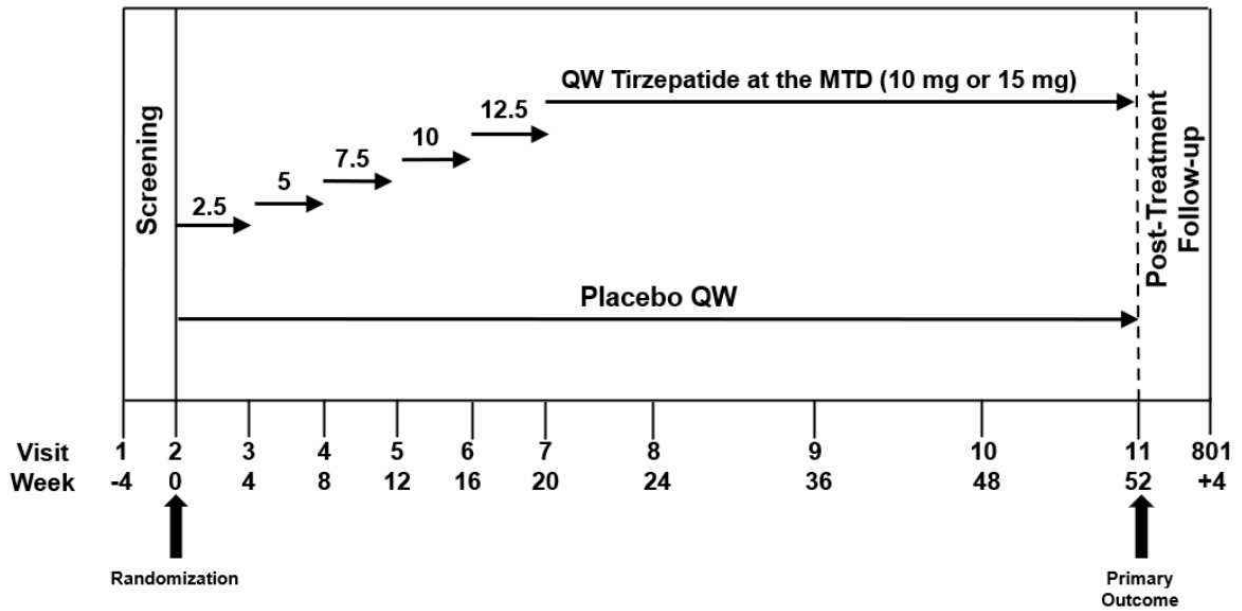
- Screening: 4 weeks
- ISA Treatment Period: 52 weeks
- Post-Treatment Follow-up Period: 4 weeks

The maximum duration of treatment is 52 weeks.



Abbreviations: ISA = intervention-specific appendix; MTD = maximum tolerated dose; PAP = positive airway pressure; QW = once weekly.

**Figure GPIF.1.1. Illustration of master protocol design for Clinical Protocol I8F-MC-GPIF.**



**Figure GPIF.1.2. Illustration of dose escalation and visit schema for Clinical Protocol I8F-MC-GPIF.**

## 2. Statistical Hypotheses

For each ISA, the primary objective is to demonstrate that tirzepatide at the MTD (10 mg or 15 mg) is superior to placebo in treating participants with OSA with respect to the change in AHI. Thus, the null and alternative hypotheses will be defined as below.

Null hypothesis: tirzepatide at the MTD (10 mg or 15 mg) is not different from the placebo with respect to the mean change from baseline in AHI at 52 weeks.

Alternative hypothesis: tirzepatide at the MTD (10 mg or 15 mg) is superior to the placebo with respect to the mean change from baseline in AHI at 52 weeks.

The treatment effect will be defined as the difference between the estimates of the mean change from baseline at 52 weeks for tirzepatide at the MTD (10 mg or 15 mg) and placebo.

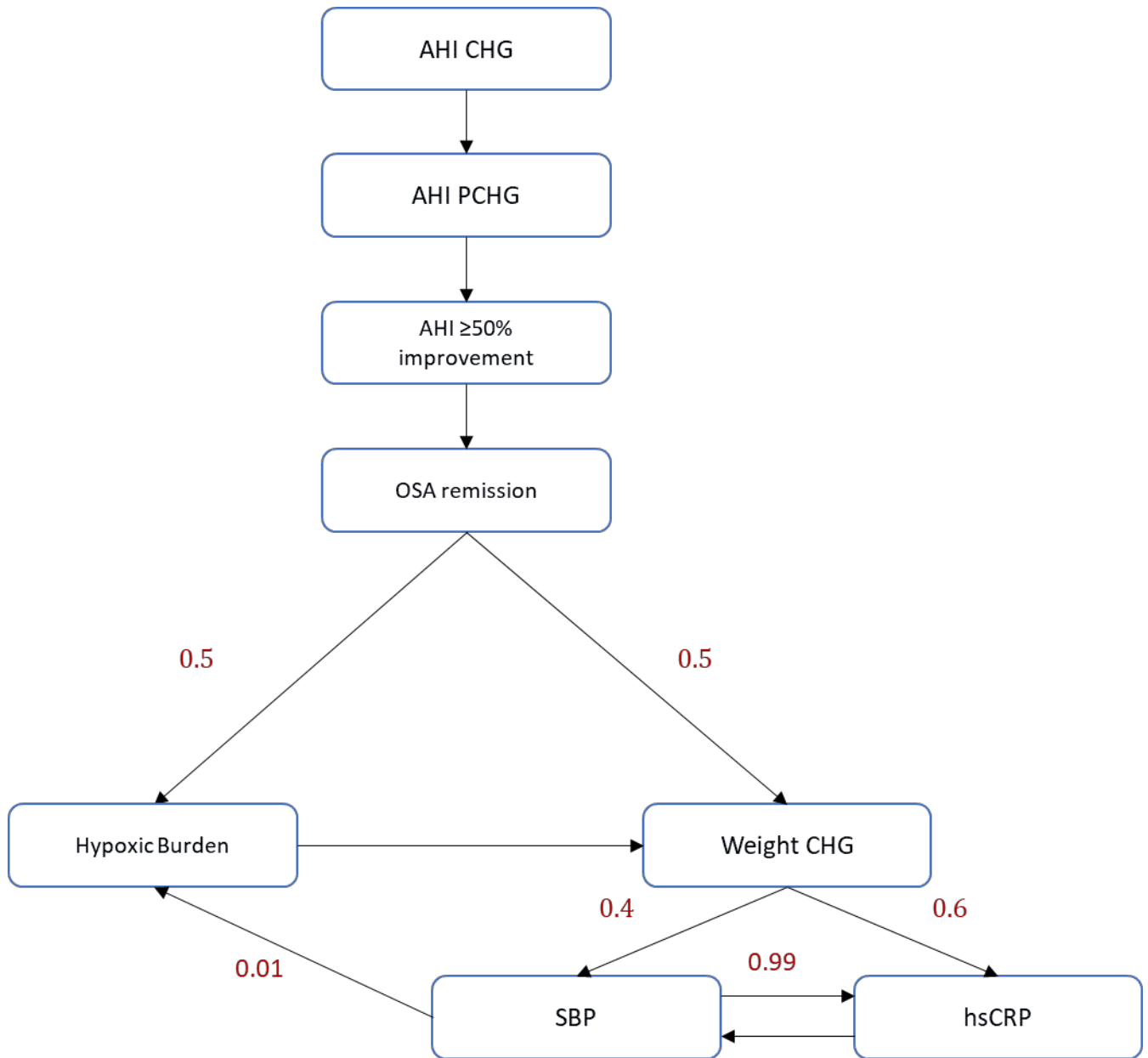
Operationally, the hypotheses will be evaluated by 2-sided tests.

### 2.1. Multiplicity Adjustment

Multiplicity adjusted analyses will be performed on the primary and key secondary objectives to control the overall family-wise Type 1 error rate at a 2-sided alpha level of 0.05 within each ISA. The graphical multiple testing procedure described in Bretz et al. (2009, 2011) will be used. This approach is a closed testing procedure; hence, it strongly controls the family-wise error rate across all hypotheses (Alosh et al. 2014).

Figure GPIF.2.1 provides the details of the graphical multiple testing procedure. Because the 2 types of estimands (treatment regimen and efficacy estimands) are intended for distinct purposes, no multiplicity adjustment will be made for conducting separate analyses on the same objectives. Unless otherwise specified, there will be no multiplicity adjustments for evaluating exploratory objectives and safety assessments.

Analysis for change in PROMIS Sleep-related impairment short form 8a and PROMIS Sleep disturbance short form 8b is specified in the integrated efficacy analysis plan to be tested subject to the submission wise error rate control strategy (Bretz and Xi 2019, Vandemeulebroecke et al. 2024) by conducting a pooled analysis across the 2 ISAs.



Abbreviations: AHI = Apnea-Hypopnea Index; CFB = change from baseline; CHG = change; GPIF = I8F-MC-GPIF; hsCRP = high-sensitivity C-reactive protein; OSA = Obstructive Sleep Apnea; PCHG = percentage change; SBP = systolic blood pressure.

**Figure GPIF.2.1 Graphical testing scheme for Study GPIF.**

### 3. Analysis Sets

Table GPIF.3.1 describes the populations that will be used for statistical analyses within each ISA of the master protocol. Additional intervention-specific populations for analyses may be described in the respective ISA.

**Table GPIF.3.1. Description of Analysis Population**

Analysis Population	Description
Entered	All participants who sign informed consent.
Randomized	All participants who are randomly assigned a study treatment (double-blind).
Modified intent-to-treat (mITT)	All randomized participants who are exposed to at least 1 dose of study intervention.

**Table GPIF.3.2. Description of Analysis Data Point Sets**

Analysis Set	Description
Full analysis set (FAS)	Data obtained during the treatment period of the set of participants from the mITT population, regardless of adherence to study intervention. For AHI related endpoints and PROs associated with PSG visits (ESS, FOSQ, PROMIS, PGIS, PGIC, SF-36v2 acute form, and EQ-5D-5L), data obtained outside the anticipated window through to the end of the study are included as part of the treatment period.
Efficacy analysis set (EAS)	Data obtained during the treatment period of the set of participants from the mITT population, excluding data after discontinuation of study intervention (last dose + 7 days) and for ISA1, excluding data after initiating PAP therapy. For AHI related endpoints and PROs associated with PSG visits (ESS, FOSQ, PROMIS, PGIS, PGIC, SF-36v2 acute form, and EQ-5D-5L), data obtained outside the anticipated window through the end of the study are included as part of the treatment period.
Safety analysis set (SS)	Data obtained during treatment and safety follow-up period of set of participants from the mITT population, regardless of adherence to study intervention.

Abbreviations: AHI = apnea-hypopnea index; EAS = efficacy analysis set; ESS = Epworth Sleepiness Scale; FAS = full analysis set; FOSQ = Functional Outcomes of Sleep Questionnaire; ISA = intervention-specific appendix; mITT = modified intent-to-treat; PAP = positive airway pressure; PGIC = Patient Global Impression of Change; PGIS = Patient Global Impression of Status; PROMIS = Patient-Reported Outcomes Measurement Information System; PRO = patient-reported outcome; PSG = polysomnography; SF-36v2 = Short-Form 36 version 2; SS = safety analysis set.

## 4. Statistical Analyses

### 4.1. General Considerations

Statistical analysis will be the responsibility of Eli Lilly and Company (Lilly) or its designee. Statistical analysis for each ISA will be conducted individually and a combined analysis with both ISAs is not planned. All analyses specified will apply to both ISAs unless the analysis is specified as ISA-specific.

The SAP will be finalized prior to the unblinding of the first ISA.

Some analyses and summaries described in this analysis plan may not be conducted if not warranted by data (for example, too few events to justify conducting an analysis). Additional exploratory analyses of the data will be conducted as deemed appropriate.

Efficacy analyses will be conducted on all participants randomly assigned to study intervention according to the treatment to which the participants are assigned and were exposed to at least 1 dose. For the “treatment regimen” estimand, the analysis will be conducted using the FAS. To minimize missing data, participants randomly assigned to study intervention who prematurely discontinue study treatment will be encouraged to remain in the study. However, some participants may choose to permanently discontinue from the study which will lead to missing endpoints. Details on handling missing values can be found in Section 4.1.2. For the “efficacy” estimand, the analysis will be conducted using the EAS.

Safety analysis will be conducted using the Safety Analysis Set. Selected safety analyses may be conducted after excluding the data after permanent discontinuation of the study intervention. For the safety related parameters, the definition of baseline and postbaseline are specified in [Table GPIF.4.1](#).



**Table GPIF.4.1. Baseline and Postbaseline Definition for Safety Analyses**

Analysis Set	Analysis Type	Baseline	Postbaseline
SS	1.1) Treatment-Emergent Adverse Events	The baseline period is defined as the start of screening and ends prior to the first dose of study treatment (typically at Week 0). If the first dose date is missing, then the randomization date will be used instead of first dose date.	Starts at or after the first dose of study treatment and ends at the end of the study period (including off-drug follow up visit).
SS	1.2) Treatment-Emergent Abnormal Laboratory Results <sup>a</sup> and Vital Signs	For laboratory results, baseline period is defined as prior to the first dose time and will include all scheduled and unscheduled measurements. If the first dose time is missing, then any data collected on the date of the first dose will be treated as baseline. For vital signs, baseline period is defined as measurements collected prior to the first dose. If the first dose date is missing, then the randomization date will be used instead of first dose date.	Postbaseline will be defined as after the baseline period through the end of the study participation. All scheduled and unscheduled measurements will be included.
SS	1.3) Change from Baseline for Laboratory Results <sup>a</sup> , and Vital Signs	The last scheduled and unscheduled nonmissing assessment recorded during the baseline period defined above (1.2).	Postbaseline will be defined as above (1.2). Only scheduled visits will be included. The ED visits are considered scheduled visits.

Abbreviations: ED = early discontinuation; SS = Safety Analysis Set.

<sup>a</sup> Immunogenicity related analysis is specified in Section 4.6.3.5.

For AHI analyses, baseline is defined as the last nonmissing measurement prior to the first dose.

The following paragraphs define selection of the PRO response which will be used for analysis at baseline and postbaseline visits. To select the baseline observation for PROs which are planned to be completed on the same day as the PSG (ESS, FOSQ, PROMIS, PGIS, SF-36v2 acute form, and EQ-5D-5L), the observation completed on the day or on the next day of the start of the baseline PSG will be selected. If multiple responses are completed within this period, the last response given within this timeframe will be selected. If no response was provided within this timeframe, the latest observation completed prior to the first dose will be selected. If a baseline still cannot be identified, the earliest observation within a 7-day window from the start of treatment date will be selected.

For postbaseline visits with a planned PSG measurement, the response for the PROs which are planned to be completed on the same day as the PSG (ESS, FOSQ, PROMIS, PGIS, PGIC, the SF-36v2 acute form, and EQ-5D-5L) will be selected in the following way.

**For Visit 11, for participants in ISA2 who do not permanently withdraw from PAP during the course of Study GPIF**, only the observations collected within the 6 days prior to the Visit 11 PSG start date, and the 7 days after the Visit 11 PSG start date, will be selected to minimize confounding with background PAP use. If there are multiple observations identified, then the observation closest to the Visit 11 PSG start date will be selected.

**For Visit 7 and Visit 11 for all participants in ISA1, Visit 7 for all participants in ISA2, and Visit 11 for participants in ISA2 who permanently withdraw from PAP during the course of the Study GPIF:** if available, the observation completed on the day or on the next day of the start of the PSG will be selected. If multiple responses are completed within this period, then the last response given within this timeframe will be selected. If no response was provided in this timeframe, the latest observation completed within the visit window will be selected. If a measurement for the visit still cannot be identified, then the observation taken within a 7-day window around the start of the visit date that is closest to the visit start date will be selected. If a measurement for Visit 11 (Week 52) still cannot be identified, then the observation collected within 30 days before Visit 11 start date and before the end of the study (including the follow-up period) that is closest to the Visit 11 start date will be selected.

For postbaseline visits without a planned PSG measurement, the response for the PROs (ESS, FOSQ, PROMIS, PGIS, PGIC, SF-36v2 acute form, and EQ-5D-5L) will be identified by selecting the latest observation completed within the visit window. If a measurement for the visit cannot be identified in this manner, then the observation within a 7-day window around the start of the visit date that is closest to the visit start date will be selected.

To select the baseline observation for the PHQ-9 and C-SSRS, if multiple responses are completed prior to the first dose and there are no differences in these responses, the observation completed most recently, prior to the first dose, will be selected. If multiple responses are completed prior to the first dose and there are differences in these responses, the approach differs based on the questionnaire. For the PHQ-9, the response with the worst total score will be selected; for the C-SSRS, the worst response for each question will be selected and each of the worst responses will be combined into a single response which will be used for analysis. For each postbaseline visit, the same approach to select a response in the case of multiple responses within the same visit window will be carried out.

For all other analyses, baseline is defined as the last nonmissing measurement prior to the first dose unless otherwise specified.

For AHI, if there are multiple observations for the same visit, then the later observation will be selected.

Statistical treatment comparisons will be performed between tirzepatide MTD and placebo. Unless otherwise noted, all tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, and the confidence interval will be calculated at a 2-sided 95% level. In statistical summaries and analyses, participants will be analyzed as randomized. Analysis models will use geographic region (US/OUS) as a covariate when applicable.

Analysis of covariance will be used to analyze continuous variables collected only at baseline and endpoint. Unless otherwise specified, the model will include treatment and strata (geographic region [US/OUS], AHI stratum [not severe (AHI <30), severe (AHI ≥30)], and gender) as fixed effects and baseline as a covariate.

MMRMs will be used to analyze continuous variables collected at baseline and more than 1 postbaseline visit. For the MMRM analysis, REML will be used to obtain model parameter estimates for continuous longitudinal variables. All the longitudinal observations at each scheduled postbaseline visit will be included in the analysis. The model will include the fixed class effects of treatment, strata (geographic region [US/OUS] and gender), visit, and treatment-by-visit interaction, as well as the continuous, fixed covariate of baseline value. For analyses of variables other than AHI, the AHI stratum will also be included in the model. Significance tests will be based on least squares means and Type III tests.

For continuous measures, summary statistics may include sample size, mean, standard deviation, median, minimum, and maximum for both the actual and the change from baseline measurements. Least-square means and standard errors derived from the analysis models will also be displayed for the change from baseline measurements. Treatment comparisons will be displayed showing the treatment difference least-square means and the 95% confidence intervals for the treatment differences, along with the p-values for the treatment comparisons.

For categorical measures, summary statistics may include sample size, frequency, and percentages. Fisher's exact test or Pearson's chi-square test will be used for treatment comparisons unless otherwise specified.

Not all analyses described in this SAP will necessarily be included in the CSRs. Any analysis described in this SAP and not provided in the CSR would be available upon request.

#### **4.1.1. Adjustment for Covariates**

The study is stratified by country/geographic region, OSA severity (not severe [AHI <30], severe [AHI ≥30]), and gender. Unless otherwise specified, the following factors will be adjusted for: geographic region (US/OUS), OSA severity (not severe [AHI <30], severe [AHI ≥30]), and gender. The value for stratification factors will be obtained from the data collected or derived from the eCRF or PSG results. In addition, the baseline value of the endpoint will be used as a covariate when appropriate.

#### **4.1.2. Handling of Dropouts or Missing Data**

For the primary and key secondary efficacy endpoint analyses aligned to the treatment regimen estimand and subject to Type 1 error rate control, missing data will be imputed based on the reason for the missing values, as described in [Table GPIF.4.2](#). For analyses aligned to the "efficacy" estimand, missing data will be considered missing at random and hence no explicit imputation will be performed for continuous endpoints. For categorical endpoints, the corresponding continuous variable associated with the missing categorical data will be considered missing at random, and multiple imputation assuming the data to be missing at random will be performed.

For exploratory endpoints and safety analyses, missing values will not be explicitly imputed unless specified otherwise.

For analyses aligned to the treatment regimen estimand, the statistical inference over multiple imputations will be guided by the method proposed by Rubin (1987). The missing values will be handled as follows:

**Table GPIF.4.2. Imputation Approaches to Handle Missing/Invalid Data for Treatment Regimen Estimand**

Missing/Invalid Data	Strategy to Handle Missing/Invalid Data	Assumptions for Missing Values	Methods to Handle Missing Values
Data missing at baseline, invalid data collected or missing data after treatment DC due to the COVID-19 pandemic (after other reasons for missing data are ruled out), technical issues (that is, sensor error on PSG) leading to invalid measurements ascertained while on treatment, missing data from participants completing the treatment period on the study drug intervention, or missing data after study DC due to inadvertent enrollment.	Hypothetical	MAR	Multiple imputation assuming MAR
Missing data due to any other reason (for example, study DC due to any reason other than COVID-19 or inadvertent enrollment).	Treatment policy	MNAR	Retrieved dropout imputation <sup>a</sup> . If there are not enough retrieved dropouts to provide a reliable imputation model, placebo-based multiple imputation will be used.

Abbreviations: COVID-19 = coronavirus disease-2019; DC = discontinuation; MAR = missing at random; MNAR = missing not at random; PSG = polysomnography; SBP = systolic blood pressure.

<sup>a</sup> Retrieved dropout imputation utilizes observed data from participants in the same treatment group who had outcome measures at Week 52 (or Week 48 for SBP) after early DC of study drug to impute the missing value.

**4.1.3. Multicenter Studies**

Randomization will be stratified by country, and geographic region (US/OUS) will be used as a covariate.

**4.1.4. Historical Illnesses and Preexisting Conditions**

The count and percentages of participants with historical illnesses and preexisting conditions will be summarized by treatment group using the MedDRA PTs nested within SOC. The SOC will be in alphabetical order. Conditions (that is, PTs) will be ordered by decreasing frequency in the tirzepatide MTD arm within the SOC. This will be summarized for all randomized participants.

#### **4.1.5. Participant Characteristics**

A listing of participant demographics for all randomized participants will be provided. The demographic and baseline clinical characteristics will also be summarized by study treatment for all randomized participants. Baseline demographic and clinical characteristics of special interest include but are not limited to:

- age (years)
- sex (female, male)
- race
- ethnicity
- height (cm)
- weight (kg)
- BMI (kg/m<sup>2</sup>)
- waist circumference (cm)
- age group (<50, ≥50)
- BMI group (<35, ≥35 and <40, ≥40 kg/m<sup>2</sup>)
- OSA severity (none [AHI <5], mild [AHI ≥5 and AHI <15], moderate [AHI ≥15 and AHI <30], or severe [AHI ≥30])
- geographic region (US/OUS), and
- country.

#### **4.1.6. Concomitant Therapy**

Concomitant medication will be summarized by treatment groups and displayed by decreasing frequency of WHODrug PTs in tirzepatide MTD arm. Baseline use of concomitant medication is defined as any medication started prior to the treatment start date and continuing on or after the treatment start date. Postbaseline concomitant medications are defined as those that are being taken any time during the postbaseline period.

In addition, medications of interest (as defined below) will be summarized by treatment groups:

- baseline use of:
  - lipid lowering therapy, by type/class and
  - antihypertensive therapy, by type/class
- utilization after baseline of
  - antihyperglycemic medication for the treatment of diabetes for participants who develop type 2 diabetes mellitus during the study
  - antidiarrheal medication, and
  - antiemetic medication.

In addition, for ISA2 participants only, a summary of PAP machine use at baseline and postbaseline PAP machine adherence will be provided. Further details are provided in Section 4.7.1. For ISA1 participants, a listing will be provided summarizing any participants who use a PAP machine during the course of the trial.

### **4.1.7. Treatment Exposure and Compliance**

#### **4.1.7.1. Study and Study Treatment Exposure**

Summary of duration of follow-up (defined as time in days from date of randomization to the date of the last study visit) will be provided by treatment group in the mITT population.

Summary of duration on study treatment (defined as time in days from date of first dose of study treatment to date of last dose of study treatment plus 7 days) will be provided by treatment group in the safety analysis set.

For the summary of duration on study treatment, the frequency and percentage of participants falling into the following categorical ranges will also be summarized by planned treatment group as well: >0 week,  $\geq 4$  weeks,  $\geq 8$  weeks,  $\geq 12$  weeks,  $\geq 16$  weeks,  $\geq 20$  weeks,  $\geq 24$  weeks,  $\geq 36$  weeks,  $\geq 48$  weeks, and  $\geq 52$  weeks.

No p-values will be reported in these summaries as they are intended to describe the study populations rather than test hypotheses.

#### **4.1.7.2. Adherence to Study Treatment**

Summary of prematurely discontinuing study treatment (including reason for discontinuation) will be provided by study treatment. A time-to-event analysis of premature study treatment discontinuation will also be conducted.

If data warrants, the counts and percentages of participants who follow the planned escalation scheme, have dose interruption, or have dose de-escalation will be summarized for the tirzepatide treatment group. This will include the percentage of participants who have 10 mg or 15 mg tirzepatide as their MTD. In addition, the proportion of participants receiving 2.5, 5, 7.5, 10, 12.5, or 15 mg may be presented by randomized tirzepatide treatment and visit during the dose escalation period.

Treatment adherence will be defined as taking at least 75% of the scheduled tirzepatide doses. Treatment adherence will be summarized descriptively over the treatment period by treatment using the mITT population.

### **4.1.8. Important Protocol Deviations**

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

## **4.2. Participant Dispositions**

The participant dispositions for the screening period, the study intervention/treatment period, and/or the follow-up period will be collected in CRFs with the corresponding primary reason.

The study completion for a participant is defined as the participant completing both the treatment period and the follow-up period, regardless of completion of study treatment.

Summaries and a listing of study disposition and study drug disposition will be provided for all randomized participants, separately for each ISA. Comparison between treatment arms will be performed using Fisher's exact test.

### 4.3. Primary Endpoint Analysis

The primary objective of this study is to test the hypothesis that tirzepatide at the MTD (10 mg or 15 mg) is superior to placebo for participants with moderate to severe OSA and obesity on the mean AHI reduction from baseline to Week 52. The primary and key secondary efficacy analyses will be guided by 2 estimands, the “treatment regimen” estimand and the “efficacy” estimand to support global regulatory submissions and publications.

#### 4.3.1. Analysis Related to the Efficacy Estimand

The primary analysis guided by the “efficacy” estimand will be conducted using the EAS. This analysis will be based on the contrast between tirzepatide at the MTD (10 mg or 15 mg) and placebo at Week 52 (Visit 11) from the MMRM analysis of mean change from baseline in AHI. All the longitudinal observations at each scheduled postbaseline visit will be included in the analysis. REML will be used to obtain model parameter estimates and the Kenward-Roger option will be used to estimate the denominator degrees of freedom. The response variable of the MMRM will be the change in AHI from baseline values obtained at each scheduled postbaseline AHI measurement.

The model will include the fixed class effects of treatment, strata (geographic region [US/OUS] and gender), visit, and treatment-by-visit interaction, as well as the continuous, fixed covariate of baseline AHI. An unstructured covariance structure will be used to model the within-participant errors. Significance tests will be based on least squares means and Type III tests. If this analysis fails to converge, the following covariance structures will be tested in order until convergence is achieved:

- Toeplitz with heterogeneity
- autoregressive with heterogeneity
- compound symmetry with heterogeneous variances
- Toeplitz
- autoregressive, and
- compound symmetry without heterogeneous variances.

Analysis aligned to each estimand will be evaluated at the full significance level of 0.05.

#### 4.3.2. Analysis Related to the Treatment Regimen Estimand

For the primary analysis guided by the “treatment regimen” estimand, the analysis will be conducted using the FAS. Missing values will be imputed based on the strategy to handle intercurrent events described in Section 4.1.2. After imputation, the primary efficacy comparison will be based on the contrast between tirzepatide at the MTD (10 mg or 15 mg) and placebo from the ANCOVA analysis of mean change from baseline to Week 52 in AHI using FAS. The ANCOVA model will include treatment and strata (geographic region [US/OUS] and gender) as fixed effects and baseline AHI as a fixed covariate. Statistical inference over multiple imputed data sets will be guided by Rubin (1987).

### 4.3.3. Sensitivity Analyses

For participants in ISA2, a sensitivity analysis will be carried out for the primary endpoint. When carrying out this sensitivity analysis using a treatment regimen estimand, participants with PAP withdrawal less than 5 days before the PSG at baseline or at Week 52 will have their data censored. Censored postbaseline data will be imputed using the approach outlined in Section 4.1.2. An ANCOVA model will be fit using the approaches outlined in Section 4.3.2. Carrying out this sensitivity analysis using the efficacy estimand, observations made with PAP withdrawal less than 5 days prior to PSG from the MMRM will be censored. Additional sensitivity analyses for ISA2 participants to accommodate participants with PAP withdrawal <5 days prior to the PSG may be considered.

Additional sensitivity analyses for both ISAs may be included as needed.

## 4.4. Secondary Endpoints Analysis

### 4.4.1. Key Secondary Endpoints

A graphical approach for multiple comparisons will be used to strongly control the overall Type 1 error (2-sided alpha level of 0.05) for testing the superior treatment effect of tirzepatide MTD over placebo including the key secondary endpoints as listed below.

- percent change in AHI at Week 52
- percent of participants with  $\geq 50\%$  AHI reduction at Week 52
- percent of participants at Week 52 with
  - AHI <5 or
  - (AHI 5 through 14 and ESS  $\leq 10$ )
- percent change from baseline to Week 52 in body weight
- change from baseline to Week 48 in SBP
- change from baseline to Week 52 in C-reactive protein (high-sensitivity C-reactive protein)
- change in sleep apnea-specific hypoxic burden (% minutes/hour)
- change in PROMIS Sleep-related impairment short form 8a, and
- change in PROMIS Sleep disturbance short form 8b.

Analytical approaches for the hierarchical assessment of PROs are described in Section 4.4.3.1 and a summary of the analysis approach for all other key secondary endpoints is provided in Section 4.4.1.1.

#### 4.4.1.1. Main Analytical Approaches

Analysis of percent change in AHI, percent change from baseline to Week 52 in body weight, change from baseline to Week 52 in log of high-sensitivity C-reactive protein, change from baseline to Week 48 in SBP, change from baseline to Week 52 in PROMIS Sleep-related impairment, change from baseline to Week 52 in PROMIS Sleep disturbance, and change from baseline to Week 52 in log of hypoxic burden will be conducted in a manner similar to the primary efficacy analyses using an ANCOVA model with treatment, strata (geographic region



[US/OUS], AHI stratum [not severe (AHI <30), severe (AHI ≥30)], and gender), and baseline of the corresponding variable as a covariate for the treatment regimen estimand. If the hypoxic burden is reported to be 0, log (0.01) will be used in place of the log of hypoxic burden. The analysis method utilizing data from both ISAs for change from baseline to Week 52 in PROMIS sleep Impairment and PROMIS Sleep disturbance is described in the integrated efficacy analysis plan.

For the efficacy estimand, the MMRM analyses will be conducted as described in Section 4.1. For both estimands, analysis of percent change in AHI will adjust for the continuous, fixed baseline value of AHI instead of the baseline AHI stratum (not severe, severe).

Comparisons at the 52-week visit between the treatments relative to the proportion of participants achieving ≥50% AHI reduction and AHI<5 or (AHI 5 through 14 and ESS ≤10) will be conducted using logistic regression analysis including the following terms as a covariate:

- treatment
- geographic region (US/OUS)
- baseline AHI, and
- gender.

Unconditional risk differences will also be provided for these endpoints using logistic regression (Ye et al. 2023).

Analysis aligned to each estimand will be evaluated at the full significance level of 0.05 contingent on reaching statistical significance of the primary objective.

#### **4.4.1.2. Sensitivity Analyses**

For participants in ISA2, sensitivity analyses for the key secondary endpoints will be carried out: percent change in AHI, clinically meaningful change in AHI, and achieving OSA remission/mild nonsymptomatic OSA.

For percent change in AHI, the sensitivity analysis using both the treatment regimen and efficacy estimand will be carried out. When using a treatment regimen estimand approach, participants with PAP withdrawal less than 5 days before the PSG at baseline or at Week 52 will have their data censored. Censored postbaseline data will be imputed using the approach outlined in Section 4.1.2. An ANCOVA model will be fit using the approaches outlined in Section 4.3.2. When using the efficacy estimand, observations made with a PAP withdrawal less than 5 days prior to the PSG or PROMIS from the MMRM will be censored.

For the binary endpoints of clinically meaningful change in AHI and achievement of OSA remission/mild nonsymptomatic OSA, the sensitivity analysis using both estimands will be carried out. PSG measurements taken after <5 days of PAP withdrawal will be censored. After censoring, analysis will be carried out as described in Section 4.4.1.1.

Additional sensitivity analyses for ISA2 participants to accommodate participants with PAP withdrawal <5 days prior to the PSG or PROMIS or ESS may be considered.

A 2-way tipping point analysis may also be utilized for the primary endpoint. This analysis will begin with the primary analysis aligned to the treatment regimen estimand and then adding positive and negative penalties simultaneously to both the tirzepatide MTD arm and the placebo arm, considering when results tip from superiority to inconclusive, and then considering the clinical plausibility of such scenarios.

Additional sensitivity analyses for both ISAs may be included as needed.

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

All primary and key secondary hypotheses will be tested with the overall family-wise Type 1 error rate at a 2-sided alpha level of 0.05 through the multiplicity control approach based on the graphical multiple testing procedure. The primary endpoint hypothesis will be tested at a 2-sided alpha level of 0.05 for statistical significance. If the primary efficacy endpoint is significant, the alpha of 0.05 will be propagated to the key secondary efficacy endpoints. The detailed graphical testing scheme is outlined in [Figure GPIF.2.1](#).

The analyses will be performed for both the treatment regimen and efficacy estimands described in Section [4.3](#) using the same graphical testing scheme. An overall 2-sided alpha of 0.05 to control Type 1 error rate separately for the treatment regimen estimand and the efficacy estimand will be used.

#### **4.4.3. Supportive Secondary Endpoints**

Unless otherwise specified, all supportive/other secondary efficacy analyses will be guided by the “efficacy” estimand and will be conducted using the EAS. Missing data will be handled using an MMRM.

**Table GPIF.4.3. Secondary Measures Not Controlled for Type 1 Error**

<b>Objective – Demonstrate Superiority of Tirzepatide MTD to Placebo for:</b>	<b>Endpoint</b>	<b>Analytical Approaches</b>
Change in excessive daytime sleepiness	Change in ESS score from baseline to Week 52	MMRM analysis described in Section 4.4.1.1 will be conducted.
Change in patient-reported functional status as assessed by FOSQ (10 items)	Change in FOSQ-10 total score from baseline to Week 52	MMRM analysis described in Section 4.4.1.1 will be conducted.
Change in patient-reported functional status as assessed by FOSQ (30 items)	Change in FOSQ (30 item) total score and all functional domain scores from baseline to Week 52	MMRM analysis described in Section 4.4.1.1 will be conducted.
Change in Body Weight	Percent of participants who achieve $\geq 10\%$ , $\geq 15\%$ , and $\geq 20\%$ body weight reduction.	Logistic models described in Section 4.4.1.1 with the following covariates: treatment, geographic region (US/OUS), baseline AHI (not severe/severe), gender, and baseline bodyweight as a covariate.
Change in Lipid Parameters	Change in: HDL-cholesterol non-HDL-cholesterol triglycerides	MMRM analysis described in Section 4.4.1.1 will be conducted.
Hierarchical assessment of PRO change	A hierarchical combination of the following: <ul style="list-style-type: none"> <li>o Change in PROMIS Sleep-related impairment short form 8a</li> <li>o Change in PROMIS Sleep disturbance short form 8b</li> </ul>	Win ratio analysis described in Section 4.4.3.1 will be conducted.
Change in supportive secondary PROs	Summary of item 8 of PROMIS Sleep Disturbance short form 8b	Counts and percentages of participants at each time point will be summarized by nominal visit and by treatment. A shift table from baseline to postbaseline will be created at each postbaseline visit.
	Change in: SF-36v2 acute form domain scores From baseline to Week 52	MMRM analysis of T-score described in Section 4.4.1.1 will be conducted. Description of T-score calculation provided in Sections 4.7.2.4 and 4.7.2.5.
	Percent of participants with improved categorical shift in: PGIS-OSA Sleepiness PGIS-OSA Fatigue PGIS-OSA Snoring From baseline to Week 52	For each question, the proportion of participants with improvements from baseline will be summarized. Shift analysis from baseline to Week 52 will also be performed.

Objective – Demonstrate Superiority of Tirzepatide MTD to Placebo for:	Endpoint	Analytical Approaches
	Proportion of participants who achieve: $\leq -8.0$ change in PROMIS Sleep-related impairment for ISA1 $\leq -10.0$ change in PROMIS Sleep-related impairment for ISA2 $\leq -7.5$ change in PROMIS Sleep disturbance for ISA1 and ISA2 From baseline to Week 52 (Threshold values are derived from blinded interim analysis)	Logistic models described in Section 4.4.1.1 with the following covariates: treatment, geographic region (US/OUS) baseline AHI (not severe/severe), gender, and baseline score as a covariate.
Change in Insulin	Change in fasting insulin from baseline to Week 52	MMRM analysis described in Section 4.4.1.1 will be conducted.
Change in DBP	Change in DBP from baseline to Week 48	MMRM analysis described in Section 4.4.1.1 will be conducted.

Abbreviations: AHI = Apnea-Hypopnea Index; DBP = diastolic blood pressure; ESS = Epworth Sleepiness Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; HDL = high-density lipoprotein; ISA = intervention-specific appendix; MMRM = mixed model repeated measures; MTD = maximum tolerated dose; PGIS-OSA = Patient Global Impression of Status Obstructive Sleep Apnea; PRO = patient-reported outcome; PROMIS = Patient-Reported Outcomes Measurement Information System; SF-36v2 = Short-Form 36 version 2.

#### 4.4.3.1. Hierarchical Assessment of PRO Change

The analysis of the hierarchical composite endpoint will be performed with the Finkelstein-Schoenfeld method, and the win ratio (Pocock et al. 2012) will be reported as the measure of treatment effect. The population level summary of win ratio will be calculated as number of pairs of tirzepatide-treated participant “wins” divided by number of pairs of placebo-treated participant “wins.”

The Finkelstein-Schoenfeld method is based on the principle that each tirzepatide-treated participant is compared with every other placebo-treated participant in a pairwise manner that proceeds in a hierarchical fashion. Differences will be calculated as tirzepatide participant value minus placebo participant value. Each pairwise comparison will proceed in the following order and a winner has:

- Stage 1: For the change from baseline at Week 52 PROMIS Sleep-related impairment score:
  - A comparison is a win when the treatment difference is  $\leq -4.9$
  - A comparison is a loss when the treatment difference is  $\geq 4.9$
  - All other cases are a tie and the comparison of PROMIS Sleep disturbance score will be conducted (that is, proceed to Stage 2).
- Stage 2: For the change from baseline at Week 52 PROMIS Sleep disturbance score:
  - A comparison is a win when the treatment difference is  $\leq -3.1$ .
  - A comparison is a loss when the treatment difference is  $\geq 3.1$ .

- In all other cases, the pair will be recorded as a tie.

Based on Donovan et al. (2020), a clinically important response of PROMIS sleep-related impairment for OSA participants is 4.9, and a clinically important response of PROMIS sleep disturbance for OSA participants is 3.1. Participants in this study had a mean BMI of 33.7 kg/m<sup>2</sup>, and one-half of them had moderate to severe OSA. Thus, the meaningful change threshold can be generalized to participants in the tirzepatide OSA trial.

For treatment policy estimand, missing values at Week 52 will be imputed through multiple imputations based on the reason of missingness with details described in Section 4.1.2.

For PRO measures, missing baseline values will be assumed to be missing at random and may be imputed through multiple imputation methods.

## 4.5. Exploratory Endpoint Analyses

Unless otherwise specified, all exploratory efficacy analyses will be guided by the “efficacy” estimand and will be conducted using the EAS. Missing data will be handled using an MMRM.

The following efficacy analyses apply to both ISAs and will be carried out separately for each ISA.

**Table GPIF.4.4. Exploratory Efficacy Analysis for Both ISAs**

<b>Objective – Demonstrate Superiority of Tirzepatide MTD to Placebo for:</b>	<b>Endpoint</b>	<b>Analytical Approaches</b>
Change in exploratory PROs	Change from baseline to Week 52 in: EQ-5D-5L utility index EQ-VAS scores	MMRM analysis described in Section 4.4.1.1 will be conducted.
	Percent of participants with improved categorical shift from baseline to Week 52 in: PGIC-OSA Sleepiness PGIC-OSA Fatigue PGIC-OSA Sleep quality PGIC-OSA Snoring	For each question, the proportion of participants with improvements from baseline will be summarized. Shift analysis from baseline to Week 52 will also be performed.
Change in parameters measured by Actigraphy (AX6)	Mean change from baseline to Week 52 in: Daytime sleep duration Daily step counts Average acceleration	MMRM analysis described in Section 4.4.1.1 will be conducted.

Abbreviations: EQ-5D-5L = EuroQoL-5 Dimension-5 Level; EQ-VAS = EuroQoL Visual Analogue Scale;

ISA = intervention-specific appendix; MMRM = Mixed model repeated measures; MTD = maximum tolerated dose; PRO = patient-reported outcome; PGIC-OSA = Patient Global Impression of Change Obstructive Sleep Apnea.

The efficacy analyses summarized in [Table GPIF.4.5](#) only apply to participants in ISA1.

**Table GPIF.4.5. Exploratory Efficacy Analysis Conducted only for ISA1 Participants**

<b>Objective – Demonstrate Superiority of Tirzepatide MTD to placebo for:</b>	<b>Endpoint</b>	<b>Analytical Approaches</b>
Change in parameters measured by WatchPAT300	Change from baseline to Week 52 in PAT-based device determinations of: <ul style="list-style-type: none"> <li>• pAHI</li> <li>• SASHB</li> </ul>	MMRM analysis described in Section 4.4.1.1 will be conducted.

Abbreviations: MMRM = mixed model repeated measures; MTD = maximum tolerated dose; pAHI = peripheral tone apnea-hypopnea index; PAT = peripheral arterial tonometry; SASHB = sleep apnea specific hypoxic burden.

## 4.6. Safety Analyses

Unless specified otherwise, safety assessments will be guided by the safety estimand. Thus, unless specified otherwise, safety analyses will be conducted in the safety analysis set (Table GPIF.3.1); all events that occur between the first dose date of study drug and the end date of study participation will be included, regardless of the adherence to study drug.

The statistical assessment of homogeneity of the distribution of categorical safety responses between tirzepatide MTD and placebo will be conducted using Fisher’s exact test, unless specified otherwise.

The mean change from baseline differences among treatments at all scheduled visits will be assessed via an MMRM using REML. The model will include treatment group, stratification factors, visit and treatment-by-visit interaction as fixed effects, and baseline value of the safety parameter as a covariate. To model the covariance structure within participants, the unstructured covariance matrix will be used. If this model fails to converge, the covariance structures specified in Section 4.3.1 will be tested in order until convergence is met. If the data does not warrant the MMRM model, then an ANCOVA model will be used.

For selected safety parameters, time-to-first-event analysis via the Cox-proportional hazards model may be conducted. Participants without the event will be censored at the end of study participation. For participants experiencing the event, the “time-to-first-event” will be the time (in days) from first dose to first occurrence of the event.

### 4.6.1. Analysis of Adverse Events

#### 4.6.1.1. Treatment-Emergent Adverse Events

A TEAE is defined as an event that first occurred or worsened in severity after the first dose of study treatment. The MedDRA LLT will be used in the treatment-emergent derivation. The maximum severity for each LLT during the baseline period including ongoing medical history will be used as baseline severity. For events with a missing severity during the baseline period, it will be treated as “mild” in severity for determining treatment-emergence. Events with a missing severity during the postbaseline period will be treated as “severe” and treatment-emergence will be determined by comparing to baseline severity.

For events occurring on the day of taking study medication for the first time, the CRF-collected information (for example, if the event starts or worsens after the first dose) will be used to determine whether the event was pre- versus posttreatment if available. If the relevant information is not available, then the events will be counted as posttreatment.

Unless otherwise specified, the counts and percentages of participants with TEAEs will be summarized by treatment using MedDRA PT nested within SOC. Statistical comparisons will be applied at both the SOC and PT levels. Events will be ordered by decreasing frequency in the tirzepatide arm within the SOC. The SOC will be in alphabetical order.

An overview of the number and percentage of participants who experienced a TEAE, SAE, death, discontinued from study treatment or study due to an AE, or with a TEAE related to study treatment will be summarized by treatment.

The counts and percentages of participants with TEAEs by maximum severity will be summarized by treatment using the MedDRA PT within the SOC. For each participant and TEAE, the maximum severity for the MedDRA PT is the maximum postbaseline severity observed from all associated LLTs mapping to the MedDRA PT. The maximum severity will be determined based on the nonmissing severities. If all severities are missing for the defined postbaseline period of interest, it will show as missing in the table.

For events that are gender specific, the denominator and computation of the percentage will only include participants of the given gender.

#### **4.6.1.2. Common Adverse Events**

The counts and percentages of participants with TEAEs, overall and common (common TEAEs occurred in  $\geq 5\%$  of participants before rounding), will be summarized by treatment using MedDRA PT. Events will be ordered by decreasing frequency in the tirzepatide arm.

#### **4.6.1.3. Deaths**

A listing of all deaths during the study will be provided. The listing will include participant identification including:

- treatment
- site number
- date of death
- age at the time of enrollment
- sex
- associated AE group identification
- time from last dose of study drug to death (if participant had discontinued study drug), and
- primary cause of death.

#### **4.6.1.4. Other Serious Adverse Events**

The counts and percentages of participants who experienced an SAE (including deaths and SAEs temporally associated or preceding deaths) during the postbaseline period will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency in the tirzepatide arm within the SOC. The SOC will be in alphabetical order.

A listing of all SAEs will be provided. The listing will include:

- treatment
- participant identification including the site number
- date of event
- age at the time of enrollment
- sex
- AE group identification
- MedDRA SOC and PT
- severity
- outcome
- relationship to study drug
- time from first dose of study drug to the event, and
- time from most recent dose to event (if participant discontinued study drug prior to the event).

#### **4.6.1.5. Discontinuation Due to Adverse Events**

The counts and percentages of participants who discontinued from study treatment or study due to an AE during the postbaseline period may be summarized by treatment group using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency in the tirzepatide arm within the SOC.

#### **4.6.2. Patient Narratives**

Patient narratives will be provided for all participants who experience any of the following “notable” events:

- death
- SAE
- pregnancy, or
- permanent discontinuation of study treatment due to AEs.

Patient narratives (patient level data and summary paragraph) will be provided for participants in the randomized population with at least 1 notable event.

#### **4.6.3. Special Safety Topics**

For AESI or special safety topics, the counts and percentages of participants will be summarized by treatment and PT with decreasing frequency in the tirzepatide arm if the overall count is 10 or more. Individual participant level data may be presented. Displays with individual participant level data may be created using various formats, such as a customized listing and/or a



customized graphical participant profile. AESI are defined in each section of special safety topics, where applicable.

#### **4.6.3.1. Exocrine Pancreas Safety**

##### **4.6.3.1.1. Pancreatic Enzyme**

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

The counts and percentages of participants with maximum postbaseline pancreatic enzyme value exceeding the following thresholds will be provided by maximum baseline pancreatic enzyme value ( $\leq$ ULN,  $>$ ULN), and postbaseline:

- $\leq 1 \times$  ULN
- $(>1$  to  $\leq 3) \times$  ULN
- $(>3$  to  $\leq 5) \times$  ULN
- $(>5$  to  $\leq 10) \times$  ULN, and
- $>10 \times$  ULN.

An MMRM analysis will be used to analyze each pancreatic enzyme with a log-transformed (postbaseline measure/baseline measure) response variable and treatment, nominal visit, treatment-by-nominal visit interaction as fixed effects.

##### **4.6.3.1.2. Pancreatitis Events**

Summaries of adjudicated and investigator-reported pancreatic events will be provided by treatment. Detailed searching criteria can be found in Appendix 2 (Section 6.2).

Treatment-emergent adjudication-confirmed pancreatitis will be considered as an AESI. Listing of participants with adjudicated pancreatitis may be provided if deemed necessary.

#### **4.6.3.2. Gastrointestinal Adverse Events**

##### **4.6.3.2.1. Nausea, Vomiting, and Diarrhea**

Summaries and analyses for incidence and severity of nausea, vomiting (including “vomiting” and “vomiting projectile”), diarrhea (including “diarrhea” and “diarrhoea”), and 3 events combined, will be provided by each treatment group.

Summary of the prevalence over time for nausea, vomiting, and diarrhea will also be presented. Time to the onset of nausea, vomiting, and diarrhea will be plotted.

##### **4.6.3.2.2. Severe Gastrointestinal Events**

The PTs under the *Gastrointestinal disorders* SOC in MedDRA will be used to identify GI AEs, and only the PTs with serious/severe treatment-emergent cases will be considered as AESIs.

The counts and percentages of participants with severe/serious treatment-emergent GI events may be summarized by treatment, or a listing may be provided.

**4.6.3.3. Hepatobiliary Disorders**

**4.6.3.3.1. Hepatobiliary Events**

The counts and percentages of participants with treatment-emergent hepatic events may be summarized by treatment using the MedDRA PTs. The detailed search criteria can be found in Appendix 2 (Section 6.2).

Events related to acute gallbladder disease may also be summarized or a listing may be provided. The search criteria can be found in Appendix 2 (Section 6.2).

Severe/serious treatment-emergent hepatic events and acute gallbladder disease will be considered as AESIs.

**4.6.3.3.2. Liver Enzymes**

Common analyses for laboratory analyte measurements described in Section 4.6.5 are applicable for the liver enzyme related measurements. This section describes additional analyses for liver enzymes.

For the postbaseline maximum value, all planned and unplanned measurements will be included. When or if multiple values are available (that is, unplanned measurement) prior to randomization, the maximum value will be used as baseline. Table GPIF.4.6 describes the planned analyses related to hepatic safety.

**Table GPIF.4.6. Summary Tables and Figures Related to Hepatic Safety**

Analysis	Population or Analysis Set
<p>Abnormal Postbaseline Categories – Hepatic Safety Parameters</p> <ul style="list-style-type: none"> <li>• ALT: The number and percentage of participants with a measurement greater than or equal to 1 time (1×), 3 times (3×), 5 times (5×), 10 times (10×), and 20 times (20×) the performing laboratory ULN during the treatment period for all participants with a postbaseline value and for subsets based on the following levels of baseline value:                             <ul style="list-style-type: none"> <li>○ participants whose nonmissing maximum baseline value is <math>\leq 1 \times \text{ULN}</math>,</li> <li>○ participants whose maximum baseline is <math>&gt; 1 \times \text{ULN}</math>,</li> <li>○ participants whose baseline values are missing.</li> </ul> </li> <li>• AST: The number and percentage of participants with a measurement greater than or equal to 1 time (1×), 3 times (3×), 5 times (5×), 10 times (10×), and 20 times (20×) the performing laboratory ULN during the treatment period for all participants with a postbaseline value and for subsets based on various baseline levels, as described above for ALT.</li> <li>• ALP: The number and percentage of participants with a measurement greater than or equal to 2 times (2×), and 3 times (3×) the performing laboratory ULN during the treatment period will be summarized for all participants with a postbaseline and for the following subsets based on the baseline values:                             <ul style="list-style-type: none"> <li>○ participants whose nonmissing maximum baseline value is <math>\leq 1 \times \text{ULN}</math>,</li> <li>○ participants whose maximum baseline is <math>&gt; 1 \times \text{ULN}</math>, but <math>&lt; 2 \times \text{ULN}</math>,</li> <li>○ participants whose maximum baseline value is <math>\geq 2 \times \text{ULN}</math>, and</li> </ul> </li> </ul>	<p>Safety Participants</p>

Analysis	Population or Analysis Set
<ul style="list-style-type: none"> <li>○ participants whose baseline values are missing.</li> <li>● TBL: The number and percentage of participants with a measurement greater than or equal to 2 times (2×), 5 times (5×), and 8 times (8×) the performing laboratory ULN during the treatment period will be summarized for all participants with a postbaseline value and the same subsets as described for ALP.</li> <li>● DBL: The number and percentage of participants with a measurement greater than or equal to 2 times (2×) and 5 times (5×) the performing laboratory ULN during the treatment period will be summarized for all participants with a postbaseline value and the same subsets as described for ALP.</li> <li>● GGT: The number and percentage of participants with a measurement greater than or equal to 2 times (2×) the performing laboratory ULN during the treatment period will be summarized for all participants with a postbaseline value.</li> </ul>	
Hepatocellular Drug-Induced Liver Injury Screening Plot (TBL vs. ALT or AST).	Safety Participants
Hepatocellular Drug-Induced Liver Injury Screening Table.	Safety Participants
Cholestatic Drug-Induced Liver Injury Screening Plot (TBL vs. ALP).	Safety Participants
Cholestatic Drug-Induced Liver Injury Screening Table.	Safety Participants
<p>Participant profiles will be created for participants meeting criteria for a comprehensive hepatic evaluation (as defined in the protocol).</p> <p>Participant profiles will include demographics, disposition, information collected on the hepatic safety CRFs (where applicable) and a display of study drug exposure, adverse events, medications, blood pressure, heart rate, and the liver -related measurements over time.</p>	Safety Participants

Abbreviations: ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CRF = case report form; DBL = direct bilirubin; GGT = gamma-glutamyl transferase; TBL = total bilirubin; ULN = upper limit of normal.

**4.6.3.4. Hypoglycemia**

The following categories in accordance with the 2020 American Diabetes Association position statement on glycemic targets (ADA 2020) will be defined in the database.

**Level 1 hypoglycemia**

**Glucose <70 mg/dL (3.9 mmol/L) and ≥54 mg/dL (3.0 mmol/L):** Level 1 hypoglycemia can alert a person to take action such as treatment with fast-acting carbohydrates. Providers should continue to counsel participants to treat hypoglycemia at this glucose alert value.

**Level 2 hypoglycemia**

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

### Level 3 hypoglycemia

**Severe hypoglycemia (in adults):** A severe event characterized by altered mental and/or physical status requiring assistance for treatment of hypoglycemia. For example, participants had altered mental status, and could not assist in their own care, or were semiconscious or unconscious, or experienced coma with or without seizures, and the assistance of another person was needed to actively administer carbohydrate, glucagon, or other resuscitative actions. Glucose measurements may not be available during such an event, but neurological recovery attributable to the restoration of glucose concentration to normal is considered sufficient evidence that the event was induced by a low glucose concentration.

- The determination of a hypoglycemic event as an episode of severe hypoglycemia, as defined above, is made by the investigator based on the medical need of the participant to have required assistance and is not predicated on the report of a participant simply having received assistance.
- If a hypoglycemic event meets the criteria of severe hypoglycemia, the investigator must record the event as serious on the AE CRF and report it to Lilly as an SAE.

### Nocturnal hypoglycemia

Nocturnal hypoglycemia is a hypoglycemia event (including severe hypoglycemia) that **occurs at night** and presumably during sleep.

To avoid duplicate reporting, all consecutive BG values <70 mg/dL (<3.9 mmol/L) occurring within a 1-hour period may be considered to be a single hypoglycemic event (Weinberg et al. 2010; Danne et al. 2013).

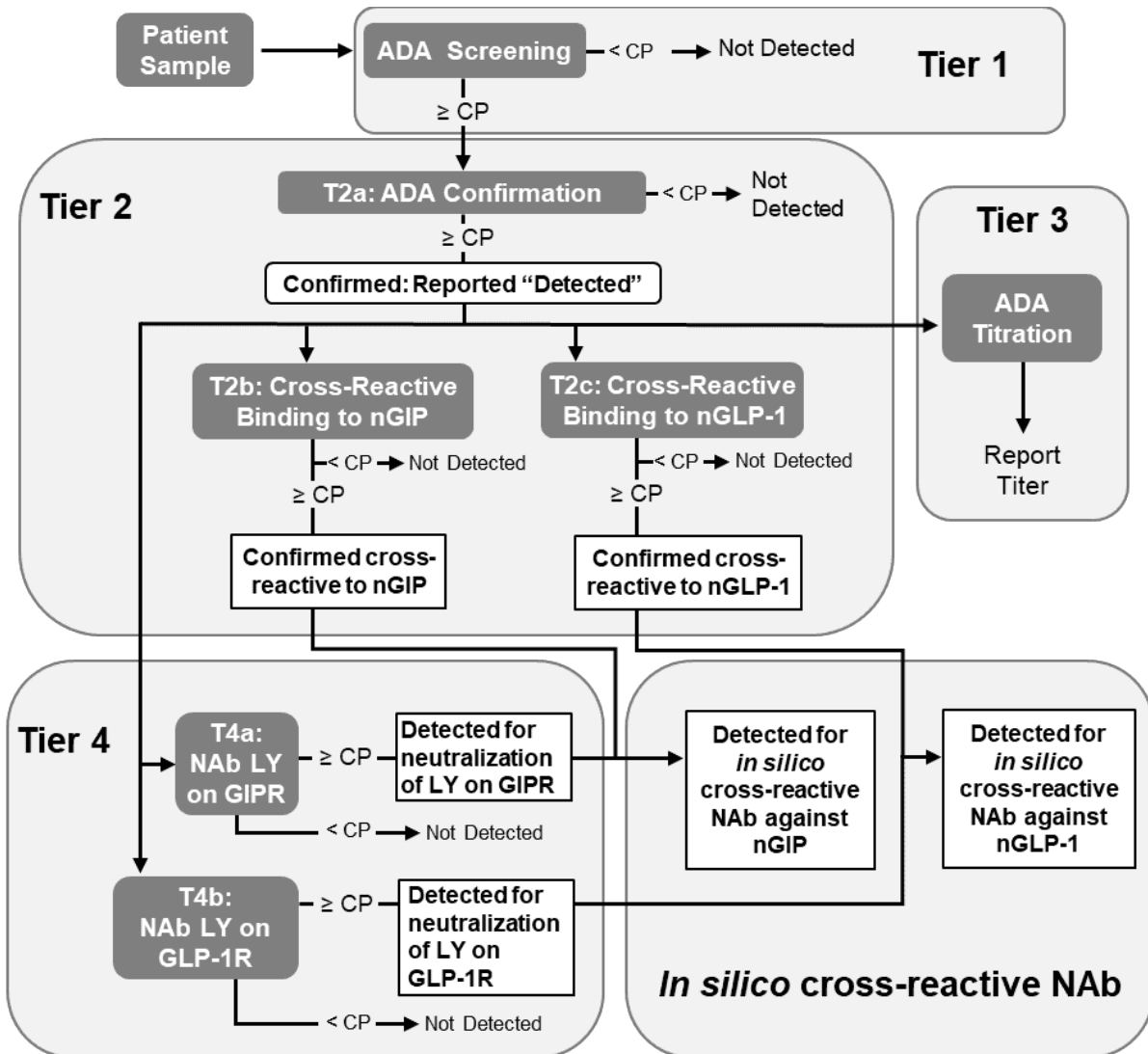
Summary and analyses of Level 2 or Level 3 hypoglycemic events will be performed.

#### 4.6.3.5. Immunogenicity

##### 4.6.3.5.1. *Definitions of Sample ADA Status*

At a high level, an individual sample is potentially examined multiple times, in a hierarchical procedure, to produce a sample ADA assay result and potentially multiple cross-reactive antibodies assay results and multiple Nab assay results.

The cut points used, the drug tolerance of each assay, and the possible values of titers are operating characteristics of the assay. [Figure GPIF.4.1](#) details a flow chart that reflects the multitiered testing approach.



Abbreviations: ADA = anti-drug antibody; CP = cut point; GIP = glucose-dependent insulinotropic polypeptide; GIPR = glucose-dependent insulinotropic polypeptide receptor; GLP-1 = glucagon-like peptide-1; GLP-1R = glucagon-like peptide-1 receptor; nGIP = native GIP; nGLP-1 = native GLP-1; LY = LY3298176; Nab = neutralizing antibodies.

**Figure GPIF.4.1. Flowchart of immunogenicity multitiered testing approach.**

Table GPIF.4.7 outlines results as reported from Tier 2a of the multitiered testing approach. Tier 4 results are reported similarly.

**Table GPIF.4.7. Sample ADA Assay Results**

Sample Laboratory Result	Explanation
Detected	ADA are detected and confirmed.
Not Detected	The raw result as reported from the laboratory indicates not detected. The clinical interpretation of such results depends on other factors (see Table GPIF.4.8).
NO TEST, QNS, and so on.	Sample exists but was unevaluable by the assay.

Abbreviations: ADA = anti-drug antibody; QNS = quantity not sufficient.

It can be the case that the presence of high concentrations of tirzepatide will affect ADA immunoassays, and conversely high levels of ADA may affect the measurement of tirzepatide concentration. Thus, a tirzepatide drug concentration, assessed from a sample drawn at the same time as the ADA sample, plays a key role in clinical interpretation of a sample when the laboratory result is Not Detected (see [Table GPIF.4.8](#)).

**Table GPIF.4.8. Sample Clinical ADA Interpretation Results**

Sample Clinical Interpretation	Explanation
ADA Present	ADA assay result is Detected
ADA Not Present	ADA assay result is Not Detected and simultaneous drug concentration is at a level that has been demonstrated to not interfere in the ADA detection method (i.e., drug concentration is below the assay’s drug tolerance level). For participants receiving placebo, drug concentration is not assessed and is assumed to be below the assay’s drug tolerance level. If drug concentration was planned but is not available for a treatment period sample, a Not Detected sample will be declared ADA Not Present.
ADA Inconclusive	ADA assay result is Not Detected but drug concentration in the sample is at a level that can cause interference in the ADA detection method.
ADA Missing	ADA sample not drawn, QNS, not tested, and so on, causing there to be no laboratory result reported or the result is reported as “no test.”

Abbreviations: ADA = antidrug antibody; QNS = quantity not sufficient.

All ADA Present samples will be evaluated for cross-reactive GIP (Tier 2b), cross-reactive GLP-1 (Tier 2c), Nab LY (tirzepatide) on GIPR (Tier 4a), and Nab LY (tirzepatide) on GLP-1R (Tier 4b).

Similar terminology to [Table GPIF.4.8](#) applies for each type of cross-reactive and Nab assay. Importantly, each of these are distinct assays and, in general, have different assay operating characteristics. The following are considered inconclusive for the Nab result:

- Nab LY on GIPR: if Nab result is not detected, and PK concentration is greater than or equal to drug tolerance limit of the Nab LY on GIPR assay
- Nab LY on GLP-1R: if Nab result is not detected, and PK concentration is greater than or equal to drug tolerance limit of the Nab LY on GLP-1R assay

For cross-reactive Nab interpretations against native GIP and GLP-1, an *in silico* method utilizing results from Tiers 2b and 2c, Tiers 4a and 4b, and tirzepatide concentrations is introduced. The *in silico* method is outlined in [Table GPIF.4.9](#).

**Table GPIF.4.9. In Silico Classification for Cross-Reactive Nab**

<i>In Silico</i> Classification	Cross-Reactive ADA Result	Nab Result	Circulating Tirzepatide Level (ng/mL)	<i>In Silico</i> Cross- Reactive Nab Interpretation
Cross-reactive Nab to nGIP	Tier 2b: “Not Detected”	Tier 4a: “Not Detected” Or Tier 4a: “Detected” or N/A or Missing	Any Value or Missing	Not Present
	Tier 2b: “Detected”	Tier 4a: “Not Detected”	< drug tolerance limit of Tier 4a assay	Not Present
	Tier 2b: “Detected”	Tier 4a: “Not Detected”	≥ drug tolerance limit of Tier 4a assay	Inconclusive
	Tier 2b: “Detected”	Tier 4a: “Detected”	< drug tolerance limit of Tier 4a assay	Present
	Tier 2b: “Detected”	Tier 4a: “Detected”	≥ drug tolerance limit of Tier 4a assay	Present
Cross-reactive Nab to nGLP-1	Tier 2c: “Not Detected”	Tier 4b: “Not Detected” Or Tier 4b: “Detected” or N/A or Missing	Any Value or Missing	Not Present
	Tier 2c: “Detected”	Tier 4b: “Not Detected”	< drug tolerance limit of Tier 4b assay	Not Present
	Tier 2c: “Detected”	Tier 4b: “Not Detected”	≥ drug tolerance limit of Tier 4b assay	Inconclusive
	Tier 2c: “Detected”	Tier 4b: “Detected”	< drug tolerance limit of Tier 4b assay	Present
	Tier 2c: “Detected”	Tier 4b: “Detected”	≥ drug tolerance limit of Tier 4b assay	Present

Abbreviations: ADA = antidrug antibody; Nab = neutralizing antibody; nGIP = native glucose-dependent insulinotropic polypeptide; nGLP-1 = native glucagon-like peptide-1; Tier 2b = cross-reactive ADA to nGIP; Tier 2c = cross-reactive ADA to nGLP-1; Tier 4a = Nab LY (tirzepatide) on GIPR; Tier 4b = Nab LY (tirzepatide) on GLP-1R.

Note: Only the drug tolerance limits of the Tier 4a and 4b assays are used for in silico classifications as they are lower than the drug tolerance limits of the Tier 2b and 2c assays, respectively.

**4.6.3.5.2. Definitions of Immunogenicity Assessment Periods**

Immunogenicity baseline observations: Baseline period for immunogenicity assessment for each participant includes all observations prior to first dose of study treatment. In instances where multiple baseline observations are collected, to determine participant ADA status the last nonmissing immunogenicity assessment prior to first administration of study drug is used to determine treatment-emergent status (see below).

Immunogenicity postbaseline period observations: Postbaseline period observations for each participant includes all observations after the first administration of study drug.

**4.6.3.5.3. Definitions of Participant ADA Status**

TE ADA-evaluable participants: A participant with a nonmissing baseline ADA result and at least 1 nonmissing postbaseline ADA result.

TE ADA-unevaluable participant: any participant who does not meet the evaluable criteria.

Baseline ADA Present (preexisting antibody): ADA detected in a sample collected up to the first dose date and time.

Baseline ADA Not Present: ADA is not detected, and the corresponding PK concentration is missing or below the drug tolerance limit in a sample collected up to the first dose date and time.

TE ADA+ participant: An evaluable participant who had a:

- baseline status of ADA Not Present and at least 1 postbaseline status of ADA Present with titer  $\geq 2 \times$  MRD, where the MRD is the minimum required dilution of the ADA assay or
- baseline and postbaseline status of ADA Present, with the postbaseline titer being 2 dilutions (4-fold) greater than the baseline titer. That is, the participant has baseline (B) status of ADA Present, with titer 1:B, and at least 1 postbaseline (P) status of ADA Present, with titer 1:P, with  $P/B \geq 4$ .

As shown in [Figure GPIF.4.1](#), a titer is expected when ADA assay result is Detected. On occasion, the corresponding assay cannot be performed, in which case a titer value will be imputed for the purpose of TE ADA determination. A baseline sample with detected ADA and no titer is imputed to be the MRD (1:10), and a postbaseline sample with ADA detected and no titer is imputed to be 1 dilution above the MRD (1:20).

TE ADA- Inconclusive participant: A TE ADA-evaluable participant is TE ADA Inconclusive if  $\geq 20\%$  of the participant's postbaseline samples, drawn predose, are ADA Inconclusive and all remaining postbaseline samples are ADA Not Present.

TE ADA- participant: A TE ADA-evaluable participant is TE ADA- when the participant is not TE ADA+ and not TE ADA Inconclusive.

For each Nab assay, the following are defined:

Nab+ participant: A participant who is TE ADA+ and has a Nab+ sample in the postbaseline period.

Nab Inconclusive participant: A participant who is TE ADA+, is not Nab+, and all samples that have TE ADA+ titer have a Nab Inconclusive sample result.

Nab- participant: A participant is neither Nab+ nor Nab Inconclusive.

Unless specified otherwise, the above-mentioned definitions of Nab are applicable to all Nab analyses, including cross-reactive Nab analyses, and cross-reactive antibodies.

#### **4.6.3.5.4. Analyses to be Performed**

The count and proportion of participants who are TE ADA+ will be tabulated by treatment group, where the proportions are relative to the number of TE ADA-evaluable participants, as defined above. The tabulation will include the count and proportion of participants with ADA Present at baseline, and the count and proportion of TE ADA+ participants exhibiting each type of cross-reactive antibodies and Nab. This analysis will be performed for the planned treatment



period. The cross-reactive Nab will include the *in silico* classification as cross-reactive Nab for summary.

Additional immunogenicity analyses as determined later may be presented. The relationship between the presence of antibodies and tirzepatide PK and pharmacodynamic response including safety and efficacy to tirzepatide may be assessed.

#### 4.6.3.6. Hypersensitivity Reactions

Two main analyses are performed in support of assessment of potential immediate hypersensitivity, including anaphylaxis as well as potential nonimmediate hypersensitivity.

**Time Period A**, of potential immediate hypersensitivity includes all TEAEs occurring from start of study drug administration up to 24 hours after end of study drug administration. For events without the hypersensitivity eCRF, only date (no time) information is collected. Among these events without time information, the event occurred on the same date as the study drug injection date will be included in Time Period A.

**Time Period B**, of potential non-immediate hypersensitivity, includes all TEAEs occurring more than 24 hours after the end of study drug administration, but prior to subsequent drug administration.

Analyses for both time periods are based on the following:

- narrow and algorithm terms in *Anaphylactic reaction* SMQ (20000021) (analysis for algorithm term only applicable for Time Period A)
- narrow terms in *Angioedema* SMQ (20000024)
- narrow terms in *Severe cutaneous adverse reactions* SMQ (20000020), and
- narrow terms in *Hypersensitivity* SMQ (20000214)

For the *Anaphylactic reaction* SMQ, each term is classified by scope (narrow, broad) and by category (A, B, C, and D). All narrow terms are category A, and all broad terms are category B, C, or D. In addition to the usual narrow and broad searches, the SMQ defines an algorithm to further refine the cases of interest. For Time Period A analysis, the *Anaphylactic reaction* SMQ algorithm will be included. The algorithm is based upon events that occur within Time Period A. The counts and percentages of participants who experienced a TEAE for the following will be analyzed for each of the 2 time periods:

- any narrow term from any 1 of the 4 SMQs indicated above (that is, combined search across narrow of all 4 SMQs), and
- any narrow scope term within each SMQ, separately (that is, narrow SMQ search). For Time Period A analysis, any term from Anaphylactic reaction SMQ algorithm.

Within each query, individual PTs that satisfied the queries will be summarized. For Time Period A analysis, the Anaphylactic reaction SMQ algorithm will be summarized. Also, a single event may satisfy multiple SMQs, in which case the event contributes to every applicable SMQ.

**4.6.3.6.1. Severe/Serious Hypersensitivity Reactions**

The severe/serious cases of hypersensitivity will be considered as AESIs. Summary of severe/serious hypersensitivity reactions or listing may be provided.

**4.6.3.7. Injection Site Reaction**

Injection site reaction, incidence and rates, and related information reported via “Injection Site Reactions” eCRF will be summarized by treatment. Information to be summarized include location of the reaction, timing of reaction relative to study drug administration, and characteristics of the injection site reaction: erythema, induration, pain, pruritus, and edema.

Patient-based analysis and event-based analysis may be provided if necessary. The patient-based analysis summarizes all ISR questionnaire forms for an individual participant with a single statistic, typically an extreme value. This analysis allows each participant to contribute only once for each parameter, at the expense of a focus on the most extreme events. By contrast, the event-based analysis summarizes all ISR questionnaire forms received, without regard to individual participants. This provides characteristics of ISR events as a proportion of all events for which questionnaire responses were provided, at the expense of some potential bias due to differential contribution of individual participants to the analysis.

The counts and percentages of participants with treatment-emergent injection site reaction will be summarized by treatment using the MedDRA PTs. Detailed searching criteria can be found in Appendix 2 (Section 6.2).

The PTs will be summarized in decreasing order of incidence for tirzepatide-treated participants.

**4.6.3.7.1. Severe/Serious Injection Site Reactions**

Severe/serious injection site reactions (for example, abscess, cellulitis, erythema, hematomas/hemorrhage, exfoliation/necrosis, pain, subcutaneous nodules, swelling, induration, inflammation) will be considered as AESI.

The counts and percentage of participants with severe/serious ISRs may be summarized by treatment, or a listing of participants with treatment-emergent severe/serious ISRs may be provided.

**4.6.3.8. Major Adverse Cardiovascular Events**

MACE reported by investigators are adjudicated by an independent clinical endpoint committee in a blinded fashion. Unreported events may also be independently identified by the clinical endpoint committee.

The following positively adjudicated MACE will be considered as AESIs:

- death due to cardiovascular AEs
- myocardial infarction
- hospitalization for unstable angina
- hospitalization for heart failure

- coronary interventions (such as coronary artery bypass graft or percutaneous coronary intervention), and
- cerebrovascular events, including cerebrovascular accident (stroke) and transient ischemic attack.

The counts and percentages of participants with adjudicated MACE may be summarized by treatment. In addition, MACE reported by investigator may also be summarized although a MACE reported by investigator that is not positively adjudicated is not considered an AESI.

A listing of participants reporting MACE events, either reported by investigator or identified by the clinical endpoint committee, may be provided.

#### **4.6.3.9. Major Depressive Disorder/Suicidal Ideation or Behavior**

The severe/serious treatment-emergent major depressive disorder/suicidal ideation or behavior will be captured as AESI. AEs will be searched using MedDRA PT terms. Detailed searching criteria can be found in Appendix 2 (Section 6.2).

The counts and percentages of participants with TEAEs will be summarized by treatment group using MedDRA PT nested within SMQ. Events will be ordered by decreasing frequency in the total tirzepatide group nested within SMQ. A listing of participants with major depressive disorder/suicidal ideation or behavior may be provided if deemed necessary.

Additionally, suicidal ideation and behavior, and depression will be assessed by the investigator via spontaneously reported AEs and through the use of the C-SSRS and the PHQ-9.

##### **4.6.3.9.1. Patient Health Questionnaire**

Total scores for the PHQ-9 range from 0 to 27 with total scores categorized as

- none (not depressed): 0 through 4
- mild: 5 through 9
- moderate: 10 through 14
- moderately severe: 15 through 19, and
- severe: 20 through 27.

Shift tables will be provided showing the counts and percentages of participants within each baseline category (maximum value) versus each postbaseline category (maximum value) by treatment.

Additionally, the following 3 outcomes of interest will be compared between treatments (based on the maximum value during baseline and postbaseline):

- any increase in depression category (that is, worsening of depression): includes participants in the none, mild, moderate, or moderately severe category during baseline and with at least 1 postbaseline measurement
- increase from No or Mild Depression to Moderate, Moderately Severe, or Severe Depression: includes participants in the none or mild depression category during baseline and with at least 1 postbaseline measurement, and

- increase from Mild or Moderate Depression to Moderately Severe or Severe Depression: includes participants in the mild or moderate depression category during baseline and with at least 1 postbaseline measurement.

#### **4.6.3.9.2. Suicidal Ideation and Behavior Solicited Through C-SSRS**

Suicide-related thoughts and behaviors occurring during the entire study period, based on the C-SSRS, will be summarized by treatment group. In particular, for each of the following suicide-related events, the counts and percentages of participants with the event will be summarized by treatment group:

- died by suicide
- nonfatal suicide attempt
- interrupted attempt
- aborted attempt
- preparatory acts or behavior
- active suicidal ideation with specific plan and intent
- active suicidal ideation with some intent to act without specific plan
- active suicidal ideation with any methods (no plan) without intent to act
- nonspecific active suicidal thoughts
- wish to be dead, and
- nonsuicidal, self-injurious behavior.

In addition, the counts and percentages of participants who experienced at least 1 of the composite measures will be presented. The participants with at least 1 postbaseline C-SSRS assessment are included. The composite measure is determined at each assessment by the “yes” or “no” responses in C-SSRS categories by the study participant:

- Category 1 – Wish to be Dead
- Category 2 – Non-specific Active Suicidal Thoughts
- Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
- Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
- Category 5 – Active Suicidal Ideation with Specific Plan and Intent
- Category 6 – Preparatory Acts or Behavior
- Category 7 – Aborted Attempt
- Category 8 – Interrupted Attempt
- Category 9 – Actual Attempt (non-fatal), and
- Category 10 – Completed Suicide.

Composite endpoints of suicidal ideation and suicidal behavior based on the above categories are defined below:

- **Suicidal ideation:** A “yes” answer at any time during study to any 1 of the 5 suicidal ideation questions (Categories 1 through 5) on the C-SSRS.
- **Suicidal behavior:** A “yes” answer at any time during study to any 1 of the 5 suicidal behavior questions (Categories 6 through 10) on the C-SSRS.

- **Suicidal ideation or behavior:** A “yes” answer at any time during study to any 1 of the 10 suicidal ideation and behavior questions (Categories 1 through 10) on the C-SSRS.

A listing contains data for each participant with suicidal ideation, suicidal behavior, or nonsuicidal self-injurious behavior during the study by treatment and visit. Data from all visits are displayed, regardless of a “yes” or “no” answer, for participants with any “yes” answer for suicidal ideation, suicidal behavior, or self-injurious behavior without suicidal intent.

#### **4.6.3.10. Renal Safety**

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 4.6.5.

In addition, 2 shift tables examining renal function will be created. A minimum-to-minimum shift table of estimated glomerular filtration rate estimated by the Chronic Kidney Disease Epidemiology Collaboration equation with units mL/min/1.73 m<sup>2</sup>, using categories (<30, ≥30 to <45, ≥45 to <60, ≥60 to <90, and ≥90 mL/min/1.73 m<sup>2</sup>). A maximum-to-maximum shift table of UACR, using the categories UACR <30 mg/g, ≥30 mg/g UACR to ≤300 mg/g, UACR >300 mg/g (respectively, these represent normal, microalbuminuria, and macroalbuminuria).

MMRM analyses as described in Section 4.6 for estimated glomerular filtration rate and log-transformed UACR will be provided. Log transformation will be performed for UACR.

##### **4.6.3.10.1. Acute Renal Events**

Because severe GI events may lead to dehydration, which could cause a deterioration in renal function including acute renal failure, dehydration events will be analyzed as described in the next section. Acute renal events associated with chronic renal failure exacerbation will also be captured.

Severe/serious renal events from the following SMQ search will be considered as AESI.

The counts and percentages of participants with acute renal events may be summarized by treatment if overall count >10 by using the MedDRA PTs contained in any of the following SMQs:

- Acute renal failure: narrow terms in *Acute renal failure* SMQ (20000003), and
- Chronic kidney disease: narrow terms in *Chronic kidney disease* SMQ (20000213).

In addition, a listing of participants with treatment-emergent acute renal events may be provided, if deemed necessary.

##### **4.6.3.10.2. Dehydration**

Dehydration events will be captured in the narrow terms in *Dehydration* SMQ (20000232). Severe/serious dehydration events will be considered as AESI. A listing of participants with treatment-emergent dehydration events may be provided.

#### 4.6.3.11. Thyroid Safety Monitoring

##### 4.6.3.11.1. Calcitonin

The purpose of calcitonin measurements is to assess the potential of tirzepatide to affect thyroid C-cell function, which may indicate development of C-cell hyperplasia and neoplasms.

Observed calcitonin data (a thyroid-specific laboratory assessment) will be summarized by treatment and nominal visit.

The counts and percentages of participants with a maximum postbaseline calcitonin value in the following thresholds will be provided by treatment and maximum baseline calcitonin value ( $\leq 20$  ng/L,  $>20$  ng/L to  $\leq 35$  ng/L,  $>35$  ng/L). Postbaseline categories are:

- $\leq 20$  ng/L
- $>20$  ng/L to  $\leq 35$  ng/L
- $>35$  ng/L to  $\leq 50$  ng/L
- $>50$  ng/L to  $\leq 100$  ng/L, and
- $>100$  ng/L.

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

Treatment-emergent thyroid malignancies and C-cell hyperplasia will be considered as AESI. Thyroid malignancies and C-cell hyperplasia will be identified using MedDRA HLT for *Thyroid neoplasms* and PT for *Thyroid C-cell hyperplasia*.

The counts and percentages of participants with treatment-emergent thyroid C-cell hyperplasia and malignancies may be summarized or a listing of participants with treatment-emergent thyroid C-cell hyperplasia and neoplasms may be provided if deemed necessary.

#### 4.6.3.12. Treatment-Emergent Arrhythmias and Cardiac Conduction Disorders

Severe/serious treatment-emergent arrhythmias and cardiac conduction disorders will be considered as AESI. The cardiovascular events will include clinically relevant rhythm and conduction disorders.

The treatment-emergent arrhythmias and cardiac conduction disorder events will be identified using the MedDRA PTs. Detailed searching criteria can be found in Appendix 2 (Section 6.2).

The counts and percentages of participants with treatment-emergent arrhythmias and cardiac conduction disorders may be summarized by treatment and PT nested within SMQ. The PT will be ordered with decreasing frequency in tirzepatide arm within SMQ. A listing of participants with treatment-emergent arrhythmias and cardiac conduction disorders may be provided if deemed necessary.

#### 4.6.3.13. Treatment-Emergent Malignancy

The malignancy events will be included using the MedDRA PTs contained in the *Malignant tumours* SMQ (20000194) narrow scope or *Tumours of unspecified malignancy* SMQ (20000195) narrow scope. Malignancy will be considered as an AESI.

The counts and percentages of participants with treatment-emergent malignancy will be summarized by treatment.

#### 4.6.3.14. Abuse Liability

To identify AE terms suggestive of abuse liability potential, narrow terms from the *Drug abuse and dependence* SMQ (20000101) will be used. The counts and percentages of participants will be summarized by treatment group in order of decreasing frequency.

These analyses will be performed for individual CSRs and the summary of clinical safety.

#### 4.6.4. Vital Signs

In the case that multiple records of an individual vital sign are collected at the same visit, they will be averaged prior to being used for data summaries and analyses.

Descriptive summaries by treatment and by nominal visit will be provided for baseline and postbaseline values as well as change from baseline values.

An MMRM and/or an ANCOVA model, as discussed in Section 4.6, using data including from the safety follow-up period might be conducted if necessary.

Counts and percentages of participants with treatment-emergent abnormal sitting SBP, sitting diastolic blood pressure, and pulse at any time during the entire study (including the off-drug follow up time period) will be summarized by treatment for participants who have both baseline and at least 1 postbaseline result. A treatment-emergent high result is defined as a change from a value less than or equal to the high limit at baseline to a value greater than the high limit at any time that meets the specified change criteria during the postbaseline period. A treatment-emergent low result is defined as a change from a value greater than or equal to the low limit at baseline to a value less than the low limit at any time that meets the specified change criteria during the postbaseline period. To assess decreases, change from the minimum value during the baseline period to the minimum value during the postbaseline period will be used. To assess increases, changes from the maximum value during the baseline period to the maximum value during the postbaseline period will be used. Both planned and unplanned measurements will be included in the analysis. The criteria for identifying participants with treatment-emergent vital sign abnormalities are stated in [Table GPIF.4.10](#).

**Table GPIF.4.10. Categorical Criteria for Abnormal Treatment-Emergent Blood Pressure and Pulse Measurements**

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

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

#### **4.6.5. Clinical Laboratory Evaluation**

Limits from the performing laboratory will be used to define low (L) and high (H). Descriptive summaries by treatment and by nominal visit will be provided for the baseline and postbaseline values as well as the change from baseline values. The associated descriptive will be presented in International System of Units and in conventional units.

For selected laboratory analyte measurements collected quantitatively, observed, and change from baseline values for each visit may be displayed in plots for participants who have both a baseline and at least 1 postbaseline planned measurement. Unplanned postbaseline measurements will be excluded from plots.

A shift table will be provided including unplanned measurements. The shift table will include the number and percentage of participants within each baseline category (low, normal, high, or missing) versus each postbaseline category (low, normal, high, or missing) by treatment. The proportion of participants shifted will be compared between treatments using Fisher's exact test.

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

A listing of abnormal findings will be created for laboratory analyte measurements, including qualitative measures. The listing will include:

- participant identification
- treatment group
- laboratory collection date
- study day
- analyte name, and
- analyte finding.

The MMRM model or ANCOVA (if MMRM model is not applicable) as described in Section 4.6 will be used for the analysis during the treatment period (excluding the safety follow-up period) for the continuous measurements for selected laboratory tests with or without log-transformed (postbaseline measure/baseline measure) response variables. For measures analyzed using log-transformed values, the results will be presented with the scale back transforming to the original, related scale.

The summary of treatment-emergent abnormal, high, or low laboratory results any time will be provided.

#### **4.6.6. Product Complaints**

A listing of all product complaints by subcategory, inclusive of device product complaints that led to an AE or that could have led to an SAE had intervention not been taken, will be provided.



## 4.7. Other Analyses

### 4.7.1. PAP Adherence

For participants in ISA2 only, the adherence to use of the PAP machine over the course of the study will be summarized. Specifically, the summary statistics for PAP adherence at baseline and at each postbaseline week, stratified by treatment arm will be provided. Additionally, the categorical shift in PAP adherence between baseline and Week 52, stratified by treatment arm using a shift table for increased, decreased, or stable PAP use will be assessed. Finally, the number and percentage of participants in ISA2 who withdraw from regular PAP use will be summarized.

### 4.7.2. Health Outcomes

The PRO questionnaires will be analyzed using the mITT population on the EAS, unless specified otherwise.

Item-level missingness will be dealt with per the instrument developers' instruction.

Additional psychometric analyses will be performed by Value, Evidence, and Outcomes at Lilly and documented in a separate analysis plan.

Analyses of actual and change from baseline in PRO scores will be conducted using linear models with baseline PRO scores, treatment, stratification factors, and other factors that may be considered relevant.

If an administrative error occurs where more than 1 PRO questionnaire is completed within the same visit window by the same participant with different responses on at least 1 item, the questionnaire with the worst response will be used (for example, the questionnaire with the highest PHQ-9 score will be used). If more than 1 PRO questionnaire is completed within the same visit window with the same response to each item, the most recent response will be used. An administrative error occurs where more than 1 PRO questionnaire is completed within the same visit window by the same participant with different responses on at least 1 item, the questionnaire with the worst response will be used (for example, the questionnaire with the highest PHQ-9 score will be used). If more than 1 PRO questionnaire is completed within the same visit window with the same response to each item, the most recent response will be used.

#### 4.7.2.1. Functional Outcomes of Sleep Questionnaire

The FOSQ will be included to assess change in FOSQ domains and total score from baseline to Week 52. The FOSQ is a 30-item sleep-specific, participant-completed questionnaire used to assess the effect of disorders associated with excessive daytime sleepiness on daily functioning in adults. It assesses the following 5 domains of:

- General productivity (8 items)
- Activity level (9 items)
- Vigilance (7 items)
- Social outcomes (2 items), and
- Intimate and sexual relationships (4 items).

The FOSQ items assess participant's current status with each item rated on a scale of 1 (extreme difficulty) to 4 (no difficulty), with an additional not applicable (0 = "I don't do this activity for other reasons") also available. Individual domain scores are calculated by taking the mean of answered, non-zero items within each domain and ranges from 1 to 4 with higher scores indicating better outcomes. A total score can be calculated by first computing the mean score for each domain, then multiplying the mean of the domain scores by 5 (Weaver et al. 1997). The change from baseline in all 5 FOSQ domain scores will be assessed.

#### **4.7.2.2. Functional Outcomes of Sleep Questionnaire, 10 Items**

The FOSQ-10 will be included to assess change in FOSQ total score domains from baseline to Week 52. The 10-item sleep-specific, participant-completed questionnaire is a shortened version of the FOSQ with the same number of domains as the parent version. Of note, the FOSQ-10 has the same 5 domains as the FOSQ, but with fewer items per domain.

Calculation of the individual domain scores and the total score for the FOSQ-10 is carried out in a similar manner to FOSQ. The domain scores are first calculated by taking the mean of the answered, non-zero items within each domain. The total score is calculated by multiplying the mean of the domain scores by 5 (for each domain which has at least 1 response).

#### **4.7.2.3. Epworth Sleepiness Scale**

The ESS is an 8-item participant-completed measure that asks the participant to rate on a scale of 0 (would never doze) to 3 (high chance of dozing), their usual chances of dozing in 8 different daytime situations, with a recall period of "in recent times." The ESS total score is the sum of the 8-item scores and ranges from 0 to 24, with higher scores indicating greater daytime sleepiness (Johns 1991). Of note, if 1 or more item scores are missing, the ESS total score is not valid and will not be included in the analysis.

#### **4.7.2.4. PROMIS Short Form v1.0 Sleep-Related Impairment 8a**

The PROMIS Short Form v1.0 Sleep-Related Impairment 8a consists of 8 items each rated on a 5-point scale ranging from "not at all" to "very much." Items have a recall period of "in the past 7 days." Individual item scores will be combined and converted to a T-score using a response pattern scoring approach (Northwestern 2016a). The T-score standardizes the raw score to a distribution with a mean of 50 and standard deviation of 10.

#### **4.7.2.5. PROMIS Short Form v1.0 Sleep Disturbance 8b**

The PROMIS Short Form v1.0 Sleep Disturbance 8b consists of 8 items each rated on a 5-point scale ranging from "not at all" to "very much," "never" to "always," or "very poor" to "very good." Items have a recall period of "in the past 7 days." Individual item scores will be combined and converted to a T-score using a response pattern scoring approach (Northwestern 2016b). The T-score standardizes the raw score to a distribution with a mean of 50 and standard deviation of 10.

For item 8 of this scale (which is a measure of sleep quality), counts and percentages of participants at each time point will be summarized by nominal visit and by treatment. A shift table from baseline to postbaseline will be created at each postbaseline visit.

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

Per copyright owner, the QualityMetric Health Outcomes™ Scoring (PRO\_CoRe V2.0) Software will be used to derive the following domain and component scores:

- Mental Component Score (MCS)
- Physical Component Score (PCS)
- Physical Functioning domain (PF)
- Role-Physical domain (RP)
- Bodily Pain domain (BP)
- General Health domain (GH)
- Vitality domain (VT)
- Social Functioning domain (SF)
- Role-Emotional domain (RE), and
- Mental Health domain (MH).

The following analyses for the actual value and change from baseline value for each domain and component score will be conducted:

- descriptive summaries by treatment group, and
- analysis described in [Table GPIF.4.3](#).

#### 4.7.2.7. Patient Global Impression of Status/Change for OSA Outcomes

The counts and percentages of participants for PGIS for Physical Activity and PGIC response categories at each time point will be summarized by nominal visit and by treatment. A shift table from baseline to postbaseline of 3 PGIS response categories (OSA Sleepiness, Fatigue, and Snoring) and 4 PGIC response categories (OSA Sleepiness, Fatigue, Snoring, and Sleep Quality) will be created at each postbaseline visit.

#### 4.7.2.8. EQ-5D-5L

For the utility score and the Visual Analog Scale scores, the following analyses of the actual value and change from baseline value will be conducted:

- descriptive summaries by treatment group, and
- ANCOVA described in [Table GPIF.4.4](#).

### 4.7.3. Subgroup Analyses

The following subgroups will be analyzed using the efficacy estimand on change in AHI values from baseline to 52-week visit if there is an adequate number of participants in each treatment by subgroup (for example, 10%):

- age (<50 years, ≥50 years)
- baseline OSA severity (not severe, severe)
- race
- ethnicity
- region of enrollment (US, OUS)
- gender (male or female)

- baseline BMI ( $<35$ ,  $\geq 35$  and  $<40$ ,  $\geq 40$  kg/m<sup>2</sup>), and
- baseline ESS (ESS  $\leq 10$ , ESS  $> 10$ ).

Analyses for change from baseline in AHI will be performed using an MMRM model that includes the same fixed effects given for the primary analysis model plus factors of subgroup, 2-way interaction of subgroup and treatment, 2-way interaction of subgroup and visit, and 3-way interaction of treatment, visit, and subgroup. The possible interaction effect of subgroup and treatment at the primary endpoint (Week 52) will be evaluated. When analyzing OSA severity (not severe, severe) as a subgroup, the baseline AHI will not be included as a covariate to avoid confounding.

#### **4.8. Interim Analyses**

The details for the interim analyses and Data Monitoring Committee will be provided in the Data Monitoring Committee Charter.

##### **4.8.1. Unblinding Plan**

Details of the blinding and unblinding are provided in the Blinding and Unblinding Plan document for Master Protocol GPIF.

#### **4.9. Changes to Protocol-Planned Analyses**

Refer to [Table GPIF.1.1](#).

## 5. Sample Size Determination

Approximately 206 participants per ISA will be randomly assigned to either tirzepatide or placebo in a 1:1 ratio (approximately 103 participants per treatment arm), and the statistical power will be evaluated for the primary efficacy endpoint and key secondary combination PRO endpoint at a 2-sided significance level of 0.05. This sample size will provide the following:

- at least 90% power to demonstrate superiority of tirzepatide at the MTD (10 mg or 15 mg) to placebo for the mean percent change from baseline in AHI, assuming 50% improvement, with a common standard deviation of 50% and a dropout rate of 25%, and
- at least 90% power to demonstrate superiority of tirzepatide at the MTD (10 mg or 15 mg) to placebo for the hierarchical combination PRO endpoint using the Finkelstein-Schoenfeld method (Finkelstein and Schoenfeld 1999), with a dropout rate of 25%.

An upper limit of approximately 70% enrollment of male participants will be used to ensure a sufficiently large sample of female participants.

## 6. Supporting Documentation

### 6.1. Appendix 1: Clinical Trial Registry Analyses

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

Analyses provided for the Clinical Trial Registry requirements include the following:

- Summary of AEs, provided as a dataset which will be converted to an XML file. Both SAEs and “Other” non-SAEs are summarized: by treatment group, by MedDRA PT.
  - An AE is considered ‘Serious’ whether or not it is a TEAE.
  - An AE is considered in the “Other” category if it is both a TEAE and is not serious. For each SAE and “Other” AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event
    - the number of participants who experienced each event term, and
    - the number of events experienced.
- For each SAE, these additional terms are provided for EudraCT:
  - the total number of occurrences causally related to treatment
  - the total number of deaths, and
  - the total number of deaths causally related to treatment.
- Consistent with [www.ClinicalTrials.gov](http://www.ClinicalTrials.gov) requirements, “Other” AEs that occur in fewer than 5% of participants/subjects in every treatment group may be excluded.
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

Demographic table including the following age ranges required by EudraCT: adults (18 to 64 years), (65 to 85 years), and (85 years and over).

### 6.2. Appendix 2: Search Criteria for Special Safety Topics

#### Arrhythmias and cardiac conduction disorders

Treatment-emergent arrhythmias, arrhythmias and cardiac conduction disorders will be considered an AESI. The cardiovascular events will include clinically relevant rhythm and conduction disorders. The treatment-emergent arrhythmias and cardiac conduction disorders events will be included using the MedDRA PT contained in any of the following SMQs:

- 1) Arrhythmias:
  - a. For symptoms: *Arrhythmia related investigations, signs and symptoms* SMQ (20000051), narrow and broad terms

- b. For supraventricular arrhythmias: In *Cardiac arrhythmia* SMQ, under tachyarrhythmia sub SMQ
  - i. *Supraventricular tachyarrhythmia* SMQ (20000057), broad and narrow terms
  - ii. *Tachyarrhythmia terms, nonspecific* SMQ (20000164), narrow terms only; and
  - iii. *Ventricular tachyarrhythmia* SMQ (20000058), narrow terms only.
- 2) Cardiac conduction disorders
  - a. *Conduction defects* SMQ (20000056), narrow terms only; and
  - b. *Cardiac conduction disorders* HLT (10000032), all PTs.

### **Hepatic TEAEs**

Treatment-emergent potentially drug-related hepatic disorders will be summarized by treatment using the MedDRA PTs contained in any of the following SMQs:

- broad and narrow terms in the *Liver related investigations, signs and symptoms* SMQ (20000008)
- broad and narrow terms in the *Cholestasis and jaundice of hepatic origin* SMQ (20000009)
- broad and narrow terms in the *Hepatitis non-infections* SMQ (20000010)
- broad and narrow terms in the *Hepatic failure, fibrosis and cirrhosis and other liver damage* SMQ (20000013), and
- narrow terms in the *Liver-related coagulation and bleeding disturbances* SMQ (20000015).

### **Acute gallbladder disease**

All events of TEAE biliary colic, cholecystitis, or other suspected events related to gallbladder disease will be summarized by treatment groups by PT with decreasing frequency under the following SMQs:

- narrow PTs in *Gallbladder related disorders* SMQ (20000124)
- narrow PTs in *Biliary tract disorders* SMQ (20000125), and
- narrow PTs in *Gallstone related disorders* SMQ (20000127).

### **Major depressive disorder/suicidal ideation**

AEs will be searched using MedDRA PTs from SMQs narrow scope: 20000037 (*Suicide/self-injury*) and 20000167 (*Depression [excl suicide and self injury]*).

### **C-cell hyperplasia and thyroid malignancies**

Thyroid malignancies and C-cell hyperplasia will be identified using MedDRA HLT for *Thyroid neoplasms* and PT for *Thyroid C-cell hyperplasia*.

### Hypersensitivity reactions

Analyses are based on the following:

- narrow and algorithm terms in *Anaphylactic reaction* SMQ (20000021)
- narrow terms in *Angioedema* SMQ (20000024)
- narrow terms in *Severe cutaneous adverse reactions* SMQ (20000020), and
- narrow terms in *Hypersensitivity* SMQ (20000214).

For the Anaphylactic reaction SMQ, each term is classified by scope (narrow, broad) and by category (A, B, C, and D). All narrow terms are category A, and all broad terms are category B, C, or D. In addition to the usual narrow and broad searches, the SMQ defines an algorithm to further refine the cases of interest. For Time Period A analysis, the *Anaphylactic reaction* SMQ algorithm will be included. The algorithm is based upon events that occur within Time Period A. The counts and percentages of participants who experienced a TEAE for the following will be analyzed for each of the 2 time periods:

- any narrow term from any 1 of the 4 SMQs indicated above (that is, combined search across narrow of all 4 SMQs), and
- any narrow scope term within each SMQ, separately (that is, narrow SMQ search). For Time Period A analysis, any term from *Anaphylactic reaction* SMQ algorithm.

### Injection site reactions

The ISR AE will be identified using the MedDRA PT in any of the following:

- HLT of *Injection site reaction*
- HLT of *Administration site reaction*, and
- HLT of *Infusion site reactions*.

### Pancreatitis events

Determination of investigator-reported events will be through the *Acute pancreatitis* MedDRA SMQ (20000022, narrow scope) and a *Chronic pancreatitis* PT search of the AE database, while adjudication-confirmed pancreatitis is found from adjudication forms.

## 6.3. Appendix 3: Magnetic Resonance Imaging Addendum

This section is applicable to the participants who are enrolled in the MRI addendum.

This addendum applies to a subset of participants (approximately 58 participants) enrolled in ISA1. MRIs for the assessment of fat dispositions will be collected at baseline and Week 52. The MRI at baseline needs to be completed prior to Visit 2 or within 7 days after Visit 2. The MRI at Week 52 may be scheduled for any day  $\pm$  14 days.

MRI analyses will be guided by the treatment policy strategy and conducted among all participants who are enrolled in the addendum, received at least 1 dose of study drug, and have baseline and at least 1 postbaseline MRI measurement. No imputation will be performed for



missing data. The participant's demographics and baseline characteristics for the MRI addendum will be summarized.

Objectives	Endpoints	Analytical Approaches
Compare the effect of once weekly tirzepatide at MTD versus placebo on the changes of soft tissues volumes, fat volumes and fat content (%) in upper airway structures and in the abdomen in participants with OSA and obesity.	Changes of absolute soft tissue volume, fat volume and fat content (%) of the following: <ul style="list-style-type: none"> <li>• Tongue</li> <li>• Soft palate</li> <li>• Pterygoid muscle</li> <li>• Lateral pharyngeal walls</li> <li>• Neck and submandibular area</li> <li>• Total, visceral and subcutaneous abdominal fat</li> </ul>	Change from baseline to Week 52 for each parameter will be compared between treatment arms using an ANCOVA approach. The model will include treatment, the stratification factors of gender, and baseline AHI (not severe/severe), and the baseline value for the parameter. Summary statistics for MRI parameters at baseline and at Week 52 will be provided. The treatment comparison at baseline will be performed using an ANOVA model.
Explore correlation of changes of soft tissue volumes, fat volumes, and fat content (%) in upper airway structures and in the abdomen with changes of AHI.	Correlations between the change in absolute soft tissue volume, fat volume, and fat content (%) for the structures listed above and the % change in AHI.	Spearman correlations between the change from baseline for each of the MRI endpoints and the % change in AHI will be calculated.

Abbreviations: AHI = Apnea-Hypopnea Index; ANCOVA = analysis of covariance; ANOVA = analysis of variance; MRI = magnetic resonance imaging; MTD = maximum tolerated dose; OSA = obstructive sleep apnea.

#### 6.4. Appendix 4: Statistical Analysis for China

Analyses will be performed for the following subpopulations:

- participants enrolled in East Asian countries/regions (China, Japan), and
- participants enrolled in China.

The analysis methods for the above-mentioned subgroups will be similar to those described for the main part of this SAP. If there is not a sufficient number of participants in the subpopulation, summary statistics will be provided.

The analyses to be included will be documented in a separate list of analyses which should include disposition, demographics, and selected efficacy and safety endpoints.

#### 6.5. Appendix 5: Statistical Analysis for Japan

Analyses will be performed for the following subpopulations:

- participants enrolled in Japan, and
- the JASSO subpopulation, participants who meet the criteria of the JASSO (not limited to the participants enrolled in Japan).

The JASSO subpopulation analysis will be performed according to the criteria of both BMI and obesity-related health problems according to the treatment flow of obesity disease in the obesity disease treatment guideline (JASSO 2022). The JASSO guideline states that drug treatment in

participants with obesity disease should follow with a BMI above or equal to 25 kg/m<sup>2</sup> (in this trial, BMI enrollment was started from 27 kg/m<sup>2</sup>), and at least 2 obesity-related health problems, or a BMI above or equal to 35 kg/m<sup>2</sup> and at least 1 obesity-related health problem out of the 11 obesity-related health problems, including OSAS, listed below. The overall population and participants with obesity disease according to the JASSO guideline will be compared.

The analysis methods for the above-mentioned subgroups will be similar to those described for the main part of this SAP. If there is not a sufficient number of participants in the subpopulation, summary statistics will be provided.

As a low number of participants were enrolled from Japan, combined analyses with both ISAs may be conducted to explore a future line extension for the OSAS indication in Japan.

The analyses to be included will be documented in a separate list of analyses which should include disposition, demographics, and selected efficacy and safety endpoints.

### **Eleven obesity-related health problems**

The JASSO guideline defines 11 health problems for the diagnosis of “Obesity Disease” in subjects who need weight reduction for a medical reason. Data collected by a specific CRF will be used to specify the obesity-related health problems for each participant. The list of the 11 health problems are:

- 1) Glucose intolerance disorder (type 2 diabetes, impaired glucose tolerance [IGT], and so on)
- 2) Dyslipidemia
- 3) Hypertension
- 4) Hyperuricemia and Gout
- 5) Cardiovascular disease, myocardial infarction and Angina
- 6) Cerebral infarction and transient ischemic attack (TIA)
- 7) Non-alcoholic fatty liver disease (NAFLD)
- 8) Menstruation disorder and infertility
- 9) Obstructive sleep apnea syndrome (OSAS) and obesity-hypoventilation syndrome
- 10) Motor dysfunction: arthritis/osteoarthritis (knee, hip joint, supine, and so on), and
- 11) Obesity-related renal disease.

## 7. References

- Alosh M, Bretz F, Huque M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693-713. <https://doi.org/10.1002/sim.5974>
- Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Stat Med*. 2009;28(4):586-604. <https://doi.org/10.1002/sim.3495>
- Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913. <https://doi.org/10.1002/bimj.201000239>
- Bretz F, Xi D. Commentary on “Statistics at FDA: reflections on the past six years”. *Stat Biopharm Res*. 2019;11(1):20-25. <https://doi.org/10.1080/19466315.2018.1562369>
- [ADA] American Diabetes Association. Glycemic targets: standards of medical care in diabetes—2020. *Diabetes Care*. 2020;43(suppl 1):S66–S76. <https://doi.org/10.2337/dc20-S006>
- Danne T, Philotheou A, Goldman D, et al. A randomized trial comparing the rate of hypoglycemia – assessed using continuous glucose monitoring – in 125 preschool children with type 1 diabetes treated with insulin glargine or NPH insulin (the PRESCHOOL study). *Pediatr Diabetes*. 2013;14(8):593-601. <https://doi.org/10.1111/pedi.12051>
- Donovan LM, Yu L, Bertisch SM, et al. Responsiveness of patient-reported outcomes to treatment among patients with type 2 diabetes mellitus and OSA. *Chest*. 2020;157(3):665-672. <https://doi.org/10.1016/j.chest.2019.11.011>
- Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. *Stat Med*. 1999;18(11):1341-1354. [https://doi.org/10.1002/\(sici\)1097-0258\(19990615\)18:11<1341::aid-sim129>3.0.co;2-7](https://doi.org/10.1002/(sici)1097-0258(19990615)18:11<1341::aid-sim129>3.0.co;2-7)
- [JASSO] Japan Society for the Study of Obesity. Guideline for the management of obesity disease. Tokyo: Life Science Publishing. 2022. (in Japanese)
- Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540-545. <https://doi.org/10.1093/sleep/14.6.540>
- [Northwestern]. Northwestern University. PROMIS Short Form v1.0 – Sleep-Related Impairment 8a. Published May 09, 2016a. Accessed December 14, 2022. [https://www.healthmeasures.net/index.php?option=com\\_instruments&view=measure&id=185&Itemid=992](https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=185&Itemid=992)
- [Northwestern]. Northwestern University. PROMIS Short Form v1.0 – Sleep Disturbance 8b. Published May 03, 2016b. Accessed December 14, 2022. [https://www.healthmeasures.net/index.php?option=com\\_instruments&view=measure&id=184&Itemid=992](https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=184&Itemid=992)
- Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. 2012;33(2):176-182. <https://doi.org/10.1093/eurheartj/her352>
- Rubin DB. *Multiple imputation for nonresponse in surveys*. John Wiley & Sons Inc.; 1987.

- Vandemeulebroecke M, Häring DA, Hua E, et al. New strategies for confirmatory testing of secondary hypotheses on combined data from multiple trials [published online February 4, 2024]. *Clin Trials*. <https://doi.org/10.1177/17407745231214382>
- Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*. 1997;20(10):835-843. <https://doi.org/10.1093/SLEEP%2F20.10.835>
- Weinberg ME, Bacchetti P, Rushakoff RJ. Frequently repeated glucose measurements overestimate the incidence of inpatient hypoglycemia and severe hyperglycemia. *J Diabetes Sci Technol*. 2010;4(3):577-582. <https://doi.org/10.1177/193229681000400311>
- Ye T, Bannick M, Yi Y, Shao J. Robust variance estimation for covariate-adjusted unconditional treatment effect in randomized clinical trials with binary outcomes. *Stat Theory Relat Fields*. 2023;7(2):159-163. <https://doi.org/10.1080/24754269.2023.2205802>

Signature Page for VV-CLIN-117998 v3.0

Approval

**PPD**

Signature Page for VV-CLIN-117998 v3.0

Approved on 05 Apr 2024 GMT

# 1. Integrated Efficacy Analysis Plan for LY3298176

Tirzepatide (LY3298176)  
Obstructive Sleep Apnea

Analysis Plan for the integrated efficacy analyses to be included in the Summary of Clinical Efficacy for LY3298176 for Obstructive Sleep Apnea.

Eli Lilly and Company  
IEAP Version 2

## Confidential Information

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

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

**Document ID:** VV-CLIN-133165

## 2. Table of Contents

<b>Section</b>	<b>Page</b>
1. Integrated Efficacy Analysis Plan for LY3298176 .....	1
2. Table of Contents .....	2
3. Revision History .....	3
4. Objective and Study Designs.....	4
4.1. Objective.....	4
4.2. Overview of Individual Studies Evaluating Clinical Efficacy .....	4
4.3. Analysis Sets .....	5
5. A Priori Statistical Methods .....	6
5.1. General Considerations .....	6
5.2. Handling of Dropouts or Missing Data .....	6
5.3. Comparison of Results of Individual Studies.....	6
5.3.1. Disposition .....	6
5.3.2. Demographics and Baseline Characteristics .....	7
5.3.3. Primary and Important Secondary Endpoints .....	7
5.4. Pooled Analysis of Data from More than One Study .....	7
5.5. Comparison of Results in Subpopulations .....	7
5.6. Analysis of Clinical Information Relevant to Dosing Recommendation.....	7
5.7. Time Course of Effect, Persistence of Effect, and/or Tolerance, Distribution of Responses.....	7
6. References .....	8

### 3. Revision History

<b>IEAP Version</b>	<b>Approval Date</b>	<b>Change</b>	<b>Rationale</b>
1	October 16, 2023	Not Applicable	Original version
2	See Date on Page 1	Added pooled analysis combining ISA1 and ISA 2 or selected key secondary endpoints	Potential inadequate statistical power to assess hypotheses within each ISA



## 4. Objective and Study Designs

### 4.1. Objective

The objective of the integrated efficacy analysis plan (IEAP) is to specify the planned integrated efficacy analysis of tirzepatide (TZP) for the indication of obstructive sleep apnea (OSA). This document describes the integrated analysis intended to be included in Summary of Clinical Efficacy (SCE)/Integrated Summary of Efficacy as part of the Common Technical Document for submission.

This document includes

- strategies for combining studies
- analyses to be performed, and
- statistical methods to be used

Lilly does not plan to pool any Intervention-Specific Appendices (ISAs) for the primary efficacy assessment and closely related key secondary endpoints because the individual ISAs are adequately powered to assess the primary efficacy objective. Lilly is planning an integrated analysis for selected efficacy parameters potentially not adequately powered within each ISA using pooled data from Study GPI1 (ISA1) and Study GPI2 (ISA2), while strongly controlling the submission-wise type 1 error rate (SWER), that is, the probability to make a false claim of success for an endpoint at the submission level (Vandemeulebroecke et al. 2024).

The side-by-side presentation of primary efficacy results across individual ISAs are planned to be included in the SCE. This document will not detail the methods to be used for analyses in individual ISAs. Those analyses will be described in the statistical analysis plan (SAP) for the study.

### 4.2. Overview of Individual Studies Evaluating Clinical Efficacy

Phase 3 study/ISAs for patients with OSA that will be included in the SCE are briefly described below.

Study I8F-MC-GPIF (GPIF) is a multicenter, randomized, parallel-arm, double-blind, placebo-controlled Phase 3 study to evaluate the efficacy and safety of TZP at the maximum tolerated dose (MTD) (10 mg or 15 mg) once weekly (QW) versus placebo in participants who have obesity and moderate-to-severe OSA.

This basket-type master protocol investigates 2 participant populations, described in 2 ISAs:

- Study GPI1 includes participants who are unwilling or are unable to use positive airway pressure (PAP) therapy.
- Study GPI2 includes participants who are on PAP therapy for at least 3 months at time of screening and plan to continue PAP therapy during the study.

Participants are to be assigned to whichever ISA they qualify for. Participants will then be randomly assigned to

- TZP at the maximum tolerated dose (10 mg or 15 mg) subcutaneously (SC) once weekly (QW), or
- placebo.

The expected total duration of study participation for each participant, including screening and the posttreatment follow-up periods, is 60 weeks across the following study periods:

- Screening: 4 weeks
- Treatment: 52 weeks
- Posttreatment follow-up: 4 weeks

The maximum duration of treatment is 52 weeks.

Efficacy measures presented in the SCE for the primary efficacy endpoint and closely related key secondary endpoints will be derived from individual ISAs for Study GPIF, without patient-level integration of data from individual ISAs. The details of these analyses are provided in the respective Study GPIF SAP.

The primary and key secondary endpoints from individual ISAs for the Study GPIF are provided in the SAP.

### **4.3. Analysis Sets**

For individual ISAs, the analysis sets are specified in the Study GPIF SAP. Pooled analyses will be conducted using the same analysis set as the set used in the individual ISAs specified in the Study GPIF SAP.

## 5. A Priori Statistical Methods

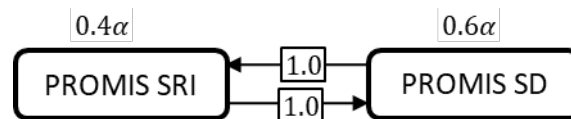
### 5.1. General Considerations

Statistical analysis will be the responsibility of Eli Lilly and Company or their appointed contract research organization.

Treatment group will be based on assignment at randomization. The multiplicity adjustment within each ISA is described in the SAP for Study GPIF. The sample size of individual ISAs was guided by the anticipated efficacy relative to the primary endpoint. Due to potentially inadequate power within each ISA, a pooled analysis will be conducted for the following key secondary endpoints:

- change from baseline in Patient-Reported Outcomes Measurement Information System (PROMIS) Sleep-Related Impairment (SRI) score at Week 52
- change from baseline in PROMIS Sleep Disturbance (SD) Score at Week 52

The pooled analyses described above are subject to the SWER (Bretz and Dong 2019; Vandemeulebroecke et al. 2024) of  $\alpha = 0.04875 = 2 \times (0.025 - 0.025^2)$  if the primary endpoint is achieved under the family wise error rate (FWER) control in both ISAs, and  $\alpha = 0.05$  if all endpoints subject to the FWER control of endpoint is achieved in both ISAs. [Figure GPIF 5.1](#) illustrates the graphical testing strategy for evaluating the key secondary endpoints subject to the SWER control.



Abbreviations: PROMIS = Patient-Reported Outcomes Measurement Information System; SD = Sleep Disturbance; SRI = Sleep-Related Impairment.

**Figure GPIF 5.1 Graphical testing scheme for pooled analysis.**

### 5.2. Handling of Dropouts or Missing Data

The details are described in the SAP for Study GPIF. The retrieved dropout imputation will be performed within the pooled treatment group across ISAs.

### 5.3. Comparison of Results of Individual Studies

There is no statistical comparison of results of individual ISAs.

#### 5.3.1. Disposition

Patient disposition will be summarized by treatment group for each ISA. The details are specified in the Study GPIF SAP, which is aligned with Program Safety Analysis Plan (PSAP).

### **5.3.2. Demographics and Baseline Characteristics**

Demographics and other baseline characteristics will be summarized by treatment group for each ISA. The details are specified in the Study GPIF SAP, which is aligned with PSAP.

### **5.3.3. Primary and Important Secondary Endpoints**

The statistical analysis method for the primary and key secondary efficacy endpoints are specified in the Study GPIF SAP.

The primary and key secondary efficacy results from each individual ISAs will be presented side-by-side to support the SCE. Selected key secondary endpoints to conduct the pooled analysis specified in Section 5.1 will also be presented in the SCE.

## **5.4. Pooled Analysis of Data from More than One Study**

All analyses will be performed and align to the “treatment regimen” estimand as described in GPIF SAP. Analysis methods and imputation approaches for handling missing/invalid data are in line with the individual ISAs specified in the Study GPIF SAP. Following imputation, all endpoints will be analyzed from the analysis of covariance model with treatment, ISA [ISA1/ISA2], geographic region [US/OUS], Apnea-Hypopnea Index (AHI) stratum (not severe [AHI <30]/severe [AHI ≥30]), and gender as fixed effects, with baseline as a covariate, using the pooled full analysis set (FAS) in each ISA. Statistical inference over multiple imputed data sets will be guided by Rubin’s rule (1987).

## **5.5. Comparison of Results in Subpopulations**

The subgroup analyses at the ISA level will be used to support the SCE. The details are specified in the Study GPIF SAP.

## **5.6. Analysis of Clinical Information Relevant to Dosing Recommendation**

The relationship between efficacy outcomes and pharmacokinetics (PK) may be presented. The details of these analyses are included in the population PK/pharmacodynamic analysis plan.

## **5.7. Time Course of Effect, Persistence of Effect, and/or Tolerance, Distribution of Responses**

Therapeutic effects of a treatment can decline over time because of tolerability issues (patients who experience adverse events and refuse treatment), from the development of drug resistance or tolerance, or because the disease tends to resolve spontaneously (FDA 2015).

Studies GPI1 and GPI2 have primary objectives at 1 year of treatment. These studies will provide evidence of the long-term sustainability of effect for TZP.

## 6. References

Bretz F, Dong X. Commentary on “Statistics at FDA: Reflections on the Past Six Years”.  
*Stat Biopharm Res.* 2019;11(1):20-25. <https://doi.org/10.1080/19466315.2018.1562369>

[FDA] United States Food and Drug Administration. Integrated summary of effectiveness: guidance for industry. October 2015. Accessed October 12, 2023.  
<https://www.fda.gov/media/72335/download>

Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: Wiley; 1987.

Vandemeulebroecke M, Häring DA, Hua E, et al. New strategies for confirmatory testing of secondary hypotheses on combined data from multiple trials [published online February 4, 2024]. <https://doi.org/10.1177/17407745231214382>

Signature Page for VV-CLIN-133165 v2.0

Approval	<b>PPD</b>
	27-Feb-2024 16:12:18 GMT+0000

Signature Page for VV-CLIN-133165 v2.0

Approved on 27 Feb 2024 GMT