Protocol C3421019

A PHASE 2B, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP, DOSE-RANGING STUDY TO EVALUATE THE EFFICACY AND SAFETY OF PF-06882961 ADMINISTRATION IN ADULTS WITH OBESITY

Statistical Analysis Plan (SAP)

[redacted]

Version: **2.0**

Date: 25 Oct 2023

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDICES

1. VERSION HISTORY

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
2 25Oct2023	Amendment 2 16Sept2022	Updates made based on addition of Cohorts 2 and 3 and Blinded Data Reviews	Rationale: The protocol was amended to add cohorts 2 and 3 of approximately 49 and 112 additional participants, respectively. The following sections were revised as part of this rationale: • Section 1.2: Endpoint column updated to reflect new timepoint schedule in Cohort 3. Text updated to reflect change of treatment duration for Cohort 3. • Section 1.3: added text verbatim and updated study design figures from the protocol amendment to reflect the changes to the protocol. • Section 2: Change in definition of end of treatment for Cohort 3. Update to text to reflect different timepoints for Cohort 3. • Section 4.2: Labels for titration arms updated and new titration arms for Cohort 3 added. • Sections 4.2 and 5: clarified that data from participants in the Cohort 1 and Cohort 2 will be combined and reported together unless otherwise stated and that Cohort 3 would be reported separately also unless otherwise stated. • Section 5.7.1: added text verbatim from the protocol amendment to specify that two separate baseline summary tables will be produced; one for Cohorts 1 & 2 and one for Cohort 3 Rationale: Changes to reflect updates based on ongoing blinded review of tables, to also provide consistency with other C342 studies. The following sections were revised as part of this rationale: • Section 2.3: t-scores based on SF-36 results won't be calculated using

			Optum scoring solution software, but internal code based on publications per internal approach. • Section 2.3.2.4: added text to specify that both pre- and post-dose populations will be from the safety analysis set. • Section 4.3: added text to specify that for the analysis of safety endpoints, the sponsor data standard rules for imputation of BLQ values will be applied. • Section 5.4.1.3: added wording to include discontinuation from IP as well as from study in exclusion from the denominator for the percentage of TEAEs of interest by week. Rationale: additional outputs needed for reporting. The following sections were revised as part of this rationale: • Sections 2.3.2.3.1 & 5.4.2.1: Bile acids and GGT have been included in the list of liver function tests for which the change from baseline is a safety laboratory endpoint of interest. • Section 5.1.1.1: added text to specify that absolute changes from baseline in body weight should be summarized as part of the main analysis. • Section 5.7.1: hip circumference has been added to be summarized at baseline. • Section 5.7.2: added text to specify that an additional table summarizing participant discontinuations due to adverse events for each system class and preferred term will be produced by treatment group and overall. • Sections 5.7.1, 5.4.2, 5.4.2.2, 5.4.3 & 5.4.4: additionally summarising parameters by overall Rationale: erroneously missed endpoints in version 1 of SAP. The following sections were revised as part of this rationale:

			• Section 5.4.1: added text that the number of participants who met the criteria for referral to a mental health professional will be listed and summarized by treatment group and time point. Minor formatting errors were also addressed (e.g. changing “Danuglipron” to “danuglipron” when included in middle of sentence).
1 10 Feb 2021	Original 14 Sep 2020	N/A	N/A

2. INTRODUCTION

This statistical analysis plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study C3421019. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

2.1. Modifications to the Analysis Plan Described in the Protocol

None.

2.2. Study Objectives, Endpoints, and Estimands

Type	Objective	Endpoint	Estimand
Primary Efficacy	*To compare the effect of multiple dose levels of PF-06882961 versus placebo on body weight in participants with obesity.*	*Percent CFB in body weight at End of Treatment.*[a]	*Estimand 1: This estimand is intended to provide a population level estimate of the mean treatment effect (PF-06882961 versus placebo) on a continuous endpoint in all evaluable participants while on treatment*
Secondary (Safety)	*To characterize the safety and tolerability of multiple dose levels of PF-06882961 administered to participants with obesity.*	• *Incidence of treatment emergent AEs [AEs and SAEs], and clinically significant abnormal laboratory, vital signs and ECG parameters.* *Assessment of mental health as determined by C-SSRS and PHQ-9.*	*There are no defined estimands for these endpoints and they will be analyzed using Pfizer data standards as applicable.*
Secondary (Efficacy)	*To compare the effect of multiple dose levels of PF-06882961 versus placebo on additional parameters of body weight in participants with obesity.*	• *Response as defined by a body weight loss of ≥5% from baseline at End of Treatment.*[a]	*Estimand 2: This estimand is intended to provide a population level estimate of the odds ratio treatment effect (PF-06882961 versus placebo) on a binary endpoint in all evaluable participants while on treatment.*

		• ***Cohorts 1 and 2****: Percent CFB in body weight at Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22.* • ***Cohort 3****: Percent CFB in body weight at Weeks 4, 8, 12, 16, 20, 24 and 28.*	*Estimand 1 as above.*
		• *Absolute CFB in waist circumference at End of Treatment.*[a]	*Estimand 3: This estimand is intended to provide a population level estimate of the mean treatment effect (PF-06882961 versus placebo) on a continuous endpoint in all evaluable participants while on treatment.*
		• *Absolute CFB in waist-to-hip ratio at End of Treatment.*[a]	*Estimand 4: This estimand will be similar to 3 above.*
Secondary (Efficacy)	*To compare the effect of multiple dose levels of PF-06882961 versus placebo on glycemic parameters in participants with obesity.*	• ***Cohorts 1 and 2****: Absolute CFB in HbA1c at Weeks 16 and 26.* • ***Cohort 3****: Absolute CFB in HbA1c at Weeks 16, 24 and 32.*	*Estimand 5: This estimand will be similar to 3 above.*
		• *Absolute CFB in FPG at each planned in-clinic study visit up through the End of Treatment visit.*[b]	*Estimand 6: This estimand will be similar to 3 above.*
Tertiary/ Exploratory	*To compare the effect of multiple dose levels of PF-06882961 versus placebo on body weight in participants with obesity who do not have major protocol deviations.*	• *Percent CFB in body weight at each planned in-clinic study visit up through the End of Treatment visit.*[b]	*Estimand 7: This estimand is intended to provide a population level estimate of the mean treatment effect (PF-06882961 versus placebo) on a continuous endpoint in evaluable participants with no major protocol deviations and while on treatment.*
Tertiary/ Exploratory	*To characterize the PK of PF-06882961 in participants with obesity.*	• *Trough plasma concentrations of PF-06882961 at time points specified in the SoA for the applicable cohort (see SoA for Cohorts 1 and 2, and SoA for Cohort 3).*	*There is no defined estimand for this endpoint and this will be analyzed using Pfizer data standards as applicable.*
Tertiary/ Exploratory	*To compare the effect of multiple dose levels of PF-06882961 versus placebo on markers of insulin resistance in participants with obesity.*	• ***Cohorts 1 and 2****: CFB in fasting plasma insulin, HOMA-IR and HOMA-B at Weeks 4, 8, 12, 16, 22 and 26.* • ***Cohort 3****: CFB in fasting plasma insulin, HOMA IR and HOMA-B at Weeks 4, 8, 12, 16, 20, 24, 28 and 32.*	*There are no defined estimands for these endpoints and they will be analyzed using Pfizer data standards as applicable.*
Tertiary/	*To compare the effect of multiple dose levels of*	• ***Cohorts 1 and 2****: Absolute CFB in waist circumference*	*Estimand 3: This estimand is defined above.*

Exploratory	*PF-06882961 versus placebo on additional parameters of body weight in participants with obesity.*	*at Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22.* • ***Cohort 3**: Percent CFB in body weight at Weeks 4, 8, 12, 16, 20, 24 and 28.*	
		• ***Cohorts 1 and 2**: Absolute CFB in waist-to-hip ratio at Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22.* • ***Cohort 3**: Percent CFB in body weight at Weeks 4, 8, 12, 16, 20, 24 and 28.*	*Estimand 4: This estimand will be similar to 3 above.*
	To compare the effect of multiple dose levels of PF-06882961 versus placebo on glycemic category in participants with obesity.	• *Shift from baseline in glycemic category (normoglycemia, pre-diabetes, or T2DM) at End of Treatment.[a]*	*There are no defined estimands for these endpoints and they will be analyzed using Pfizer data standards as applicable.*
	To compare the effect of multiple dose levels of PF-06882961 versus placebo on blood pressure (BP) *and lipid levels in participants with obesity.*	• *CFB in systolic and diastolic BP at each planned in-clinic study visit up through the End of Treatment visit.* • ***Cohorts 1 and 2**: CFB in lipid parameters at Weeks 4, 8, 12, 16, 22 and 26.* • ***Cohort 3**: CFB in lipid parameters at Weeks 4, 8, 12, 16, 20, 24, 28 and 32.*	*There are no defined estimands for these endpoints and they will be analyzed using Pfizer data standards as applicable.*
	To compare the effect of multiple dose levels of PF-06882961 versus placebo on PCOAs in participants with obesity.	• ***Cohorts 1 and 2:** CFB in SF-36v2® subscales and summary scores at Weeks 16 and 26.* • ***Cohort 3**: CFB in SF-36v2® subscales and summary scores at Weeks 16, 24 and 32.*	*There are no defined estimands for these endpoints and they will be analyzed using Pfizer data standards as applicable.*
	To compare the effect of multiple dose levels of PF-06882961 versus placebo on PCOAs in participants with obesity (US sites only).	• ***Cohorts 1 and 2**: CFB in Eating-Related Factors Daily Diary at Weeks 4, 8, 12, 16, 22 and 26.* • ***Cohort 3**: CFB in Eating-Related Factors Daily Diary at Weeks 4, 8, 12, 16, 20, 24, 28 and 32.* • *CFB in PGI-S at each planned in-clinic study visit up through the End of Treatment visit.[b]* • ***Cohorts 1 and 2**: CFB in PROMIS® Fatigue at Weeks 16 and 26.* • ***Cohort 3**: CFB in PROMIS® Fatigue at Weeks 16, 24 and 32.* • ***Cohorts 1 and 2**: CFB in PROMIS® Physical*	*There are no defined estimands for these endpoints and they will be analyzed using Pfizer data standards as applicable.*

		Function at Weeks 16 and 26. • ***Cohort 3**: CFB in PROMIS® Physical Function at Weeks 16, 24 and 32.* • ***Cohorts 1 and 2**: CFB in IWQOL-Lite-CT© at Weeks 16 and 26.* • ***Cohort 3**: CFB in IWQOL Lite-CT© at Weeks 16, 24 and 32.* • ***Cohorts 1 and 2**: PGI-C at Weeks 4, 16 and 26.* • ***Cohort 3**: PGI-C at Weeks 4, 16, 24 and 32.*	
***Note:** For all endpoints, baseline is defined as the result closest prior to dosing at V3 (Day 1).* *a. End of Treatment defined as Week 26 for Cohorts 1 and 2, and as Week 32 for Cohort 3.* *b. For Cohorts 1 and 2, this includes Weeks 2, 4, 6, 8, 10, 12, 16, 18, 22 and 26; for Cohort 3, this includes Weeks 4, 8, 12, 16, 20, 24, 28 and 32.*			

2.2.1. Primary Estimand(s)

The primary estimand (Estimand 1) will be the population average treatment effect on the percent change from baseline (CFB) *in body weight at End of Treatment of PF-06882961* (danuglipron) *compared to placebo in all evaluable participants while on treatment. For Cohorts 1 and 2, End of Treatment is defined as Week 26. For Cohort 3, End of Treatment is defined as Week 32. This reflects a combination of the 'Hypothetical' and 'While on treatment' strategies as outlined in the ICH-E9 (R1) guidance*.[1] It includes the following 4 attributes:

- Population: Defined by the inclusion and exclusion criteria to reflect participants with obesity
- Variable: Percent change from baseline in body weight at End of Treatment
- Intercurrent events: All data after an intercurrent event of discontinuation of investigational product (IP) will be censored
- Population-level summary: Mean percent change from baseline in each danuglipron arm compared to placebo

2.2.2. Secondary Estimand(s)

2.2.2.1. Estimand related to achieving Body Weight loss ≥5% at End of Treatment

A secondary estimand (Estimand 2) will be the population odds ratio of the treatment effect of achieving a body weight loss ≥5% from baseline at End of Treatment of PF-06882961 (danuglipron) *compared to placebo in all evaluable participants while on treatment. This reflects a combination of the 'Hypothetical' and 'While on treatment' strategies as outlined in the ICH-E9 (R1) guidance*.[1] It includes the following 4 attributes:

- Population: Defined by the inclusion and exclusion criteria to reflect participants with obesity

- Variable: Response as defined by a body weight loss of ≥5% from baseline at End of Treatment

- Intercurrent events: All data after an intercurrent event of discontinuation of investigational product (IP) will be censored. Missing or censored values at End of Treatment will be imputed (assuming missing at random [MAR]) to determine a response classification.

- Population-level summary: Odds ratio for a response between danuglipron (each arm considered separately) and placebo

2.2.2.2. Estimands related to Waist Circumference at End of Treatment

Estimand 3 is similar to Estimand 1, except the population average treatment effect is on the absolute CFB in waist circumference at End of Treatment. The population based treatment effect will be the mean CFB in each PF-06882961 (danuglipron) *arm compared to placebo.*

2.2.2.3. Estimands related to Waist-to-hip ratio, HbA1c and Fasting Plasma Glucose

Estimands 4, 5 and 6 will utilize the same approach as Estimand 3 (Section 2.2.2.2) *for the associated endpoint(s).*

2.2.3. Additional Estimand(s)

An additional exploratory estimand (Estimand 7) will be analyzed, which will be the population average treatment effect on the percent CFB in body weight at End of Treatment of PF-06882961 (danuglipron) *compared to placebo in evaluable participants with no major protocol deviations and while on treatment. This reflects a combination of the 'Hypothetical', 'While on treatment' and 'Principal stratum' strategies as outlined in the ICH-E9 (R1) guidance*.[1]

It includes the following 4 attributes:

- Population: Defined by the inclusion and exclusion criteria to reflect participants with obesity who do not have major protocol deviations

- Variable: Percent change from baseline in body weight at End of Treatment

- Intercurrent events: All data after an intercurrent event of discontinuation of investigational product (IP) will be censored

- Population-level summary: Mean percent change from baseline in each danuglipron arm compared to placebo

2.3. Study Design

This is a Phase 2b, multicenter, randomized, double-blind, placebo-controlled, parallel group, study to assess efficacy, safety, tolerability and PK of twice daily oral administration of PF-06882961 (danuglipron) *in adult participants with obesity. There are 3 Cohorts included in the study: Cohort 1 under the original protocol, Cohort 2 added under Amendment 1, and Cohort 3 added under Amendment 2.*

For all cohorts, following the initial screening period to confirm eligibility (up to 4 weeks), screening continues with a 2-week placebo run-in period which will be single-blinded (participant). Eligible participants who maintain acceptable compliance during the run-in period are randomized on Day 1 to double-blinded study intervention.

For participants in Cohorts 1 and 2, the total duration of dosing with double-blinded study intervention is 26 weeks, followed by an approximate 4-week follow-up. The total duration of participation in this study for the participants in Cohorts 1 and 2 is approximately 32 weeks, including the 2-week single-blind run-in period, but excluding the initial screening period.

For Cohort 3, added under protocol Amendment 2, participants will be randomized to placebo or 1 of 3 PF-06882961 (danuglipron) *arms with 4-week dose titration steps. The total duration of dosing with double-blinded study intervention is 32 weeks, followed by an approximate 4-week follow-up. The total duration of study participation for participants in Cohort 3 is approximately 38 weeks, including the 2-week single-blind run-in period, but excluding the initial screening period.*

All dosing regimens for all cohorts include dose titration to enhance tolerability of PF-06882961 (danuglipron)*, where the dose level is increased at set intervals of 1, 2 or 4 weeks until a target dose is achieved. For all cohorts, dosing is to occur with food twice daily. For the 5 1-week titration PF-06882961* (danuglipron) *groups in Cohort 1, the target dose levels for each of the arms are 40 mg BID, 80 mg BID, 120 mg BID, 160 mg BID, and 200 mg BID, with titration from a starting dose of 10 mg BID, and up to 10 weeks of the dosing duration used for titration. For the 3 2-week titration PF-06882961* (danuglipron) *groups in Cohorts 1 and 2, target dose levels for each of the arms are 120 mg BID, 160 mg BID, and 200 mg BID, with titration from a starting dose of 10 mg BID, and up to 20 weeks used for titration. Both Cohorts 1 and 2 also include a placebo arm. For Cohort 3, participants will be enrolled into 4 study arms, and target dose levels are placebo and PF-06882961* (danuglipron) *doses of 80 mg BID, 140 mg BID, and 200 mg BID. For Cohort 3, titration will occur at 4-week intervals from a starting dose of 10 mg BID, and up to 20 weeks will be used for titration to reach the target dose.*

There will be total of approximately 581 participants across all 3 cohorts of this study. In Cohort 1, approximately 420 participants (approximately 60 in each of the placebo and 5 1-week titration PF-06882961 (danuglipron) *arms and approximately 20 participants in each of the 3 2-week titration arms) were randomized. All dosing arms, including placebo, were enrolled simultaneously in Cohort 1. Subsequently, in Cohort 2, approximately 49 additional participants were randomized to the placebo and 2-week titration arms (approximately 7 in the placebo and approximately 14 in each of the 3 2-week titration arms). Combining Cohorts 1 and 2, there are a total of approximately 469 participants randomized (approximately 67 in the placebo, approximately 60 each of the 5 1-week titration arms and approximately 34 in each of the 3 2-week titration arms). In Cohort 3, approximately 112 participants will be randomized to 4 additional study arms (approximately 16 in the placebo and approximately 32 in each of 3 PF-06882961* (danuglipron) *arms with 4-week titration steps).*

Eligible participants in the Cohort 1 will be randomized to double-blinded study intervention in 1 of 9 arms in a 3:3:3:3:3:3:1:1:1 ratio (placebo: 5 standard titration danuglipron arms: 3 2-week titration arms). Eligible participants in Cohort 2 will be randomized to double-blinded study intervention in 1 of 4 arms in a 1:2:2:2 ratio (1 placebo: 3 2-week titration arms). Similarly for Cohort 3, eligible participants will be randomized to double-blinded study

intervention in 1 of 4 arms in a 1:2:2:2 ratio (1 placebo: 3 4-week titration arms). *Randomization is stratified according to biological gender (female versus male), as both tolerability and efficacy may be different between men and women. This stratification is intended to balance the numbers of participants of either gender across treatment arms. In addition, no more than approximately 70% of the trial population will consist of 1 gender, in order to permit adequate representation of both men and women in the trial population.*

Figure 1: Overall Study Schema for Cohorts 1 and 2

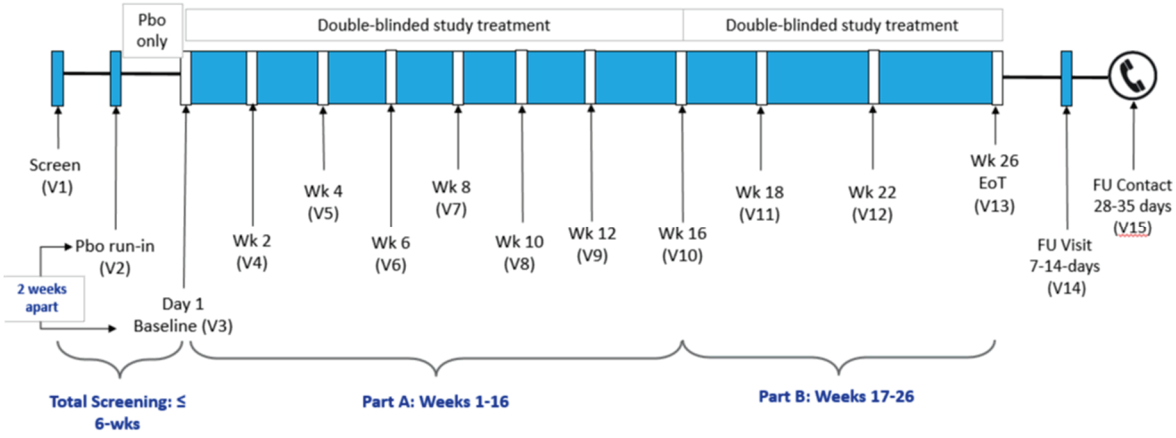


Figure 2: Overall Study Schema for Cohort 3

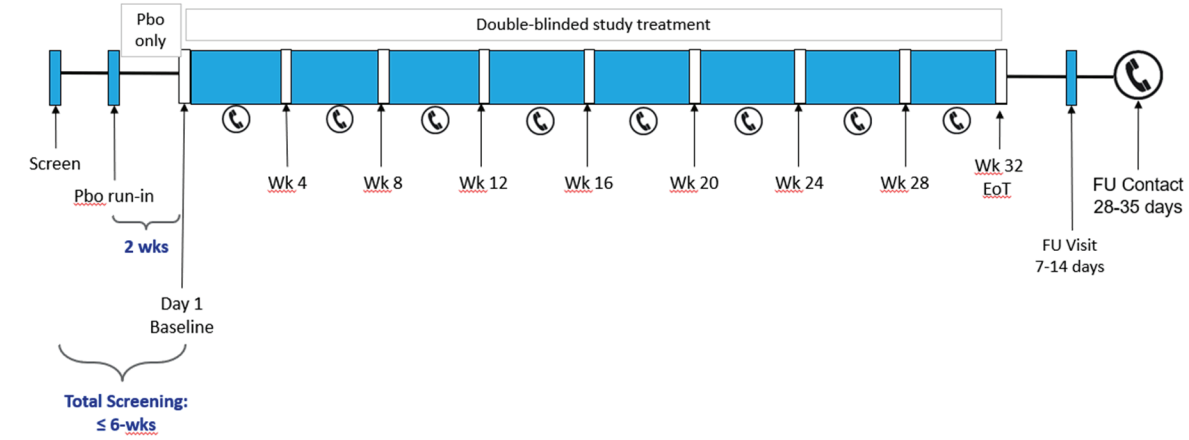


Figure 3: Titration Schema for Cohorts 1 and 2

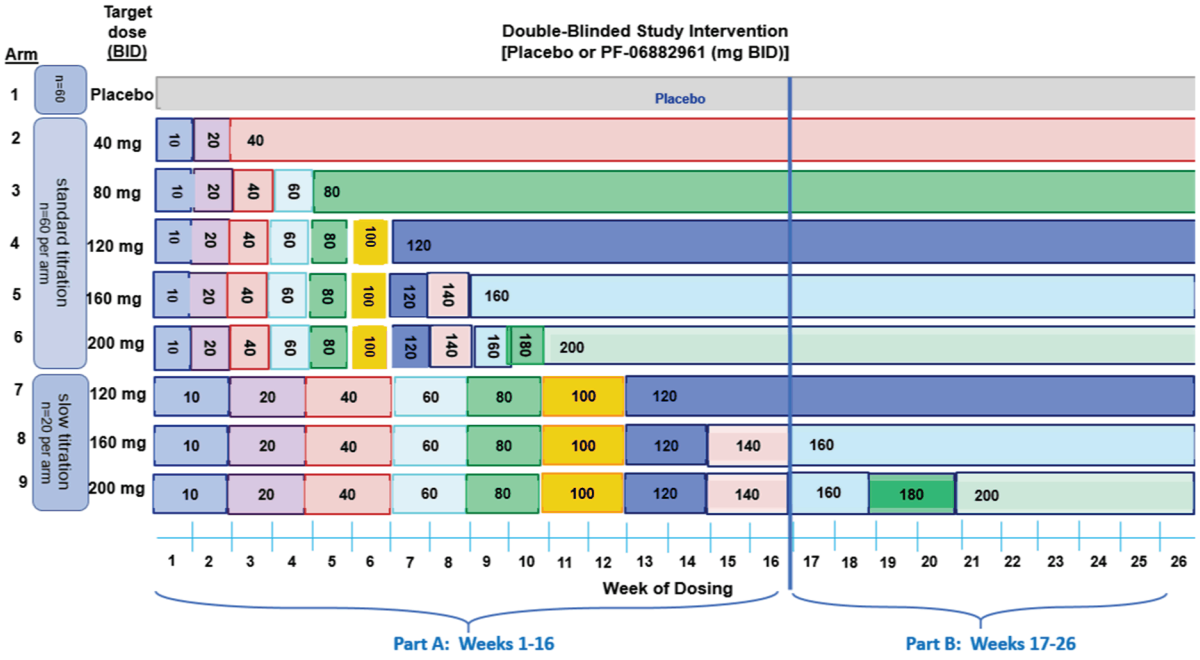
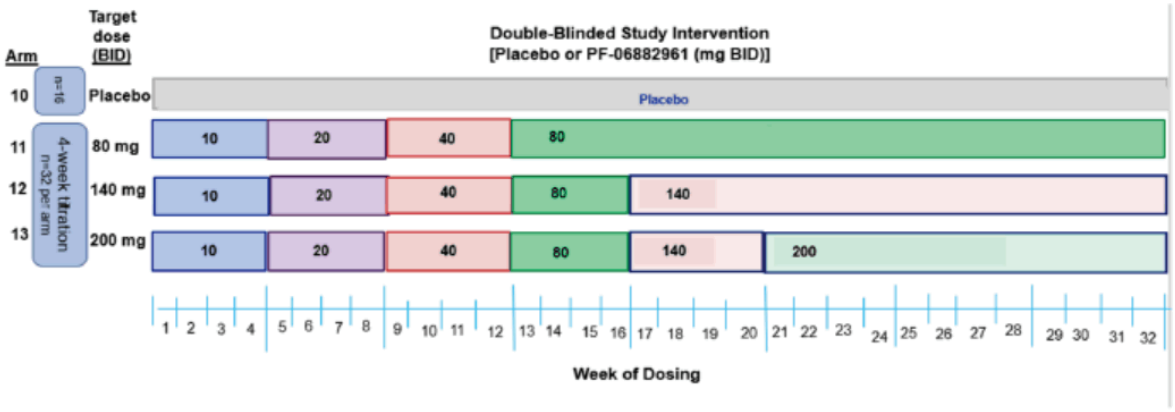


Figure 4: Titration Schema for Cohort 3



3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

For all endpoints, unless otherwise specified below, *baseline is defined as the result closest prior to dosing at V3 (Day 1).*

The "End of Treatment" definition is applied at the endpoint level (not the subject-level) and is defined for relevant endpoints as Week 26 for Cohorts 1 & 2, and as Week 32 for Cohort 3.

3.1. Primary Endpoint(s)

- Percent change from baseline in body weight at End of Treatment.

The average of the duplicate body weight readings collected at each assessment time will be calculated prior to summaries/analysis. If one of the two duplicates are missing, the non-

missing value will be used, and missing values will not be imputed. Both the absolute change and percent change from baseline in body weight will be calculated.

For the MMRM model, all body weight values (including baseline) will be $\log_e$-transformed prior to analysis (i.e. the outcome in the model will be the difference of the $\log_e$ absolute value at the time point of interest minus the $\log_e$ baseline). All LSMeans and LSMean differences (including confidence intervals) will be back-transformed to obtain the ratio to baseline or placebo (as applicable). The percent change from baseline or percent difference to placebo (as applicable) will then be calculated as follows:

Percent change = 100*(back-transformed LSMean – 1)

For descriptive summaries, percentage change from baseline will be calculated as the End of Treatment measurement minus the baseline value divided by the baseline value multiplied by 100.

The absolute change from baseline in body weight at End of Treatment will also be calculated for additional reporting.

3.2. Secondary Endpoint(s)

3.2.1. Efficacy Endpoint(s)

- Response as defined by a body weight loss of ≥5% from baseline at End of Treatment

This endpoint will have two levels: ‘Response’ and ‘Non-response’. The former will be based on participants having a body weight loss of ≥5% from baseline at End of Treatment, otherwise participants with a loss of <5% will be classed as having a ‘Non-response’. Participants with an intercurrent event of discontinuation of IP prior to, or on, the End of Treatment week will have their End of Treatment value censored (if not missing). Missing or censored values at End of Treatment will be imputed as described in Section 5.3. Percent change will be calculated as described in Section 3.1 (based on the descriptive summaries approach).

- Percent CFB in body weight at:
  - Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22 for Cohorts 1 and 2
  - Weeks 4, 8, 12, 16, 20, 24 and 28 for Cohort 3.

Data will be analysed on the natural log scale and the percent change from baseline will be calculated as described in Section 3.1.

- Absolute CFB in waist circumference at End of Treatment.

The average of the triplicate waist measurements collected at each assessment time will be calculated prior to summaries/analysis. If any of the triplicates are missing, the non-missing value(s) will be used, and missing values will not be imputed.

- Absolute CFB in waist-to-hip ratio at End of Treatment.

The average of the triplicate waist and hip measurements collected at each assessment time will be calculated prior to calculating the waist-to-hip ratio at each assessment time. If any of the triplicates are missing, the non-missing value(s) will be used, and missing values will not be imputed.

- Absolute CFB in HbA1c at:
    - Weeks 16 and 26 for Cohorts 1 and 2.
    - Weeks 16, 24 and 32 for Cohort 3.
- Absolute CFB in Fasting Plasma Glucose (FPG) at each planned in-clinic study visit up through the End of Treatment visit

3.2.2. Safety Endpoint(s)

3.2.2.1. Adverse Event

If an adverse event has a start time of 00:00 and a start date that is equal to the date of first administration of IP, the adverse event start time will be imputed with the time of the first dose of administration of IP.

If an adverse event has an end time of 00:00 then the end time will be imputed as 23:59.

An adverse event is considered treatment emergent (TEAE) relative to a given treatment if the event starts during the effective duration of treatment (i.e. starting on or after the first dose but before the last dose plus lag time).

The effective duration of treatment is determined by the lag time. Any event occurring within the lag time, whether this occurs during a break in treatment or at the end of treatment, is attributed to the corresponding treatment period. The lag time is defined by the Pfizer Standard of 365 days post last dose of IP.

Adverse events occurring during the placebo run-in period (i.e. starting from Day -14, inclusive, up to and before the first dose of active treatment on Day 1) will be considered non-treatment emergent.

A 3-tier approach will be used to summarize TEAEs. Under this approach, TEAEs are classified into 1 of 3 tiers. Different analyses will be performed for different tiers (see Section 6.3.1.

Tier 1 events: These are prespecified events of clinical importance and are maintained in a list in the product's Safety Review Plan.

Tier 2 events: These are events that are not tier 1 but are "common." A Medical Dictionary for Regulatory Activities (MedDRA) preferred term (PT) is defined as a Tier 2 event if there are at least 5% of participants reporting the event in any of the standard titration or placebo groups or at least 2 participants reporting the event in any of the slow titration groups.

Tier 3 events: These are events that are neither Tier 1 nor Tier 2 events.

3.2.2.1.1. Adverse events of interest

The TEAEs of interest for additional reporting are: nausea, vomiting and diarrhoea (as defined based on preferred term). Based on emerging blinded data reviews, other AEs of interest may be added to this list, which would be documented with a SAP amendment or documented in the changes to planned analysis section in the CSR.

3.2.2.1.2. Subset Reporting Interval

A summary of adverse events occurring during a subset of the main reporting interval will be based on a shorter reporting window. This interval will include all TEAEs (reported separately) if:

- the event starts after or on the first dose, but doesn't start more than two days after the last dose of IP (defined as either completing the 26 weeks or 32 weeks (depending on cohort) of treatment or discontinuing from IP earlier).

3.2.2.2. Hypoglycemic Monitoring

Hypoglycemia AEs will be recorded in the AE Case Report Form (CRF) with details of the event captured on the Hypoglycemic Event Details CRF. Details of when these will be recorded are given in the protocol Section 8.2.5.

For programming purposes, the hypoglycemic AE categories are based on the following:

- Severe Hypoglycemia: Severe is checked in the severity criteria of the CRF. This assessment will be made by the PI based on the protocol definition.

- Documented Symptomatic Hypoglycemia: If (1 – Did the participant have symptoms of hypoglycemia?) Yes and (2 – Was the blood glucose measured?) Yes and result <70 mg/dL on the CRF, but hypoglycemia is not classified as severe.

- Asymptomatic Hypoglycemia: If (1) No and (2) Yes and result <70 mg/dL on the CRF, but hypoglycemia is not classified as severe.

- Probable Symptomatic Hypoglycemia: If (1) Yes and (2) No and (2b – If blood glucose was not measured, did symptoms resolve when treated with carbohydrate or glucagon?) Yes on the CRF, but hypoglycemia is not classified as severe.

3.2.2.3. Laboratory Data

Safety laboratory tests (hematology, chemistry, urine testing and other clinical laboratory tests) will be performed as described in the protocol.

To determine if there are any clinically significant laboratory abnormalities, the safety tests will be assessed against the criteria specified in the sponsor reporting standards. The assessment will take into account whether each participant's baseline test result is within or outside the laboratory reference range for the particular laboratory parameter.

Baseline for all laboratory measurements will be defined as the result closest prior to dosing at Visit 3 (Day 1).

3.2.2.3.1. Change from Baseline Summaries

Focused change from baseline summaries (including both absolute changes from baseline and percent change from baseline, calculated separately) of the following safety laboratory endpoints will be assessed:

- Change from baseline in calcitonin to all post-dose time points as per the SOA
- Change from baseline in amylase to all post-dose time points as per the SOA
- Change from baseline in lipase to all post-dose time points as per the SOA
- Change from baseline in thyroid stimulating hormone (TSH) to all post-dose time points as per the SOA
- Change from baseline in free thyroxine (free T4) to all post-dose time points as per the SOA
- Change from baseline in lipid profile (total cholesterol, direct LDL cholesterol, HDL cholesterol and triglycerides) to all post-dose time points as per the SOA
- Change from baseline in liver function tests (alanine transaminase [ALT], aspartate aminotransferase [AST], alkaline phosphatase, total bilirubin, bile acids and gamma-glutamyl tranferase [GGT]) to all post-dose time points as per the SOA
- Change from baseline in estimated glomerular filtration rate (eGFR) to all post-dose time points as per the SOA

3.2.2.3.2. Clinical Laboratory Parameters of Interest

For the specific laboratory parameters listed in the table below, the following endpoints will be derived:

- Abnormalities defined as either a “Flag Level” or “Alert Level” as in the table below

Parameter	Flag Level	Alert Level	Conventional Units
Amylase	> ULN	Pfizer standard flag for PCC	U/L
Calcitonin	> ULN	Pfizer standard flag for PCC	ng/L
Lipase	> ULN	Pfizer standard flag for PCC	U/L
HbA1c	-	>10	%
Fasting Plasma Glucose	-	< 54	mg/dL
	-	> 240	
Alanine aminotransferase	> ULN > 2x ULN > 3x ULN > 5x ULN > 5x ULN > 5x ULN	- - - >8x ULN >10x ULN >20x ULN	

Aspartate Aminotransferase	> ULN ≥ 2x ULN ≥ 3x ULN ≥ 5x ULN ≥ 5x ULN ≥ 5x ULN	- - - > 8x ULN > 10x ULN > 20x ULN	U/L
Alkaline Phosphatase	≥ 2x ULN ≥ 3x ULN ≥ 5x ULN	Pfizer standard flag for PCC	U/L
Creatinine	≥0.3 relative to baseline	≥0.4 relative to baseline	mg/dL
Gamma Glutamyl Transferase	≥ ULN	Pfizer standard flag for PCC	U/L
Total Bilirubin	> 1.5 ULN	Pfizer standard flag for PCC	mg/dL
Serum Total Bile Acids	> ULN	Pfizer standard flag for PCC	µmol/L
Direct and Indirect Bilirubin	> ULN	-	mg/dL

PCC – potential clinical concern
ULN – upper limit of normal as determined by the central laboratory

These endpoints will be derived using both pre- and post-dose data separately. Post-dose will include all post-baseline data including unplanned readings and pre-dose will include all data from the placebo run-in defined by including all values from Visit 2 to pre-dose, including the baseline measurement and unplanned readings. Note, both pre- and post-dose populations will be from the safety analysis set (defined in Section 4).

3.2.2.1. Vital Signs

Vital sign measurements (blood pressure and pulse rate) will be taken as detailed in the Schedule of Activities given in the protocol. The average of the triplicate measurements collected at each appropriate assessment time will be calculated for each vital sign parameter.

Baseline will be defined as the average of the triplicate measurements at the visit closest prior to dosing at Visit 3 (Day 1).

Changes from baseline for supine systolic and diastolic blood pressure and pulse rate will be calculated for each post baseline measurement.

3.2.2.2. Electrocardiogram (ECG)

Standard 12-lead ECG (including heart rate, QT, QTcF, PR and QRS interval) will be obtained at times detailed in the Schedule of Activities given in the protocol.

Baseline will be defined as the measurement at the visit closest prior to dosing at Visit 3 (Day 1).

Change from baseline for heart rate, QT, QTcF, PR and QRS interval will be calculated for each post baseline measurement.

3.2.2.3. Assessment of Mental Health as determined by Columbia-Suicide Severity Rating Scale (C-SSRS) and Patient Health Questionnaire-9 (PHQ-9).

The C-SSRS is a validated tool to evaluate suicidal ideation and behaviour. Data relevant to the assessment of suicidality will be mapped to the Columbia-Classification Algorithm of Suicide Assessment (C-CASA) codes as given in Appendix 5.

For this endpoint the screening visit will be labelled as 'Lifetime' in tables and the recent history (i.e. past 12 months) will also be reported separately. Baseline is defined as the last pre-dose measurement.

The PHQ-9 is a 9 item self-report scale for the assessment of depressive symptoms. The PHQ-9 will be completed by participants and reviewed by site staff at the pre-defined time points outlined in the schedule of activities.

The PHQ-9 total score will be derived for each time point separately by summing the responses to the 9 questions.

3.3. Other Safety Endpoint(s)

Not applicable.

3.4. Other Endpoint(s) (or, Exploratory Endpoints)

- Trough plasma concentrations of danuglipron at time points specified in the SoA.
- CFB in fasting plasma insulin, homeostatic model assessment of insulin resistance (HOMA-IR) and homeostatic model assessment of beta-cell function (HOMA-B) at:
    - Weeks 4, 8, 12, 16, 22 and 26 for Cohorts 1 and 2.
    - Weeks 4, 8, 16, 20, 24, 28 and 32 for Cohort 3.

HOMA-IR is calculated as: (fasting plasma insulin [FPI] × fasting plasma glucose [FPG])/405

HOMA-B is calculated as: (360 × FPI) / (FPG – 63)

In both cases above, FPI concentration is in mIU/L and the FPG concentration is in mg/dL.

- Absolute CFB in waist circumference at:
    - Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22 for Cohorts 1 and 2.
    - Weeks 4, 8, 12, 16, 20, 24 and 28 for Cohort 3.

The average of the triplicate waist measurements collected at each assessment time will be calculated prior to summaries/analysis. If any of the triplicates are missing, the non-missing value(s) will be used, and missing values will not be imputed.

- Absolute CFB in waist-to-hip ratio at:
    - Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22 for Cohorts 1 and 2.

- Weeks 4, 8, 12, 16, 20, 24 and 28 for Cohort 3.

The average of the triplicate waist and hip measurements collected at each assessment time will be calculated prior to calculating the waist-to-hip ratio at each assessment time. If any of the triplicates are missing, the non-missing value(s) will be used, and missing values will not be imputed.

- Shift from baseline in glycemic category (normoglycemia, pre-diabetes, or T2DM) at End of Treatment

For the purposes of reporting the shift in glycemic categories[2], the following definitions will be used:

Normoglycemia: FPG < 100 mg/dL and HbA1c < 5.7%;
Pre-diabetes: FPG 100-125 mg/dL (both inclusive) or HbA1c 5.7-6.4% (both inclusive);
T2DM: FPG ≥126 mg/dL or HbA1c ≥6.5%.

For discordant results between FPG and HbA1c with respect to the glycemic category criteria, the higher category will be used (eg, a participant with FPG of 124 mg/dL and HbA1c of 6.6% would be classified as T2DM).

- CFB in systolic and diastolic BP at each planned in-clinic study visit up through the End of Treatment visit.

- CFB in lipid parameters at:

  - Weeks 4, 8, 12, 16, 22 and 26 for Cohorts 1 and 2.

  - Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3.

- CFB in SF-36v2® subscales and summary scores at:

  - Weeks 16 and 26 at Weeks 16 and 26 for Cohorts 1 and 2.

  - Weeks 16, 24 and 32 for Cohort 3.

The SF-36v2®[3,4] is a 36-item questionnaire that measures functional health and well-being from a patient's perspective. It is a generic questionnaire and can be used across age (adults), disease, and treatment groups. The questionnaire consists of 8 health domain scales: physical functioning (10 items), role-physical (4 items), bodily pain (2 items), general health (5 items), vitality (4 items), social functioning (2 items), role-emotional (3 items), mental health (5 items), reported health transition (1 item) and 2, physical and mental component summary scores. Each health domain scale raw score is transformed to 0-100 scale which can then be converted to norm-based T-scores (Mean=50, SD=10).

The following transformed norm-based T-scores will be derived: physical functioning (PF); role-physical (RP); bodily pain (BP); general health (GH); vitality (VT); social functioning (SF); role-emotional (RE); mental health (MH); reported health transition; physical component summary (PCS); and mental component summary scores (MCS).

- CFB in Eating-Related Factors Daily Diary at:
    - Weeks 4, 8, 12, 16, 22 and 26 for Cohorts 1 and 2.
    - Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3.

The Eating-Related Factors Daily Diary is a daily, self-administered questionnaire that measures eating-related factors. The diary consists of 5 items that ask participants to evaluate their hunger, appetite, fullness, and cravings in the past 24 hours.

At baseline and at the timepoints listed above, the 7-day average score will be calculated separately for each item. This average score will be based on the 7 days leading up to the respective week (i.e. for Week 4 it would be Day 22 up to Day 28) and will be considered missing when 4 or more days are missing in a 7 day period. Baseline is defined as the 7 days leading up to dosing on Day 1 (i.e. Day -7 up to and including Day -1). The change from baseline for each post-dose time point will then be calculated for each of the 5 items separately.

- CFB in PGI-S at each planned in-clinic study visit up through the End of Treatment visit.

The PGI-S is a supportive item recommended by FDA for use as an anchor measure to generate an appropriate threshold that represents meaningful within-patient change in the target patient population. There are 4 PGI-S items that ask participants to evaluate the severity of their hunger, appetite, fullness, and cravings in the past 7 days.

The PGI-S should be converted to a numeric scale using the following mapping:

Response	Numeric value
None	0
Mild	1
Moderate	2
Severe	3

The change from baseline to each post-dose time point will then be calculated.

- CFB in PROMIS® Fatigue at:
    - Weeks 16 and 26 for Cohorts 1 and 2.
    - Weeks 16, 24 and 32 for Cohort 3.

The PROMIS® Fatigue[5] Custom 9-item Version is a self-reported measure that assesses a range of symptoms in the past 7 days from mild subjective feelings of tiredness to an overwhelming, debilitating, and sustained sense of exhaustion that likely decreases one's ability to execute daily activities and function normally in family or social roles. A global raw score ranging from 9 to 45 is calculated and can be translated into a T-score (Mean=50, SD=10) using the applicable score conversion table.

See Appendix 6 for derivation of the PROMIS Fatigue T-scores.

- CFB in PROMIS® Physical Function at:
    - Weeks 16 and 26 for Cohorts 1 and 2.
    - Weeks 16, 24 and 32 for Cohort 3.

The PROMIS® Physical Function[6] Custom 13-item Version is a self-reported measure that assesses capability rather than actual performance of physical activities. It includes assessment of one's ability to walk, climb stairs, run/jog, exercise as well as instrumental activities of daily living. A global raw score ranging from 13 to 65 is calculated and can be translated into a T-score (Mean=50, SD=10) using the applicable score conversion table.

See Appendix 7 for derivation of the PROMIS Physical Function T-scores.

- CFB in IWQOL-Lite-CT© at:
    - Weeks 16 and 26 for Cohorts 1 and 2.
    - Weeks 16, 24 and 32 for Cohort 3.

The IWQOL-Lite-CT©[7] was developed for use in the context of clinical trials for obesity, with 20 items addressing concerns that are specifically relevant to the study population.

The responses to the 20 items will be summarised at each visit using the following 3 scores: physical composite score; physical function score; and psychosocial composite score. These are derived as outlined in Appendix 8. The change from baseline for each post-dose time point (as applicable) will then be calculated for each of the 3 scores.

- PGI-C at:
    - Weeks 4, 16 and 26 for Cohorts 1 and 2.
    - Weeks 4, 16, 24 and 32 for Cohort 3.

The PGI-C is a supportive item recommended by FDA for use as an anchor measure to generate an appropriate threshold that represents meaningful within-patient change in the target patient population. There are 4 PGI-C items that ask participants to rate the overall change in their hunger, appetite, fullness, and cravings since they started taking the study medication.

3.5. Baseline Variables

Baseline measures will be included as a covariate in all applicable statistical models along with the stratification variable of biological gender (female versus male) at screening.

The statistician may conduct further exploratory analyses into the effect of covariates and factors (such as age and strata) on the efficacy endpoints. If conducted, and considered relevant to the clinical study report (CSR), the methods will be fully justified and discussed within the report.

4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS)

Data for all participants will be assessed to determine if participants meet the criteria for inclusion in each analysis population prior to unblinding and releasing the database and classifications will be documented per standard operating procedures.

Participant Analysis Set	Description
Enrolled	*"Enrolled" means a participant's agreement to participate in a clinical study following completion of the informed consent process. Potential participants who are screened for the purpose of determining eligibility for the study, but do participate in the study, are not considered enrolled, unless otherwise specified by the protocol.*
Randomly assigned to study intervention	*All participants randomly assigned to study intervention regardless of whether or not study intervention was administered.*
Evaluable	*All participants randomly assigned to study intervention and who take at least 1 dose of study intervention. Participants will be analyzed according to the randomized study intervention.*
Safety	*All participants randomly assigned to study intervention and who take at least 1 dose of study intervention. Participants will be analyzed according to the study intervention they actually received.*

Defined Analysis Set	Description
Estimand Set 1	*All participants randomly assigned to study intervention and who take at least 1 dose of study intervention.For participants who discontinue study intervention, all subsequent values will be censored.*
Estimand Set 2	*All participants randomly assigned to study intervention and who take at least 1 dose of study intervention and do not have major protocol deviations. Major protocol deviations will be reviewed prior to database lock to identify such participants* (described in Section 4.1). *For participants who discontinue study intervention, all subsequent values will be censored.*
PK Concentration Set	*All participants randomly assigned to study intervention and who take at least 1 dose of PF-06882961* (danuglipron) *and in whom at least 1 concentration value is reported.*

4.1. Identifying Major Protocol Deviations

Major protocol deviations will be identified by review of each protocol deviation on a case-by-case basis to determine whether the associated subject should be excluded from the Estimand Set 2. The list of potential deviations that could be classified as major are: falsely enrolled into study despite violating inclusion/exclusion criteria; non-compliance with study medication; and use of prohibited medications. Other deviations that are identified during the study based on emerging data that are deemed to qualify as major may also be considered. The project statistician will ensure that the list of subjects and the reason for exclusion is clearly documented and provided to programming prior to unblinding and releasing the database.

4.2. Strata Misallocations

Participants who are randomized to the wrong stratum (male/female), in error, will have the incorrect stratum assignment remain in IMPALA but the clinical database will include the correct stratum. The latter will subsequently be used for all relevant analyses.

5. GENERAL METHODOLOGY AND CONVENTIONS

5.1. Hypotheses and Decision Rules

The null hypothesis that there is no difference between danuglipron and placebo will be tested for the primary endpoint and efficacy related secondary and tertiary endpoints. The alternative hypothesis is that danuglipron is superior (i.e. greater reduction) to placebo on the primary endpoint.

Statistical inference will be based on the primary endpoint: percent change from baseline in body weight at End of Treatment.

The Type I error rate (α-level) used for the statistical inference will be 5% (1-sided).

Each treatment group of danuglipron will be tested separately compared to placebo. The slow titration arms will be considered separate treatment groups, to the standard titration arms.

No adjustment for multiple comparisons will be made.

For all other endpoints there will be no formal hypothesis testing.

5.2. General Methods

The analyses related to the primary, secondary and exploratory endpoints will be based on the appropriate population for analysis (see Section 4).

Unless otherwise stated, all summaries, analyses and plots will be presented by treatment group with separate tables for Cohorts 1 & 2 combined and Cohort 3. The following treatment group labels (or similar) will be used:

Cohorts 1 & 2

Placebo
PF-06882961 40mg BID (1-week titration)

PF-06882961 80mg BID (1-week titration)
PF-06882961 120mg BID (1-week titration)
PF-06882961 160mg BID (1-week titration)
PF-06882961 200mg BID (1-week titration)
PF-06882961 120mg BID (2-week titration)
PF-06882961 160mg BID (2-week titration)
PF-06882961 200mg BID (2-week titration)

Cohort 3

Placebo
PF-06882961 80mg BID (4-week titration)
PF-06882961 140mg BID (4-week titration)
PF-06882961 200mg BID (4-week titration)

5.2.1. Analyses for Continuous Endpoints

Continuous variables will be presented using summary statistics: number of observations, arithmetic mean, standard deviation, median and range (minimum and maximum) values. For endpoints to be analysed on the natural $\log_e$ scale, the geometric mean and geometric coefficient of variation will additionally be calculated (excluding change from baseline summaries).

5.2.2. Analyses for Categorical Endpoints

Categorical variables will be presented using summary statistics: number of observations, counts and percentages.

5.2.3. Mixed Model Repeated Measures (MMRM)

The MMRM model(s) will include treatment, time, strata (females versus males) and treatment-by-time interaction as fixed effects, baseline as a covariate and the baseline-by-time interaction with time fitted as a repeated effect and participant as a random effect.

An unstructured covariance matrix will be used to estimate the variances and covariance within participant across time points. If convergence is not obtained or model fit is not adequate, then other covariance structures will be investigated as necessary. The Kenward-Roger approximation will be used for estimating degrees of freedom for the model parameters.

Missing values (e.g. due to censoring) will be implicitly imputed as part of the MMRM model fitting.

The Least Squares Means (LSMeans) together with 90% confidence intervals, standard errors and p-values (2-sided) will be obtained for each treatment group at each time point.

Differences in LSMeans between each treatment group of danuglipron relative to placebo at each time point, together with 90% confidence intervals, standard errors and p-values (2-sided), will also be obtained.

Example SAS code is provided in Appendix 3.

5.2.4. Logistic Regression

The logistic regression model(s) will include a term for treatment, strata (females versus males), and baseline will be included as a covariate.

Missing values (e.g. due to censoring) will be imputed for missing data using a multiple imputation method as described in Section 5.3.

The odds ratio for each treatment group of danuglipron relative to placebo and corresponding 90% confidence interval will be obtained.

Example SAS code is provided in Appendix 3.

5.2.5. Emax Model

The 4-parameter dose-response Emax model will be used to characterize the percent change from baseline (%CFB) dose-response relationships with dose included as a continuous variable.

The model structure will take the form:

$$log_e CFB = log_e(E_0) + \frac{log_e(E_{max}) \times dose^{Hill}}{ED_{50}{}^{Hill} + dose^{Hill}}$$

E_0 is the placebo effect, *dose* is the target randomized dose, E_{max} is the maximum effect, ED_{50} is the dose producing 50% of the maximum effect and *Hill* is the slope parameter.

The model will be applied to the raw LSMean results (i.e. on the log_e-scale) from the primary MMRM model (Section 5.2.3) utilizing a Bayesian methodology approach with weakly informative priors as described in Appendix 2. Only the raw LSMean results from the placebo (combined Cohorts 1 & 2 only) and 1-week titration treatment groups (from Cohort 1) will be included in the model, i.e. the results from the 2-week and 4-week titration treatment groups and placebo from Cohort 3 will not be included.

Estimates of the model parameters of E_0, E_{max}, ED_{50} and *Hill* and their 95% credible intervals will be produced.

The posterior medians and 90% credible intervals (5^{th} and 95^{th} percentiles of the relevant posterior distribution) will be reported for each target randomized dose (including Placebo) and their differences relative to placebo. These will all be presented on the %scale by back transforming modelled results. Both will be reported in tables and plotted in separate figures.

If convergence cannot be obtained or visual inspection of the data does not support a dose-response Emax relationship the following options will be considered in order: (1) assume the hill parameter is 1 and remove from the model (giving a 3-parameter dose-response Emax model); (2) utilize alternative priors to those provided in the appendix 2; (3) otherwise this Emax model will not be reported and alternative model structures may be investigated and the subsequent analyses may be included in the CSR with rationale for the eventual model selected.

5.2.6. Emax Model (baseline interaction)

An additional 4-parameter dose-response Emax model will be used to characterize percent change from baseline dose-response relationships, assuming an interaction between baseline body weight and dose.

The model structure will be the same as Section 5.2.5.

The model will be applied to the raw LSMean results (i.e: on the $\log_e$-scale) from individual MMRM models that have been applied to each treatment group separately. The MMRM models will include time, strata, baseline (on the $\log_e$-scale) as a covariate and the baseline-by-time interaction with time fitted as a repeated effect and participant as a random effect. Only the raw LSMean results from placebo (combined Cohorts 1 & 2 only) and 1-week titration treatment groups (from Cohort 1) will be included in the model, i.e. the results from the 2-week and 4-week titration treatment groups and placebo from Cohort 3 will not be included.

An unstructured covariance matrix will be used to estimate the variances and covariance within participant across time points for each separate MMRM. If convergence is not obtained or model fit is not adequate, then other covariance structures will be investigated as necessary. The Kenward-Roger approximation will be used for estimating degrees of freedom for the model parameters.

The raw LSMean results from the separate MMRM models (placebo and standard titration treatment groups only, as above) will be calculated for the average baseline body weight (i.e. across all participants in the Estimand Set 1A) and then fitted with the Emax model above utilizing a Bayesian methodology approach with weakly informative priors as described in Appendix 2

Estimates of the model parameters of E_0, E_{max}, ED_{50} and *Hill* and their 95% credible intervals will be produced.

The posterior medians and 90% credible intervals (5^{th} and 95^{th} percentiles of the relevant posterior distribution) will be reported for each target randomized dose (including Placebo) and their differences relative to placebo. These will all be presented on the %scale by back transforming modelled results. Both will be reported in tables and plotted in separate figures.

If convergence cannot be obtained or visual inspection of the data does not support a dose-response Emax relationship the following options will be considered in order: (1) assume the hill parameter is 1 and remove from the model (giving a 3-parameter dose-response Emax model); (2) utilize alternative priors to those provided in the appendix 2; (3) otherwise an Emax model (with a baseline interaction) will not be reported.

5.2.7. Cumulative Incidence Plots

Cumulative Incidence Plots will be produced based on the time to the event of interest (starting from the time of start of dosing on Day 1) for each treatment group separately and will be plotted on the same graph. This will be based on plotting the cumulative incidence function (with no competing risks), which will be presented as a % on the y-axis. No statistical testing for differences between treatment groups will be considered.

Details of censoring are included in Section 6 and example SAS code is provided in Appendix 3.

5.3. Methods to Manage Missing Data

Details of efficacy data to be censored is described in Section 3.2.

For applicable continuous endpoints modelled with an MMRM, missing/censored values will be imputed as part of the analysis method.

For the analysis of safety endpoints, the sponsor data standard rules for imputation of BLQ values will be applied.

In all PK data presentations (except listings), concentrations below the limit of quantification (BLQ) will be set to zero. In listings BLQ values will be reported as "<LLQ", where LLQ will be replaced with the value for the lower limit of quantification (LLQ).

Where a dose has been administered on site and actual dose, frequency, planned dose, planned dose frequency, end date time and start date time are missing, these missing values will be imputed with the details of the dose planned as per the protocol.

5.3.1. Multiple Imputation

For summarizing the proportion of responses in body weight, all data from Estimand Set 1 will be included (i.e. all time points up to and including End of Treatment). A multiple imputation (MI) method will be implemented, using a multivariate imputation method by chained equations, which is still valid with an arbitrary missing data pattern. The model will include: treatment, baseline body weight and strata. Twenty sets of imputations of each missing value will be constructed from the MI method and the proportion of responses by treatment group will be determined with associated standard errors utilizing a normal approximation and will be combined using standard multiple imputation techniques proposed by Rubin[8] to yield overall estimates. If there is more missing data in the study than anticipated, the number of imputation sets may be increased as required.

For logistic regression of body weight, all data from Estimand Set 1 will be included (i.e. all time points up to and including End of Treatment). The same imputed datasets as produced for the proportion of responses above will be utilized, where a Logistic Regression model (as described in Section 5.2.4) will be applied to each of the 20 imputed datasets separately. Parameter estimates of the log odds ratios for each treatment group relative to placebo will be combined using standard multiple imputation techniques proposed by Rubin[8] to yield overall estimates of the log odds ratios and their associated standard errors will be used to create 90% confidence intervals on the log-odds scale. The log odds ratios and log odds 90% confidence intervals will be back transformed into odds ratios and associated 90% confidence intervals for final reporting.

6. ANALYSES AND SUMMARIES

Data collected before baseline will only be listed, unless otherwise stated.

For all endpoints, baseline is defined in Section 3.

For Cohorts 1 and 2, End of Treatment is defined as Week 26. For Cohort 3, End of Treatment is defined as Week 32.

Data from participants enrolled into the original cohort (Cohort 1) and Cohort 2 will be combined for reporting, where data from the placebo and 2-week titration arms will be combined from the 2 cohorts and analyzed together. Data from participants enrolled into Cohort 3 will be analyzed and reported separately unless otherwise specified.

6.1. Primary Endpoint(s)

6.1.1. Percent Change from Baseline in Body Weight at End of Treatment

6.1.1.1. Main Analysis

In all cases the Estimand Set 1, as specified in Section 2.2.1 will be utilised.

Absolute changes from baseline and percent changes from baseline in body weight will be summarised descriptively by treatment group and time point as described in Section 5.2.1, which will be presented separately for the combined Cohorts 1 & 2 and Cohort 3. Tables will present all data from the screening (visit 1, absolute tables only), beginning of the placebo run-un (visit 2, absolute tables only), baseline and post-baseline time points (including follow-up, which will be restricted to participants who completed treatment up to and including the End of Treatment visit).

The primary analysis will be an MMRM (as described in Section 5.2.3) applied to the changes from baseline in body weight on the natural log scale at Weeks 2, 4, 6, 8, 10, 12, 16, 18, 22 and 26 for Cohorts 1 & 2, and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3, that will be used to estimate the treatment effect related to the primary Estimand 1 (as described in Section 2.2.1). The MMRM results will be back-transformed as described in Section 3.1 for presentations.

The following results from the above primary analysis will be plotted separately for the combined Cohorts 1 & 2 and Cohort 3 and for Cohorts 1, 2 and 3 combined. The plots for Cohort 1 & 2 and Cohort 3 separately will not be included in the CSR.

- Profile plots of the back-transformed LSMeans (including 90% confidence intervals) representing percent change from baseline over time, with a separate line for each treatment group
- Profile plots of the back-transformed LSMean differences to Placebo (including 90% confidence intervals), representing percent change to Placebo over time, with a separate line for each treatment group of danuglipron
- Plot of the back-transformed LSMeans (including 90% confidence intervals) representing percent change from baseline at End of Treatment versus treatment group on the x-axis (applicable to analyses related to body weight only)
- Plot of the back-transformed LSMean differences to Placebo (including 90% confidence intervals) representing percent change to Placebo at End of Treatment, with a separate point for each treatment group of danuglipron (applicable to analyses related to body weight only)

Standard SAS output will be provided to support the main statistical summary table for the primary analysis model but will not be included in the CSR.

Statistical Model Diagnostics

The presence of outliers will be investigated for this analysis. An outlier will be defined as any response data value with a studentized (conditional) residual greater than 3, or less than -3. A listing will be presented of any participants meeting these criteria and will be included with standard SAS outputs. The assumptions of normality will be verified graphically using residual plots. For each fitted model, a set of conditional studentized residual plots will be produced, including residual plot, histogram of normality, QQ plot and summary of fit statistics. The residual plots will not be included in the clinical study report.

If there are outliers or major deviations from normality, then the effect of these on the conclusions may be investigated through alternative transformations and/or analyses excluding outliers. Justification for any alternative to the planned analysis will be given in the report of the study.

6.1.1.2. Sensitivity/Supplementary Analyses

For all of the following analyses, standard SAS output will be provided to support the statistical summaries produced but will not be included in the CSR.

The following supplementary analyses to the primary endpoint will be carried out:

- To assume a dose-response relationship with body weight (without a baseline interaction) an Emax model will be applied to the raw LSMeans at End of Treatment from the main MMRM analysis as described in Section 5.2.5. Final results from the Emax model will be back transformed using the same formula as provided in 3.1. This will only be applied in reference to percent change from baseline in body weight for Cohorts 1 & 2.

- To assume a dose-response relationship with body weight, including an interaction between baseline body weight and dose, an Emax model will be applied to the raw LSMeans at End of Treatment from separate MMRM models at the overall baseline body weight as described in Section 5.2.6. Final results from the Emax model will be back transformed using the same formula as provided in 3.1. This will only be applied in reference to percent change from baseline in body weight for Cohorts 1 & 2.

- To model absolute changes from baseline in body weight, a separate MMRM model (as described in Section 5.2.3) will be applied to the absolute change from baseline in body weight at Weeks 2, 4, 6, 8, 10, 12, 16, 18, 22 and 26 for Cohorts 1 & 2, and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3 separately, where baseline absolute body weight will be the baseline term. Similar outputs to Section 6.1.1.1 will be reported from the MMRM model where back-transformation of LSMeans and LSMean differences will not be applicable, along with standard SAS output (which will not be included in the CSR).

The additional exploratory estimand to the primary endpoint will be carried out:

An additional exploratory analysis that assesses the primary endpoint in evaluable participants with no major protocol deviations while on treatment (Section 2.2.3) will be performed to estimate the treatment effect related to the exploratory Estimand 7 (as described in Section 2.2.3). The same summary and analysis output as the main analysis (Section 6.1.1.1) will be performed and reported but applied to the Estimand Set 2 (as described in Section 4).

6.2. Secondary Endpoint(s)

6.2.1. Response as defined by a body weight loss of ≥5% from baseline at End of Treatment

In all cases the Estimand Set 1 as specified in Section 4 will be utilised.

The proportion of responses at End of Treatment will be summarised descriptively by treatment group as described in Section 5.2.2, where no imputation for missing values will be conducted. The proportion of responses at End of Treatment will also be reported after multiple imputation for missing values as per Section 5.3.

A logistic regression model (as described in Section 5.2.4) will be applied to the response at End of Treatment that will be used to estimate the treatment effect related to the secondary Estimand 2 (as described in Section 2.2.2.1), where missing values will be imputed using multiple imputation as per Section 5.3. Baseline body weight and strata will be included in the model.

In addition, standard SAS output will be provided to support the main statistical summary table for the logistic regression model and multiple imputation approach but will not be included in the CSR.

6.2.2. Percent CFB in Body Weight over time

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Percent CFB in body weight will be summarized at Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22 for Cohorts 1 & 2 and for Weeks 4, 8, 12, 16, 20, 24 and 28 for Cohort 3. Summaries and analysis results will be reported as part of the main analysis approach for the primary endpoint (Section 6.1.1.1).

6.2.3. Absolute CFB in waist circumference at End of Treatment

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute changes from baseline in waist circumference will be summarised and analysed similar to body weight in Section 6.1.1.1, including follow-up.

An MMRM (as described in Section 5.2.3) will be applied to the absolute change from baseline at Weeks 2, 4, 6, 8, 10, 12, 16, 18, 22 and 26 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3 separately, where baseline waist circumference will be the baseline term. Results similar to body weight in Section 6.1.1.1, where back-transformation of LSMeans and LSMean differences will not be applicable, will be reported along with standard SAS output (where the figures and standard SAS output will not be included in the CSR).

The MMRM results from the End of Treatment visit will be extracted in order to estimate the treatment effect related to Estimand 3.

6.2.4. Absolute CFB in waist-to-hip ratio at End of Treatment

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute changes from baseline in waist-to-hip ratio will be summarised and analysed similar to body weight in Section 6.1.1.1, including follow-up.

An MMRM (as described in Section 5.2.3) will be applied to the absolute change from baseline at Weeks 2, 4, 6, 8, 10, 12, 16, 18, 22 and 26 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3 separately, where baseline waist-to-hip ratio will be the baseline term. Results similar to body weight in Section 6.1.1.1, where back-transformation of LSMeans and LSMean differences will not be applicable, will be reported along with standard SAS output (where figures and standard SAS output will not be included in the CSR).

The MMRM results from the End of Treatment visit will be extracted in order to estimate the treatment effect related to Estimand 4.

6.2.5. Absolute CFB in HbA1c

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute changes from baseline in HbA1c will be summarised and analysed similar to body weight (Weeks 16 and 26 for Cohorts 1 & 2 and Weeks 16, 24 and 32 for Cohort 3) in Section 6.1.1.1, including follow-up.

An MMRM (as described in Section 5.2.3) will be applied to the absolute change from baseline at the time points specified above where baseline HbA1c will be the baseline term. Results similar to body weight in Section 6.1.1.1, where back-transformation of LSMeans and LSMean differences will not be applicable, will be reported along with standard SAS output (where figures and standard SAS output will not be included in the CSR).

The MMRM results will be extracted in order to estimate the treatment effect related to Estimand 5.

6.2.6. Absolute CFB in FPG over time

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute changes from baseline in FPG will be summarised and analysed similar to body weight in Section 6.1.1.1, including follow-up.

An MMRM (as described in Section 5.2.3) will be applied to the absolute change from baseline at Weeks 2, 4, 6, 8, 10, 12, 16, 18, 22 and 26 for Cohorts 1 & 2 and separately, at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3, where baseline FPG will be the baseline term. Results similar to body weight in Section 6.1.1.1, where back-transformation of LSMeans and LSMean differences will not be applicable, will be reported along with standard SAS output (where figures and standard SAS output will not be included in the CSR).

The MMRM results will be extracted in order to estimate the treatment effect related to Estimand 6.

6.3. Other Safety Summaries and Analyses Endpoint(s)

The AEs of interest are defined in Section 3.2.2.1.1. Figures described in this section may not be included in the CSR.

6.3.1. Adverse Events

Adverse events (Tier 1, 2 and 3 AEs as described in Section 3.2.2.1) will be summarized by treatment group and overall and in accordance with sponsor reporting standards using the safety analysis set defined in Section 4. The adverse events (AEs) will be sorted in descending frequency within a system organ class.

Incidence and severity of treatment emergent adverse event (TEAE) tables will additionally be produced ('All causality' and 'Treatment related', separately) to summarise the total number of adverse events by preferred term, which will be reported by treatment group and overall.

The following tables and figures related to TEAEs classed as Tier 1 or 2 will be ordered in descending point estimate of risk difference within System Organ Class. If two or more events have the same frequency, they will be sorted alphabetically by preferred term.

TEAEs classed as Tier 1 events will be tabulated by treatment group. Number of participants and percent will be presented, along with the risk difference between each treatment group of danuglipron and placebo. 95% confidence intervals and p-values will also be presented for the comparison. No adjustment for multiplicity will be used.

Tier 1 TEAEs will also be presented graphically. The TEAEs will be presented in a two-panel plot, the left panel will give the proportions of TEAEs observed in a treatment group of danuglipron and separately placebo while the right panel will display the 95% confidence interval for the risk differences for each TEAE. A vertical line corresponding to the value of 0 will be added to the right-hand plot. For the graphical display for tier 1 events, a column containing the p-value for each event will be added to the right-hand side of the forest plot in the right panel. Each panel will be paged by treatment group of danuglipron.

TEAEs classed as Tier 2 events will be summarised and graphically presented similar to Tier 1 events, but no p-values will be presented. No adjustments for multiplicity will also be used.

For Tier 1 and Tier 2 outputs, footnotes will be included on the tables to provide proper interpretation of p-values (Tier 1 only) and confidence intervals and to describe how the comparison was conducted, e.g. "P-values and confidence intervals are not adjusted for multiplicity and should be used for screening purpose only. 95% Confidence intervals are provided to help gauge the precision of the estimates for Risk Difference. Risk Difference is computed as danuglipron versus placebo."

The incidence of TEAEs by preferred term with more than 5% occurrences by preferred term will be presented. The TEAEs to include in tables will be based on those occurring in more than 5% of participants in either (1) Cohorts 1 and 2 combined; or (2) more than 5% of participants in Cohort 3.

TEAEs with more than 5% occurrences in any treatment arm will additionally be presented, split by the titration step at the time of TEAE onset.

It should be recognized that most studies are not designed to reliably demonstrate a causal relationship between the use of a pharmaceutical product and an adverse event or a group of adverse events. Except for select events in unique situations, studies do not employ formal adjudication procedures for the purpose of event classification. As such, safety analysis is generally considered as an exploratory analysis and its purpose is to generate hypotheses for further investigation. The 3-tier approach facilitates this exploratory analysis.

6.3.1.1. Forest plots on the Incidence of adverse events of interest

Forest plots displaying the incidence of TEAEs of interest ("All Causality" only) by treatment group will be produced separately by AE. Each forest plot will present all treatment groups separately and will include 95% confidence intervals (calculated using the Blyth-Still-Casella method[9]).

6.3.1.2. Incidence of first occurrence of adverse events of interest by week

The number and percentage of participants who experience the first occurrence of a TEAE of interest each week (up to and including the follow-up visit) will be tabulated by treatment group and week, where percentage will be defined in two ways. These will be included in the same table and produced separately for each TEAE of interest.

To calculate the percentage each week, the numerator will be the total number of participants with the TEAE of interest that was also their first post-dose occurrence of the event during that respective week. The two approaches for the denominator are: (1) the total overall number of evaluable participants in the Safety Analysis Set for the respective treatment group; and (2) the total number of participants who had not discontinued from the study prior to that respective week will be the denominator (note if a participant did discontinue from the study during that respective week they would be included in the denominator).

A figure of the percentage per week will also be produced, with week on the x-axis and a separate line for each treatment group. The percentage will represent the above definition (2) of the denominator but based on blinded data reviews this figure may be adapted or additionally produced based on definition (1). These figures will be produced separately for each TEAE of interest.

6.3.1.3. Prevalence of adverse events of interest by week

The number and percentage of participants who are experiencing a TEAE of interest (either onset or ongoing) each week (up to and including the follow-up visit, with week as appropriate based on the End of Treatment definition) will be tabulated by treatment group and week. A separate table representing the number and percentage by severity (mild, moderate or severe) will also be produced. These tables will be produced separately for each TEAE of interest.

To calculate the percentage each week, the total number of participants who experience the TEAE of interest at any time during the respective week will be the numerator and the total number of participants who had not discontinued from IP and/or the study prior to that

respective week will be the denominator (note if a participant did discontinue from IP and/or the study during that respective week they would be included in the denominator).

A figure of the overall percentage (i.e. not by severity) per week will be produced, with week on the x-axis and a separate line for each treatment group. A separate figure of the percentage per week will also be produced, with week on the x-axis and a separate line for each severity, paged by treatment group. These figures will be produced separately for each TEAE of interest.

6.3.1.4. Time to First Occurrence/Recurrence of AEs of Interest

Exploratory summaries on the time to the first occurrence of AEs of interest will be produced in Cumulative Incidence Plots as described in Section 5.2.7. Participants who discontinue from the study or discontinue from IP prior to the start of the AE event of interest will be censored at the discontinuation/initiation date.

A separate plot for each AE of interest will be produced separately.

The above will also be produced separately for the time to the first recurrence of the AEs of interest.

6.3.1.5. Time to discontinuation from IP due to Gastrointestinal Disorders AEs

Exploratory summaries on the time to discontinuation from IP (regardless of study discontinuation or continuation) due to Gastrointestinal Disorders AEs (defined as based on System Organ Class) will be produced with a Cumulative Incidence Plot as described in Section 5.2.7. Participants who discontinue will be censored at the associated discontinuation date.

6.3.1.6. Subset Reporting Interval

The subset adverse events, described in Section 3.2.2.1.2, will be summarised as above (excluding the 3-tier reporting system) where standard tables and incidence tables will be produced for both ‘All causality’ and ‘Treatment related’.

6.3.1.7. Hypoglycemic Adverse Events

The hypoglycemic AEs will be listed in a separate table and summarized categorically by treatment group and overall as per Section 5.2.1.

6.3.2. Laboratory Data

Laboratory data from will be listed and summarized by treatment group and overall, in accordance with the sponsor reporting standards.

6.3.2.1. Focused Laboratory Summaries on Endpoints of Interest

Absolute values, changes from baseline and percent changes from baseline in calcitonin, amylase, lipase, TSH, free T4, lipid profile, liver function tests and eGFR (as outlined in Section 3.2.2.3) will be summarized by treatment group and time point as per Section 5.2.1 (unplanned readings will not be included in these summaries). Follow-up data will be including in the summaries with data from all participants in the safety analysis set. Tables will be paged by parameter.

The following box and whisker plots for each parameter will be produced:

- Absolute values over time by treatment group
- Change from baseline over time by treatment group
- Percentage change from baseline over time by treatment group.

Spaghetti plots for individual data versus time, by treatment, will be produced for ALT, ALP, AST and total bilirubin:

Plots of absolute values of ALT and AST versus change from baseline in weight by treatment will also be produced.

MMRM models (as described in Section 5.2.3) will be applied to the change from baseline at Weeks 4, 8, 12, 16, 22 and 26 (as applicable) for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3 separately and for each parameter separately using the safety analysis set (as defined in Section 4). For ALT, AST, alkaline phosphatase and GGT, the MMRM models will instead be applied to the change from baseline on the natural log scale and the MMRM results will be back-transformed for reporting, similar to the primary endpoint as described in Section 6.1.1.1.

From each model, the LSMeans together with 90% confidence intervals, standard errors and p-values will be obtained for each treatment group at each time point. Differences in LSMeans between each treatment group of danuglipron relative to placebo at each time point, together with 90% confidence intervals, standard errors and 2-sided p-values will also be obtained. A plot of the LSMean differences to Placebo (including 90% confidence intervals) over time, with a separate line for each treatment group of danuglipron will be produced separately for each parameter. Standard SAS outputs will be provided to support the main statistical summary tables but will not be included in the CSR.

6.3.2.2. Clinical Laboratory Parameters of Interest

An additional summary table of the number of participants (from the Safety Analysis Set as defined in Section 4) with “Flag Level” or “Alert Level” abnormalities for each of the clinical laboratory parameters of interest as specified in Section 3.2.2.4 will be produced. This table will summarise the number of participants with “Flag level” or “Alert level” abnormalities separately and by treatment group and placebo run-in and overall as per Section 5.2.2.

6.3.3. Vital Signs

Average of the triplicate measurements (where applicable) will be used in analyses.

Absolute values and changes from baseline in supine systolic and diastolic blood pressure and pulse rate will be summarized by treatment group and time point, according to sponsor reporting standards. Tables will be paged by parameter.

Mean changes from baseline for systolic and diastolic blood pressure and pulse rate will be plotted against time point. On each plot there will be 1 line for each treatment group with all groups on the same plot including placebo.

MMRM models (as described in Section 5.2.3) will be applied to the change from baseline for supine systolic and diastolic blood pressure and pulse rate separately using the safety analysis set (as defined in Section 4). For Cohorts 1 & 2, Weeks 0 (i.e. 2-6H post-dose only), 2, 4, 6, 8 (0 and 2-6H separately), 10, 12, 16 (0 and 2-6H separately), 18, 22 and 26 (0 and 2-6H separately) assessments will be included in the model. For Cohort 3, Weeks 0 (0 and 2-6H separately), 4, 8 (0 and 2-6H separately), 12, 16 (0 and 2-6H separately), 20, 24 (0 and 2-6H separately), 28 and 32 (0 and 2-6H separately) assessments will be included. From each model, the LSMeans together with 90% confidence intervals, standard errors and p-values will be obtained for each treatment group at each time point. Differences in LSMeans between each treatment group of danuglipron relative to placebo at each time point, together with 90% confidence intervals, standard errors and 2-sided p-values will also be obtained. A plot of the LSMean differences to Placebo (including 90% confidence intervals) over time, with a separate line for each treatment group of danuglipron will be produced separately for each parameter. Standard SAS outputs will be provided to support the main statistical summary tables but will not be included in the CSR.

Maximum absolute values and maximum changes from baseline for vital signs, over all measurements taken post-dose will also be tabulated by treatment group and overall using categories as defined in the Appendix 4. Numbers and percentages of participants meeting the categorical criteria will be provided. All planned and unplanned post-dose time points will be counted in these categorical summaries. All values meeting the criteria of potential clinical concern will be listed.

Maximum decrease and increase from baseline for supine systolic and diastolic blood pressures, and maximum increase from baseline for supine pulse rate will be summarized by treatment group, according to sponsor reporting standards.

6.3.4. Electrocardiograms

Average of the triplicate measurements (where applicable) will be used in analyses. Absolute values and changes from baseline in QT, heart rate, QTcF, PR and QRS will be summarized by treatment group and time point using sponsor reporting standards. Tables will be paged by parameter.

Mean changes from baseline in QT, heart rate and QTcF will be plotted against time point. On each plot there will be 1 line for each treatment group with all treatments included on the same plot including placebo.

MMRM models (as described in Section 5.2.3) will be applied to the change from baseline at Weeks 2, 4, 6, 8, 10, 12, 16, 18, 22 and 26 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3 separately, for QT, heart rate and QTcF separately using the safety analysis set (as defined in Section 4). From each model, the LSMeans together with 90% confidence intervals, standard errors and p-values will be obtained for each treatment group at each time point. Differences in LSMeans between each treatment group of danuglipron relative to placebo at each time point, together with 90% confidence intervals, standard errors and 2-sided p-values will also be obtained. A plot of the LSMean differences to Placebo (including 90% confidence intervals) over time, with a separate line for each treatment group of danuglipron will be produced separately for each parameter. Standard SAS outputs will be provided to support the main statistical summary tables but will not be included in the CSR.

ECG endpoints and changes from baseline (QTcF, PR and QRS) will also be summarized descriptively by treatment group and overall using categories as defined in the Appendix 4 (for QTc these correspond to the Pfizer Guidance[10]). Numbers and percentages of participants meeting the categorical criteria will be provided. All planned and unplanned post-dose time points will be counted in these categorical summaries. All values meeting the criteria of potential clinical concern will be listed.

Maximum absolute value (post-dose) and the maximum increase from baseline for QTcF, PR and QRS will be summarized by treatment group according to sponsor reporting standards.

Listings of participants with any single post-dose value > 500 msec will also be produced for QTcF.

6.3.5. Assessment of mental health as determined by C-SSRS and PHQ-9.

Screening, baseline and post-baseline C-SSRS data (mapped to C-CASA scores as described in Section3.2.2.3) using the safety population defined in Section 4, will be separately summarized categorically by treatment group and time point as outlined in Section 5.2.2.

Baseline and post-baseline PHQ-9 data (responses to each of the 9 items) using the safety population defined in Section 4 will be summarized categorically for each question separately by treatment group and time point as outlined in Section 5.2.2. The PHQ-9 total score as defined in Section 3.2.4 will additionally be summarized descriptively by treatment group and time point as outlined in Section 5.2.1.

The number of participants who met the criteria for referral to a mental health professional will be listed and summarized by treatment group and time point as outlined in Section 5.2.2.

6.4. Other Endpoint(s) (or, Exploratory Endpoint[s])

6.4.1. Trough plasma concentrations of danuglipron at time points specified in the SoA

Danuglipron concentrations will be characterized by C_{trough} and will be summarized by treatment group and by time point using the following descriptive statistics: number of participants contributing at each time point, arithmetic mean, median, minimum, maximum, Q1, Q3, standard deviation, geometric mean, and geometric CV (%).

Median C_{trough} versus time point will be plotted by treatment group including error bars representing the inter-quartile range (i.e. Q1 to Q3).

C_{trough} values from PK samples that were collected after discontinuation of investigational product will be listed but excluded from summarization.

The post-dose random plasma concentration data will only be listed (and not summarized); these data are noted for use in supplemental population-PK analyses.

In addition, as permitted by data and determined by the sponsor, PK-PD relationship between plasma concentrations of danuglipron and effect on primary, secondary and tertiary endpoints may be characterized using a population PK-PD approach. The objective of such

an analysis, if conducted, would aim to explore potential demographic determinants (eg, age, gender, and weight) influencing the observed PK or PD variability in response to danuglipron. The population PK-PD analysis, if conducted, will be reported separately from the main CSR.

6.4.2. CFB in fasting plasma insulin, HOMA-IR and HOMA-B at Weeks 4, 8, 12, 16, 22 and 26

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute changes from baseline in fasting plasma insulin, HOMA-IR and HOMA-B will be summarised and analysed separately similar to body weight in Section 6.1.1.1, including follow-up, where HOMA-IR and HOMA-B are as defined in Section 3.2.2.

MMRM models (as described in Section 5.2.3) will be applied to each parameter separately, where each will include the absolute change from baseline at Weeks 4, 8, 12, 16, 22 and 26 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3 separately, and the baseline term will be the baseline for the associated parameter. Results similar to body weight in Section 6.1.1.1, where back-transformation of LSMeans and LSMean differences will not be applicable, will be reported separately along with standard SAS output (where the latter will not be included in the CSR).

6.4.3. Absolute CFB in waist circumference over time.

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

CFB in waist circumference will be summarized at Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22 for Cohorts 1 & 2, and at Weeks 4, 8, 12, 16, 20, 24 and 28 for Cohort 3 separately. Summary and analysis results will be reported as described in Section 6.2.3.

6.4.4. Absolute CFB in waist-to-hip ratio over time

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

CFB in waist-to-hip ratio will be summarized at Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24 and 28 for Cohort 3 separately. Summary and analysis results will be reported as described in Section 6.2.4.

6.4.5. Shift from baseline in glycemic category (normoglycemia, pre-diabetes, or T2DM) at End of Treatment

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

A shift table for the baseline glycemic category to the End of Treatment glycemic category will be presented by treatment group.

6.4.6. CFB in systolic and diastolic BP over time

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute and change from baseline in systolic and diastolic blood pressure will be summarised descriptively and analyzed as described in Section 6.3.3 at Weeks 2, 4, 6, 8,

10, 12, 16, 18, 22 and 26 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3 separately.

6.4.7. CFB in lipid parameters over time

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute and change from baseline in lipid parameters will be summarised descriptively and analyzed as described in Section 6.3.2.1 at Weeks 2, 4, 6, 8, 10, 12, 16, 18, 22 and 26 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3 separately.

6.4.8. CFB in SF-36v2® subscales and summary scores

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute and change from baseline in the 11 SF-36v2 transformed norm-based T-scores (covering 8 domains, health transition and 2 components) will be summarised descriptively by treatment group and time point as described in Section 5.2.1, at Weeks 16 and 26 for Cohorts 1 & 2 and at Weeks 16, 24 and 32 at Cohort 3 separately.

6.4.9. CFB in Eating-Related Factors Daily Diary

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute and change from baseline in the 7-day average scores for each of the 5 individual items from the Eating-Related Factors Daily Diary will be summarized descriptively by treatment group and time point for each item separately as described in Section 5.2.1, at Weeks 4, 8, 12, 16, 22 and 26 for Cohorts 1 & 2 and at Weeks 4, 8,12, 16, 20, 24, 28 and 32 at Cohort 3 separately.

6.4.10. CFB in PGI-S at each planned clinic visit up through End of Treatment visit

In all cases a dataset similar to Estimand Set 1 as specified in Section 4will be utilised.

PGI-S will be summarized categorically by treatment group and time point for each item as described in Section 5.2.2. Shift tables will also be produced to summarise the categorical changes for each item at each post-dose time-point separately.

6.4.11. CFB in PROMIS® Fatigue

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute and change from baseline in the PROMIS® Fatigue T-scores will be summarized descriptively by treatment group and time point as described in Section 5.2.1, at Weeks 16 and 26 for Cohorts 1 & 2 and at Weeks 16, 24 and 32 for Cohort 3 separately.

6.4.12. CFB in PROMIS® Physical Function

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute and change from baseline in the PROMIS® Physical Function T-scores will be summarized descriptively by treatment group and time point as described in Section 5.2.1, at Weeks 16 and 26 for Cohorts 1 & 2 and at Weeks 16, 24 and 32 for Cohort 3 separately.

6.4.13. CFB in IWQOL-Lite-CT©

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

Absolute and change from baseline for each of the 3 composite scores from the IWQOL-Lite-CT© will be separately summarized descriptively by treatment group and time point as described in Section 5.2.1, at Weeks 16 and 26 for Cohorts 1 & 2 and at Weeks 16, 24 and 32 for Cohort 3 separately.

6.4.14. PGI-C

In all cases a dataset similar to Estimand Set 1 as specified in Section 4 will be utilised.

PGI-C data will be summarized categorically by treatment group and time point for each item as described in Section 5.2.2, at Weeks 4, 16 and 26 for Cohorts 1 & 2 and at Weeks 4, 16, 24 and 32 for Cohort 3 separately.

6.5. Subset Analyses

Absolute values, percentage changes from baseline and absolute changes from baseline in body weight will be summarised descriptively by treatment group, time point and biological sex as described in Section 5.2.1. The tables will present time points as described in Section 6.1.1.1 and baseline is as defined in Section 3.

6.6. Baseline and Other Summaries and Analyses

6.6.1. Baseline Summaries

Baseline tables summarizing the following will be produced by treatment group and overall: age; gender; race; ethnicity; height; body weight; body mass index; waist circumference; waist-to-hip ratio; hip circumference; country; HbA1c; fasting plasma glucose; systolic blood pressure; diastolic blood pressure; pulse rate; total cholesterol; direct LDL cholesterol; HDL cholesterol and triglycerides. The tables should be produced separately for: (1) Cohorts 1 & 2, and (2) Cohort 3.

6.6.2. Study Conduct and Participant Disposition

Participant evaluation groups will show participant disposition for each phase of the study (screening, double blind treatment and follow-up) and will additionally show which participants were analyzed for efficacy (Estimand Set 1 and 2) as well as for safety. Frequency counts and percentages will be supplied for participant discontinuation(s) by treatment group and overall for: (1) Cohorts 1 & 2, and (2) Cohort 3 data, which will be reported in accordance with the sponsor reporting standards.

6.6.3. Study Treatment Exposure

Banked biospecimens will be collected and retained for future analyses, but will not be analyzed specifically for this study and will not be included in the CSR.

6.6.4. Concomitant Medications and Nondrug Treatments

All prior and concomitant medication(s) as well as non-drug treatment(s) will be reported according to current sponsor reporting standards.

6.6.1. Treatment Compliance

Treatment compliance (in %) will be summarised descriptively by treatment group and overall, at time point as described in Section 5.2.1. The time points will be: the Placebo run-in phase; Weeks 2, 4, 6, 8, 10, 12, 16, 18, 22 and 26 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3 separately.

Treatment compliance is defined for each time point separately as: (actual number of IP tablets / planned number of IP tablets) × 100%.

Missing information pertaining to doses administered on site will be imputed as per Section 5.3.

6.6.2. Discontinuations

Participant discontinuations, temporary discontinuations or dose reductions due to adverse events will be detailed and summarised by treatment group and overall.

Data will be reported in accordance with the sponsor reporting standards, where an additional table summarizing participant discontinuations of IP due to adverse events for each system organ class and preferred term will be produced by treatment group and overall also.

Exploratory summaries on the time to discontinuation from the study and time to discontinuation of IP (regardless of study discontinuation or continuation) will be produced separately using Cumulative Incidence Plots as described in Section 5.2.7. For both, participants who discontinue from the study/IP will be censored at the associated discontinuation date.

7. INTERIM ANALYSES

7.1. Introduction

An interim analysis will be performed at least once annually while the study is ongoing, after at least 25% of participants are randomized, with further details provided in the IRC charter. This interim analysis will assess, at a minimum, unblinded safety of the randomized participants. Additional interim analyses for safety and/or efficacy may be performed if needed. Further details will be provided in the IRC charter.

Interim analysis results may be used for internal business decisions including, but not limited to: stopping a dose level, future study planning, stopping for futility, stopping for early success, conducting a sample size re-estimation, or adapting the study after the interim analysis. Before any interim analysis is instigated, the details of the objectives, decision criteria, information dissemination plan, and method of maintaining the study blind as per Pfizer's SOPs will be documented and approved in an IRC charter. In addition, the analysis details must be documented and approved in an interim analysis SAP or final SAP.

7.2. Interim Analyses and Summaries

The interim analyses will be performed using the methodology specified in this SAP and will be outlined in this SAP (with an amendment) or an interim analysis SAP. Any substantial deviations from the SAP's methodology will be fully justified and outlined. Details of the ongoing unblinded safety reviews will be provided in the IRC Charter and/or the interim analysis SAP.

8. REFERENCES

1. ICH Harmonised Guideline E9 (R1); Estimands and Sensitivity Analysis in Clinical Trials; 16 June 2017.

2. American Diabetes Association. Standards of medical care in diabetes – 2019. Diabetes Care 2019;42(Suppl 1):S1-S193.

3. Ware J, Kosinski M, Bjorner J, et al. Development. User's Manual for the SF-36v2® Health Survey. Lincoln (RI): QualityMetric Incorporated; 2007.

4. Ware J, Kosinski M, Bjorner J, et al. SF-36v2® Health Survey: Administration guide for clinical trial investigators: Lincoln (RI): QualityMetric Incorporated; 2008.

5. Patient Reported Outcomes Measurements Information System, FATIGUE: A brief guide to the PROMIS® Fatigue instruments; Feb 28, 2019.

6. Patient-Reported Outcomes Measurements Information System, PHYSICAL FUNCTION, A brief guide to the PROMIS® Physical Function instruments; July 18, 2019.

7. Kolotkin RL, Williams VSL, Ervin CM, et al. Validation of a new measure of quality of life in obesity trials: Impact of Weight on Quality of Life-Lite Clinical Trials Version. Clin Obes 2019;9:e12310.

8. Rubin DB. Multiple Imputation for Nonresponse in Surveys. 1987; Wiley, New York.

9. Blyth CR, Still HA. Binomial confidence intervals. Journal of the American Statististical Association. 78:108-116, 1983, & Casella G. Refining binomial confidence intervals. The Canadian Journal of Statistics, 14:113-129, 1986

10. Pfizer Guidance for Evaluation of QT / QTc Interval Prolongation and Proarrhythmic Potential for Non-antiarrhythmic Drugs; Members of the Cardiovascular Safety & Advisory Council (CVSAC); January 26, 2018.

11. Thomas, N. and Wu, J. (2019). clinDR: Simulation and Analysis Tools for Clinical Dose Response Modeling. R package version 1.9. https://CRAN.R-project.org/package=clinDR

12. Wu, J, Banerjee, A., Jin, B., Menon, S., Martin, S., and Heatherington, A., (2017), Clinical dose-response for a broad set of biological products: A model-based meta-analysis, *Statistical Methods in Medical Research,* Vol. 7, No. 9, 2694-2721.

13. Thomas, N., and Roy, D. (2017). Analysis of clinical dose-response in small-molecule drug development: 2009-2014. *Statistics in Biopharmaceutical Research,* Vol. 9, No. 2, 137-146.

14. Thomas, N., Sweeney, K., and Somayaji, V. (2014). Meta-analysis of clinical dose response in a large drug development portfolio, *Statistics in Biopharmaceutical Research,* Vol. 6, No.4*,* 302-317*.*

15. HealthMeasures. http://www.healthmeasures.net/score-and-interpret/calculatescores/scoring-instructions. 2020.

Appendix 1. Summary of Efficacy Analyses

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Percent change from baseline in body weight at End of Treatment (and also at Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24 and 28 for Cohort 3.)	Summary	Estimand Set 1	DFI: Censor all values post discontinuation MV: Not imputed	N/A
	Primary analysis			MMRM
	Supplementary analyses			Emax (x 2)
	Exploratory analysis (summary)	Estimand Set 2	MPD: Remove participants DFI: Censor all values post discontinuation MV: Not imputed	N/A
	Exploratory analysis			MMRM
Response as defined by a body weight loss of ≥5% from baseline at End of Treatment	Summary	Estimand Set 1	DFI: Censor all values post discontinuation MV: Not imputed	N/A
	Summary		DFI: Censor all values post discontinuation MV: Imputed using multiple imputation	N/A
	Secondary analysis			Logistic regression
Absolute change from baseline in waist circumference at End of Treatment (and at Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24 and 28 for Cohort 3.)	Summary	Estimand Set 1	DFI: Censor all values post discontinuation MV: Not imputed	N/A
	Secondary analysis			MMRM
Absolute change from baseline in waist-to-hip ratio at End of Treatment (and at Weeks 2, 4, 6, 8, 10, 12, 16, 18 and 22 for Cohorts 1 & 2 and at	Summary	Estimand Set 1	DFI: Censor all values post discontinuation MV: Not imputed	N/A
	Secondary analysis			MMRM

Weeks 4, 8, 12, 16, 20, 24 and 28 for Cohort 3.)				
Absolute change from baseline in HbA1c at Weeks 16 and 26 for Cohorts 1 & 2 and at Weeks 16, 24 and 32 for Cohort 3.	Summary Secondary analysis	Estimand Set 1	DFI: Censor all values post discontinuation MV: Not imputed	N/A MMRM
Absolute change from baseline in FPG at Weeks 2, 4, 6, 8, 10, 12, 16, 18, 22 and 26 for Cohorts 1 & 2 and at Weeks 4, 8, 12, 16, 20, 24, 28 and 32 for Cohort 3.	Summary Secondary analysis	Estimand Set 1	DFI: Censor all values post discontinuation MV: Not imputed	N/A MMRM
CFB in fasting plasma insulin, HOMA-IR and HOMA-B at Weeks 4, 8, 12, 16, 22 and 26 for Cohort 1 & 2 and at Weeks 4, 8, 12, 15, 20, 24 and 28 for Cohort 3.	Summary Exploratory analysis	Estimand Set 1	DFI: Censor all values post discontinuation MV: Not imputed	N/A MMRM

Abbreviations: DFI = discontinuation from IP; MPD = major protocol deviations; MV = missing values (note this includes missing values produced through censoring).

End of Treatment is *defined as Week 26 for Cohorts 1 and 2 and as Week 32 for Cohort 3.*

Appendix 2. Bayesian Statistical Methodology Details

Emax model without baseline interaction (Section 5.2.5)

A dataset (either .txt or .csv) of the following format should be produced by programming for use in R by the reporting statistician and QC statistician. Note, column headers should be labelled as specified below (including capitalization), as R is case sensitive:

Dose	Mean	SE
0	0	0.3
40	-0.01	0.4
80	-0.02	0.5
120	-0.03	0.3
160	-0.05	0.4
200	-0.1	0.5

Note the danuglipron dose groups refers to the results from the 1-week titration arms only, from the model described in Section 6.1.1.1. The Placebo group is the combined data from Cohorts 1 & 2. The residual standard deviation at End of Treatment from the unstructured covariance matrix from the associated MMRM will also be provided to the statisticians. The 4-parameter Emax model will be fit using the latest version (currently v2.1) of the clinDR package.[11] This analysis will be conducted by the study statistician. A different statistician will conduct QC of the analysis. The outputs of the analysis will be provided as .txt files to the programming team for inclusion in the final CSR table and figure formats.

Compound-specific information will be used to specify independent prior distributions for the placebo response (E_0), and the difference in response between the highest dose (*dTarget*=200 mg) and placebo, denoted by *difTarget*, based on data from the C3421002 study. Non-informative priors will be used for these parameters. Note that the E_{max} parameter is derived from the other parameters and is thus not explicitly supplied. The residual standard deviation, *sigma*, is assigned a uniform prior distribution over a range we are confident will include the population value.

Parameter	Prior
E_0	Normal(Mean = 0, SD = 6.8)
difTarget	Normal(Mean = -0.067, SD = 6.8)
sigma	Uniform(lb = 0.017, ub = 0.272)

In addition, the projected ED_{50} is P_{50} = ~6mg based on data from the C3421002 study. It will be combined with the predictive prior distribution for the log(ED_{50}/P_{50}), obtained from meta-data on approximately 225 compounds (from 3 references[12,13,14]), to specify an informative prior distribution for the ED_{50}. The distribution of the Hill parameter is the predictive distribution from the meta-data. The current distributions are listed below. They are the default distributions in clinDR. These default distributions will be updated if the meta-data and their analysis are updated before the completion of the current study.

Parameter	Prior
log(Hill)	*t*(Mean = 0, SD = 0.84, df = 5)
$log(ED_{50})$	log(P_{50}) + *t*(Mean = 0, SD = 1.74, df = 5)

The bivariate predictive distribution of these parameters also includes a prior correlation, which is currently -0.43 based on the analysis of the meta-data, which also would be updated if the historical analysis is updated.

The default burn-in and number of samples will be utilized along with thinning of 20, which will include 3 chains to assess convergence. Model diagnostics will be examined including trace and auto-correlations plots. If these raise concerns over model convergence, additional burn-ins, samples and thinning will be attempted to improve convergence. Changes to the priors above may also be considered (e.g. increase precision of E_0 and *difTarget*) to improve convergence if deemed necessary. The final diagnostic plots will not be included in the clinical study report.

The following R code is included as an example that will be used as a basis for the analysis:

```
library(clinDR)
compileStanModels()
mmrmRes <- read.csv("LSmeans.csv",header=T,stringsAsFactors=F)
# Determine 'effective' subject numbers based on MMRM SD at End of Treatment:
mmrm_sd <- 0.068 # Provided by programming
mmrmRes$N <- trunc((mmrm_sd/mmrmRes$StdErr)^2,0)

# Set-up priors and MCMC options:
prior_mmrm <- emaxPrior.control(epmu=0, epsca=6.8, difTargetmu=-0.067, difTargetsca=6.8,
dTarget=200, effDF=9999, p50=6, sigmalow=0.017, sigmaup=0.272)
mcmc_mmrm <- mcmc.control(chains=3,thin=20,seed=169)

### Run Emax model: ###
emaxMMRM <-
fitEmaxB(mmrmRes$Estimate,mmrmRes$dose,prior_mmrm,modType=4,count=mmrmRes$N,msSat
=mmrm_sd^2,mcmc=mcmc_mmrm)

# Diagnostics and output:
stan_trace(emaxMMRM$estanfit) # Look at trace
stan_dens(emaxMMRM$estanfit) # Look at densities
stan_ac(emaxMMRM$estanfit) # Look at autocorrelation
summary(emaxMMRM) # Summary of model parameters
plot(emaxMMRM) # Look at fitted vs. observed data
emaxMMRMout <- predict(emaxMMRM,dosevec=mmrmRes$dose,clev=0.90) # Get dose predictions
```

Model parameters, posterior medians and 90%/95% credible intervals as specified in Section 5.2.5 will be output and provided back to the programming team after QC is complete.

Emax model with baseline interaction (Section 5.2.6)

A dataset of the same format to above should be produced by programming for use in R by the reporting statistician and QC statistician. Note, column headers should be labelled as specified (including capitalization), as R is case sensitive.

The residual standard deviations at End of Treatment from the unstructured covariance matrices from *each* of the associated MMRM models will also be provided to the statisticians.

The same process for priors, model fitting, checking of model convergence and QC as above will be implemented for this model along with similar R code as to above. The only major difference in R code is the requirement to take the average of the residual standard deviations from the separate MMRM outputs which will be used as the global standard deviation in model fitting:

```
mmrmDoseRes <- read.csv(Dose_LSmeans.csv",header=T,stringsAsFactors=F)
# Covariance matrix:
mmrmDoseCov <- read.csv(Dose_CovParams.csv",header=T,stringsAsFactors=F)
# Calculate global SD:
mmrm_global_sd <- mean(mmrmDoseCov$SD)
# Merge datasets:
mmrmDoseRes <- merge(mmrmDoseRes,mmrmDoseCov,by="dose")
mmrmDoseRes$N1 <- trunc((mmrmDoseRes$SD/mmrmDoseRes$StdErr)^2,0)
mmrmDoseRes$Glob_SD <- mmrm_global_sd
mmrmDoseRes$N <- trunc((mmrmDoseRes$Glob_SD/mmrmDoseRes$StdErr)^2,0)
…
emaxMMRM_Int <-
fitEmaxB(mmrmDoseRes$Estimate,mmrmDoseRes$dose,prior_mmrm,modType=4,count=mmrmDos
eRes$N,msSat=mmrm_global_sd^2,mcmc=mcmc_mmrm)
```

Model parameters, posterior medians and 90%/95% credible intervals as specified in Section 5.2.6 will be output and provided back to the programming team after QC is complete.

Appendix 3. Traditional Statistical Methodology Details

The following SAS code is to be used as a guide for implementation.

Example SAS code for MMRM Model (strata included):

```
proc mixed data = dataset method=reml;
        class subject treatment time strata;
        model cfb = treatment base time base* time time *treatment strata/ddfm=kr residual
outp=resid_out;
        repeated time /subject=subjid type = un;
        lsmeans treatment*time/diff cl alpha=0.1;
        ods output lsmeans=lsmeans_out;
        ods output diffs=diffs_out;
        ods output CovParms=CovParms_out;
run;
```

Example SAS code for Cumulative Incidence Plots:

```
proc lifetest data = dataset method=km plots=cif(test) outcif=cifatrisk intervals=0 to 20 by 2;
        strata treatment;
        time day*censor(1)/eventcode=0;
run;
```

NOTE: the censor variable has a value = 1 when the related time is censored and has a value = 0 when the event of interest occurs. There should be no other values available for this censored variable in this dataset (including missing values). If required, missing observations should be removed prior to analysis.

Example SAS code for Proportion of Responses (with multiple imputation):

Assume the SAS dataset is in a long format. The variable 'treatment' should be coded similar to 'Placebo', 'Danuglipron 40 mg BID', 'Danuglipron 80 mg BID',… 'Danuglipron 200 mg BID', 'Danuglipron 120 mg BID (2-week titration)', 'Danuglipron 160 mg BID (2-week titration)' and 'Danuglipron 200 mg BID (2-week titration)' so that when sorted in descending order 'Placebo' comes first.

```
proc sort data=analysis out= analysisl;
        by subjid time treatment strata;
run;

* Create wide dataset for multiple imputation;
proc transpose data= analysisl out=analysisw  prefix=week;
        by subjid treatment strata;
        id time;
        var cfb_log;
run;

* Perform multiple imputation;
proc mi data=analysisw seed=169 nimpute=20 out= analysis_mi;
        class treatment strata;
        fcs nbiter=10 reg(/details);
        var base treatment strata week2 week4 week6 week8 week10 week12 week16 week18
week22 week26;
run;

* Determine responders and non-responders at End of Treatment for all imputed datasets;
data analysis_mi_26;
        set analysis_mi;
        value = week26;
        if value < log(0.95) and week26 ne . then resp = 1;
        if value >= log(0.95) and week26 ne . then resp = 0;
run;

* Create datasets and combine proportions;
proc freq data=analysis_mi_26;
        tables _imputation_*dose*resp/out=prop_mi outpct;
run;

data prop_mi_0;
        set prop_mi;
        if resp=0;
        keep _imputation_ treatment count;
        rename count=count_0;
run;
data prop_mi_1;
        set prop_mi;
        if resp=1;
        keep _imputation_ treatment count;
        rename count=count_1;
run;
data prop_mi_combined;
```

```
        merge prop_mi_0 prop_mi_1;
        by _imputation_ treatment;
        total = count_0 + count_1;
        p = count_1/total;
        q = count_0/total;
        se_p = sqrt((p*q)/total);
run;

proc sort data=prop_mi_combined;
        by treatment _imputation_;
proc mianalyze data=prop_mi_combined  alpha=0.1;
    by treatment;
    modeleffects p;
    stderr se_p;
    ods output parameterestimates=prop_mi_out;
run;
```

Example SAS code for Logistic Regression Model (with multiple imputation):

The same imputed datasets as produced above for the 'Proportion of Responses' will be utilized. Note: the imputed dataset should be ordered so that the group 'Placebo' and 'Resp'=1 comes first to ensure that the reference group is Placebo and the odds ratios are for a response = 1.

```
* Fit logistic regressions to each imputed dataset and combine results;
proc sort data=analysis_mi_26;
        by _imputation_ descending treatment descending resp;
proc logistic data = analysis_mi_26 order=data;
        by _imputation_;
        class resp treatment strata;
        model resp = treatment strata base/alpha=0.1;
        oddsratio dose/diff=all;
        ods output OddsRatiosWald=OddsRatiosWald_mi;
run;

data OddsRatiosWald_mi;
        set OddsRatiosWald_mi;
        if effect = . then delete;
        estimate = log(OddsRatioEst);
        SE = (log(UpperCL) - log(LowerCL))/(2*QUANTILE('NORMAL',0.95));
run;

proc sort data=OddsRatiosWald_mi;
        by effect _imputation_;
proc mianalyze data=OddsRatiosWald_mi  alpha=0.1;
    by effect;
    modeleffects estimate;
    stderr SE;
    ods output parameterestimates=OddsRatiosWald_mi_out;
run;
```

Appendix 4. Categorical Classes for ECG and Vital Signs of Potential Clinical Concern

Categories for QTcF

Absolute value of QTcF (msec)	>450 and ≤480	>480 and ≤500	>500
Increase from baseline in QTcF (msec)	>30 and ≤60	>60	

Categories for PR and QRS

PR (ms)	max. ≥300	
PR (ms) increase from baseline	Baseline >200 and max. ≥25% increase	Baseline ≤200 and max. ≥50% increase
QRS (ms)	max. ≥140	
QRS (ms) increase from baseline	≥50% increase	

Categories for Vital Signs

Supine Systolic BP (mm Hg)	min. <90	
Supine Systolic BP (mm Hg) change from baseline	max. decrease ≥30	max. increase ≥30
Supine Diastolic BP (mm Hg)	min. <50	
Supine Diastolic BP (mm Hg) change from baseline	max. decrease ≥20	max. increase ≥20
Supine pulse rate (bpm)	min. <40	max. >120

Appendix 5. C-SSRS Mapped to C-CASA - Suicidal Ideation and Behavior Events and Codes

Table 2. C-SSRS Mapped to C-CASA (Suicidality Events and Codes)

Event Code	C-CASA Event	C-SSRS Response
Suicidal Ideation		
1	Passive	"Yes" on "Wish to be dead"
2	Active: Nonspecific (no method, intent, or plan)	"Yes" on "Non-Specific Active Suicidal Thoughts"
3	Active: Method, but no intent or plan	"Yes" on "Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act"
4	Active: Method and intent, but no plan	"Yes" on "Active Suicidal Ideation with Some Intent to Act, without Specific Plan"
5	Active: Method, intent, and plan[1]	"Yes" on "Active Suicidal Ideation with Specific Plan and Intent"
Suicidal Behavior		
1	Completed suicide	"Yes" on "Completed Suicide"
2	Suicide attempt	"Yes" on "Actual Attempt"
3	Interrupted attempt	"Yes" on "Interrupted attempt"
4	Aborted attempt	"Yes" on "Aborted attempt"
5	Preparatory actions toward imminent suicidal behaviors	"Yes" on "Preparatory Acts or Behavior"
Self-injurious behavior, no suicidal intent		
	Self-injurious behavior, no suicidal intent	"Yes" on "Has subject engaged in Non-suicidal Self-Injurious Behavior?"

[1]According to C-SSRS, the definition of plan includes intent (i.e., intent to complete the suicide is implicit with the concept of plan). Thus, there is no need for the category method and plan, but no intent.

Appendix 6. PROMIS-Fatigue

The raw score for each patient should be translated into a T-Score using the table below (i.e. 9 would be 29.1 on the T-Score scale), which was provided by the Health Measures Scoring Service[15]:

Raw	Scale
9	29.1
10	32.8
11	36
12	38.3
13	40.3
14	42
15	43.7
16	45.1
17	46.4

18	47.6
19	48.8
20	50
21	51.1
22	52.2
23	53.3
24	54.4
25	55.5
26	56.6
27	57.6
28	58.7
29	59.8
30	60.9
31	62
32	63.1
33	64.2
34	65.3
35	66.4
36	67.6
37	68.8
38	70.1
39	71.4
40	72.8
41	74.3
42	76
43	78
44	80.3
45	83

Appendix 7. PROMIS-Physical Function

The raw score for each patient should be translated into a T-Score using the table below (i.e. 13 would be 21.5 on the T-Score scale), which was provided by the Health Measures Scoring Service[15]:

Raw	Scale
13	21.5
14	25.1
15	26.8
16	28.1
17	29.1
18	30
19	30.7

20	31.4
21	32.1
22	32.7
23	33.3
24	33.8
25	34.3
26	34.8
27	35.3
28	35.8
29	36.3
30	36.8
31	37.2
32	37.7
33	38.1
34	38.6
35	39
36	39.5
37	39.9
38	40.4
39	40.8
40	41.3
41	41.8
42	42.2
43	42.7
44	43.2
45	43.7
46	44.2
47	44.7
48	45.2
49	45.7
50	46.3
51	46.9
52	47.5
53	48.1
54	48.8
55	49.5
56	50.3
57	51.1
58	52
59	53
60	54.1
61	55.4

62	56.8
63	58.7
64	61
65	66

Appendix 8. IwQOL-Lite CT

The Physical composite score comprises Items 1-5, 16, and 17. The Physical Function composite comprises Items 1-3, 16, and 17, with Item 4 ("Uncomfortable in small seats") and Item 5 ("Bodily pain") excluded. The Psychosocial composite comprises Items 6-8, 9-15, 18, 19, and 20.

Each response is assigned a score as follows: 1= "Never" or "Not at all true"; 2= "Rarely" or "A little true"; 3= "Sometimes" or "Moderately true"; 4= "Usually" or "Mostly true"; 5= "Always" or "Completely true".

The scoring algorithm is as follows:

- Scores for each composite should only be compared if a minimum of 50% of the iterms for that composite are non-missing, for the IWQOL-Lite-CT total score if a minimum of 75% of all items are non-missing. Psychosocial 7 of 13; Physical 4 of 7; Physical Function 3 of 5; Total 15 of 20).

- If the minimum number of items is answered for a composite, the composite score is calculated by transforming the composite item average to the worst (0) to 100 (best) metric using the following formula for every patient at each time point:

$$100 \times (S_{max} - C_{avg})/(S_{max} - S_{min})$$

where

- C_{avg} is the raw average score of all non-missing item responses in the composite; this average must be a number between 1 and 5, inclusive

- S_{max} is the maximum possible raw score value, i.e., 5

- S_{min} is the minimum possible raw score value, i.e., 1

- Inserting the maximum and minimum possible score values, the formula is reduced to 100 x $(5-C_{avg})/4$.

Appendix 9. List of Abbreviations

Abbreviation	Term
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine transaminase
AST	aspartate aminotransferase
BID	twice daily
BLQ	below the limit of quantitation
BP	blood pressure/ bodily pain
C-CASA	Columbia Classification Algorithm of Suicide Assessment
CFB	Change from baseline
CRF	case report form
CSR	clinical study report
C-SSRS	Columbia-suicide severity rating scale
C_{trough}	Lowest concentration
CV	coefficient of variation
DFI	discontinuation from IP
dL	Deciliter
ECG	electrocardiogram
E_0	Placebo effect
ED_{50}	Dose producing 50% of Emax
eGFR	estimated glomerular flitration rate
E_{max}	Maximum response
FDA	Food and Drug Administration (United States)
FPG	Fasting Plasma Glucose
FPI	Fasting Plasma Insulin
GGT	Gamma-glutamyl transpeptidase
GH	Growth hormone
H	hours
HbA1c	Hemoglobin A1c
HOMA-B	Homeostatis model assessment of β-cell function
HOMA-IR	Homeostatis model assessment -insulin resistance
ICH	International Council for Harmonisation
IRC	internal review committee
IP	Investigational Product
IWQOL Lite-CT	Impact of Weight on Quality of Life – Lite Clinical Trials Version
L	liter

LDL	low-density lipoprotein
LLQ	lower limit of quantitation
LS	least-squares
MAR	Missing at random
MCS	mental component summary scores
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
MH	mental health
MI	multiple imputation
mIU	milli-international units
MMRM	mixed-effects model with repeated measures
μmol	micromol
MPD	major protocol deviations
msec	millisecond
MV	Missing values
N/A	not applicable
PCC	potential clinical concern
PCOA	patient centered outcome assessment
PCS	physical component summary
PD	protocol deviation
PF	physical functioning
PGI-C	Patient's global impression of change
PGI-S	Patient's global impression of severity
PHQ-9	Patient health questionnaire-9
PK	Pharmacokinetic
PROMIS	Patient-reported outcomes measurements information system
PT	preferred term
Q1	First quartile
Q3	Third quartile
QC	Quality control
QQ	quantile-quantile
QTc	corrected QT
QTcF	corrected QT (Fridericia method)
RE	role-emotional
RP	role-physical
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation

SF	social functioning
SF-36	Short Form-36 Health Survey
SoA	schedule of activities
SOP	standard operating procedure
T2DM	Type 2 diabetes mellitus
T4	thyroxine
TEAE	treatment emergent adverse event
TSH	thyroid stimulating hormone
U	Units
ULN	upper limit of normal
US	United States
VT	vitality