Protocol C3421005

A 16-WEEK, PHASE 2B, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL GROUP STUDY TO EVALUATE THE EFFICACY AND SAFETY OF TWICE DAILY PF-06882961 ADMINISTRATION IN ADULTS WITH TYPE 2 DIABETES MELLITUS INADEQUATELY CONTROLLED ON METFORMIN OR DIET AND EXERCISE

Statistical Analysis Plan
(SAP)

Version: 2.0

Date: 21 JUL 2021

NOTE: *Italicized* text within this document has been taken verbatim from the Protocol.

[redacted]

TABLE OF CONTENTS

CCI [redacted]

[redacted]

LIST OF TABLES

LIST OF FIGURES

No table of figures entries found.

APPENDICES

1. VERSION HISTORY

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
2 21 Jul 2021	Amendment 1 19 May 2020	Updates based on SAP template update, protocol amendment, A&R plan review and Blinded Data Reviews.	Minor updates to page header and footers and addition of list of abbreviations appendix table. *Rationale:* update to latest SAP template. CCI [redacted] Sections 2.1.2.1, 3.2, 4 & 6.3.4: Updated and clarified Secondary Estimand definition in reference to intercurrent events for a response in HbA1c (and also separately, Body Weight). *Rationale:* SAP wording was not consistent with protocol, reverted back to protocol wording and added some clarifications. CCI [redacted] Sections 3.5.1, 3.5.1.1 & 3.5.6: Minor updates related to adverse event CCI [redacted] definitions. *Rationale:* minor clarifications based on occurrences identified during BDR reviews. Sections 3.5.3.1 & 6.6.2.1: Added %change from baseline summaries of labs of interest, added box and whisker plots, added MMRM analyses and the parameters bile acids, eGFR and GGT. *Rationale:* permit a more comprehensive review of additional safety parameters.

[redacted]

			Sections 3.5.3.2 & 6.6.2.2: Added a table summarizing specific clinical laboratory parameters of interest where values met a flag or alert level. *Rationale:* permit a more comprehensive review of lab parameters of interest. Sections 3.5.4 & 3.5.5: Added that averages of triplicates would be used for baseline vitals and ECGs. *Rationale:* minor clarification. Section 4.1: Removed listing comparing randomization stratum to concomitant medication categorization. *Rationale:* clinical database was to have the correct randomization stratum included, so listing won't provide value. Sections 5.1 & Section 6.6.1: Added description of hypotheses and error rates for secondary endpoints. *Rationale:* pre-specify type I error rates for secondary objectives. Section 5.2.3: Added p-values to be reported for various outputs and clarified two-sided. *Rationale:* aid in future publication writing. Sections 5.2.7, 6.5.6, 6.6.1.3 & Appendix 4: Re-named Kaplan-Meier Curve Plots to Cumulative Incidence Plots and provided more details along with SAS code. *Rationale:* refer to the plots using the correct terminology with more clarity. Section 6.1.1.1.1: Clarified details on summaries, added box and whisker plots and added figures restricted to Week 16. *Rationale:* provide more information on primary and other key endpoints. Section 6.5.2: Added that COVID-19 related outputs would follow sponsor reporting

[redacted]

			standards. *Rationale:* ensure appropriate reporting of COVID-19 related events. Section 6.5.3: Removed Demographic data summary. *Rationale:* Section duplicated information included in Section 6.5.1 Sections 6.5.2, 6.5.4, 6.5.5, 6.5.6, 6.6.1.2, 6.6.1.4, 6.6.2, 6.6.2.2 & 6.6.5: Added overall summary. *Rationale:* provide an overview of entire study population. Section 6.6.1.4: Added information on how percentage was calculated for % of participants with ongoing AEs of interest. *Rationale:* clearly specify algorithm. Section 6.6.1.5: Added summaries on time to discontinuation from IP due to gastrointestinal disorders AEs. *Rationale:* describe more information the timing of discontinuation for certain AEs. Sections 6.6.3 & 6.6.4: Removed summaries of differences to placebo for vital signs and ECG parameters and replaced with MMRM analyses. *Rationale:* provide a more comprehensive assessment of these safety parameters. CCI [redacted]
1 8 Oct 2019	Original 3 Apr 2019	N/A	N/A

[redacted]

2. INTRODUCTION

This statistical analysis plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study C3421005. 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. Study Objectives, Endpoints, and Estimands

Objectives	***Endpoints***	***Estimands***
Primary:	***Primary:***	***Primary:***
To compare the effect of multiple dose levels of PF-06882961 versus placebo on glycated hemoglobin (HbA1c) in participants with type 2 diabetes mellitus (*T2DM*) *on stable doses of metformin and/or diet and exercise.*	*Change from baseline in HbA1c at Week 16.*	*Estimand 1A: This estimand is intended to provide a population level estimate of the mean treatment effect (PF-06882961 versus placebo) on a continuous endpoint in participants on stable doses of metformin and/or diet and exercise without the benefit of glycemic rescue medication while on treatment.*
Secondary:	***Secondary:***	***Secondary:***
To compare the effect of multiple dose levels of PF-06882961 versus placebo on glycemic control in participants with T2DM on stable doses of metformin and/or diet and exercise.	*Response as defined by an HbA1c <7% at Week 16.*	*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 responder endpoint in participants on stable doses of metformin and/or diet and exercise without the benefit of glycemic rescue medication.*
	Change from baseline in HbA1c at Weeks 2, 4, 6, 8 and 12.	*Estimand 1A as above.*
	Change from baseline in fasting plasma glucose at Weeks, 2, 4, 6, 8, 12 and 16.	*Estimand 3: This estimand will be the same as 1A.*
To compare the effect of multiple dose levels of PF-06882961 versus placebo on body weight in participants with T2DM on stable doses of metformin and/or diet and exercise.	*Change from baseline in body weight at Weeks 2, 4, 6, 8, 12 and 16.*	*Estimand 4: This estimand will be the same as 1A.*
To characterize the safety and tolerability of multiple dose levels of PF-06882961 administered to participants with T2DM on stable	*Incidence of treatment emergent adverse events [adverse events (AEs) and serious adverse events (SAEs)], clinical laboratory abnormalities, vital signs (blood pressure and pulse rate) and electrocardiogram (ECG)*	*There are no defined estimands for these endpoints and they will be analyzed using Pfizer data standards as applicable.*

[redacted]

doses of metformin and/or diet and exercise.	*parameters (heart rate, QT, QTcF, PR and QRS intervals).*	
CCI		
CCI		

CCI [redacted]	[redacted]	
*For **all** endpoints, baseline is defined as the result closest prior to dosing at Visit 3 (Day 1).*		

2.1.1. Primary Estimand(s)

The primary estimand (Estimand 1A) will be the population average treatment effect on the change from baseline in HbA1c at Week 16 of PF-06882961 compared to placebo in the absence of glycemic rescue medication while on treatment and stable doses of background metformin and/or diet and exercise. This reflects a combination of the 'Hypothetical' and 'While on treatment' strategies as outlined in the ICH-E9 (R1) draft guidance (1). It includes the following 4 attributes:

- Population: Defined by the inclusion and exclusion criteria to reflect participants with T2DM who are on a background of metformin and/or diet & exercise
- Variable: Change from baseline in HbA1c (%) at Week 16
- Intercurrent events: All data after an intercurrent event of glycemic rescue medication and/or discontinuation of investigational product (IP) will be censored
- Population-level summary: Difference of variable means between PF-06882961 (each dose considered separately) and placebo

2.1.2. Secondary Estimand(s)

2.1.2.1. Estimand Related to Achieving HbA1c <7% at Week 16

A secondary estimand (Estimand 2) will be the population odds ratio of the treatment effect of achieving HbA1c <7% at Week 16 of PF-06882961 compared to placebo in the absence of glycemic rescue medication in participants while on treatment and on stable doses of background metformin and/or diet and exercise. This reflects a combination of the 'Hypothetical' and 'While on treatment' strategies as outlined in the ICH-E9 (R1) draft guidance (1). It includes the following 4 attributes:

- Population: Defined by the inclusion and exclusion criteria to reflect participants with T2DM who are on a background of metformin and/or diet & exercise
- Variable: Response as defined by an HbA1c <7% at Week 16
- Intercurrent events: Any participant with an intercurrent event of glycemic rescue medication and/or discontinuation of IP prior to Week 16 will have all subsequent data

[redacted]

censored if not missing. Missing or censored values at Week 16 will be imputed (assuming missing at random [MAR]) to determine a response classification.

- Population-level summary: Odds ratio for a response between PF-06882961 (each dose considered separately) and placebo

2.1.2.2. Estimands Related to Fasting Plasma Glucose and Body Weight

These estimands (Estimands 3 and 4) will utilize the same approach as Estimand 1A (Section 2.1.1) for the associated endpoint(s).

CCI [redacted]

2.2. Study Design

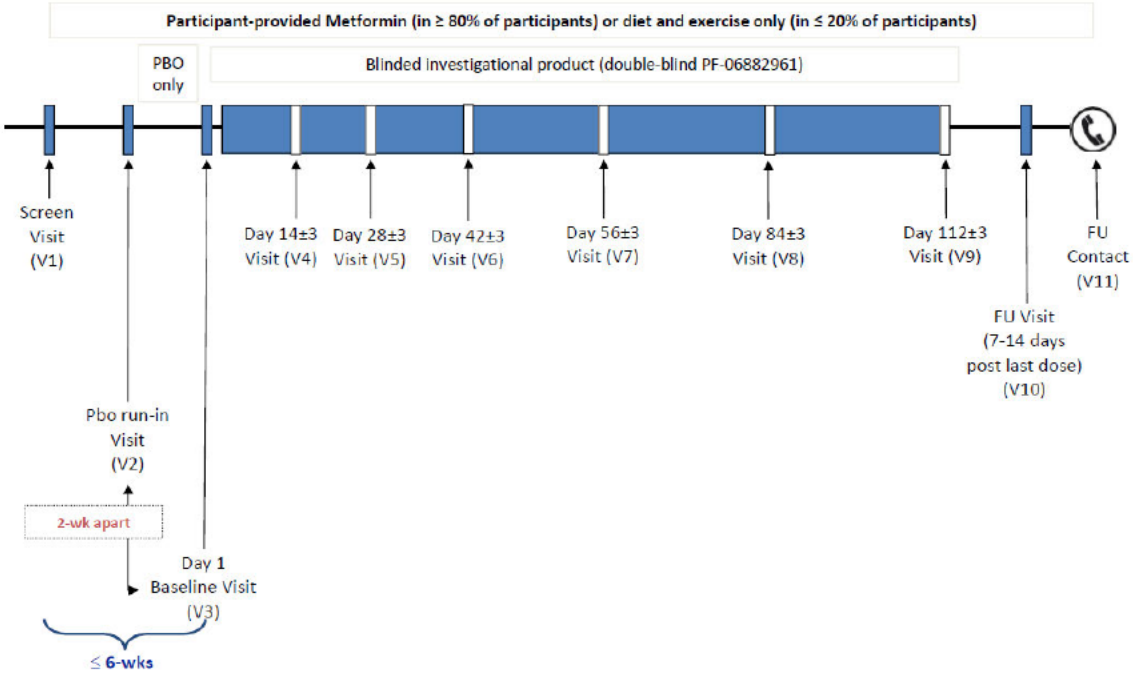
This Phase 2b, multi-center, randomized, double-blind, placebo-controlled, 6 - arm, parallel group, study will assess efficacy and safety of twice daily administration of PF-06882961 in adult participants with T2DM inadequately controlled on metformin monotherapy or diet and exercise alone. At least 80% of the enrolled participants are required to be on metformin prior to screening.

[redacted]

Following the screening period to confirm eligibility (up to 4-weeks), the study will consist of a 2-week placebo run-in period, prior to randomization on Day 1. The treatment period will be 16 weeks, followed by an approximate 4-week follow-up. The total duration of participation in this study is approximately 22 weeks, not including the screening period.

Participants will be randomized to one of the following 6 treatment arms (in a 1:1:1:1:1:1 ratio): placebo, 2.5 mg BID, 10 mg BID, 40 mg BID, 80 mg BID or 120 mg BID. *Dosing will occur with food twice daily, and up to 6 weeks of the 16-week dosing duration will be used for dose titration to maximize tolerability of PF-06882961.*

Approximately 400 participants (approximately 67 participants per treatment arm) will be randomized to ensure completion of approximately 300 participants (approximately 50 participants per treatment arm). Randomization will be stratified according to background diabetes treatment (presence or absence of metformin therapy).



3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

Baseline is defined as the result closest prior to dosing at Visit 3 (Day 1) for all endpoints.

[redacted]

3.1. Primary Endpoint(s)

The primary efficacy endpoint is the change from baseline in HbA1c at Week 16.

3.2. Secondary Endpoint(s)

- Response as defined by an HbA1c <7% at Week 16

This endpoint will have two levels: 'Response' and 'Non-response'. The former will be based on participants having an HbA1c < 7% at Week 16, otherwise participants with a value ≥ 7% will be classed as having a 'Non-response'. Participants with an intercurrent event of starting glycemic rescue medication and/or discontinuation of IP prior to Week 16 will have all subsequent data censored if not missing. If glycemic rescue medication and/or discontinuation of IP occurs at Week 16, the data will not be censored. The intercurrent event that occurs first during the study will take precedence for determining the censoring of data. Missing or censored values at Week 16 will be imputed as described in Section 5.3.

- Change from baseline in HbA1c at Weeks 2, 4, 6, 8 and 12
- Change from baseline in fasting plasma glucose at Weeks 2, 4, 6, 8, 12 and 16
- Change from baseline in body weight at Weeks 2, 4, 6, 8, 12 and 16

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 %change from baseline in body weight will be calculated.

CCI [redacted]

CCI [redacted]

3.4. Baseline Variables

Baseline measures will be included as a covariate in all applicable statistical models along with the stratification variable of whether or not a participant was on background metformin at screening (see Section 4.1).

The statistician may conduct further exploratory analyses into the effect of covariates and factors (such as gender, 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.

3.5. Safety Endpoints

3.5.1. Adverse Events

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 6.6.1).

Tier 1 events: These are prespecified events of clinical importance and are maintained in a list in the product's Safety Review Plan (or similar, e.g. Safety Surveillance 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 treatment group.

[redacted]

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

3.5.1.1. 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 will include all TEAEs starting 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 16 weeks of treatment or discontinuing from IP earlier).

3.5.2. 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 8.2.5.2.

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.

3.5.3. 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).

3.5.3.1. 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 (ALT, AST, alkaline phosphatase, total bilirubin, bile acids and gamma-glutamyl transferase [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

3.5.3.2. 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
Fasting Plasma Glucose	< 70	≤ 54	mg/dL
	≥ 270	≥ 270	
Amylase	>ULN	Pfizer standard flag for PCC	U/L
Lipase	>ULN	Pfizer standard flag for PCC	U/L
Calcitonin	>ULN	Pfizer standard flag for PCC	pg/mL

[redacted]

Alanine Aminotransferase	≥ 2 ULN	Pfizer standard flag for PCC	U/L
Aspartate Aminotransferase	≥ 2 ULN	Pfizer standard flag for PCC	U/L
Alkaline Phosphatase	≥ 2 ULN	Pfizer standard flag for PCC	U/L
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

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).

3.5.4. 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.

3.5.5. 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. The average of the triplicate readings collected at each appropriate assessment time will be calculated for each ECG parameter.

Baseline will be defined as the average of the triplicate measurements 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.

CCI [redacted]

[redacted]

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.

Population	***Description***
Enrolled	*All participants who sign the ICD.*
Randomly assigned to investigational product (IP)	*All participants randomly assigned to IP regardless of whether or not IP was administered.*
Evaluable	*All participants randomly assigned to IP and who take at least 1 dose of IP. Participants will be analyzed according to the randomized intervention.*
Safety Analysis Set	*All participants randomly assigned to IP and who take at least 1 dose of IP. Participants will be analyzed according to the product they actually received.*

[redacted]

Defined Population for Analysis	*Description*
Estimand Set 1A	*All* evaluable *participants. For participants who discontinue IP and/or receive glycemic rescue medication, all subsequent values will be censored.*
CCI [redacted]	[redacted]
[redacted]	[redacted]
Estimand Set 2	All evaluable participants. Participants who receive glycemic rescue medication and/or discontinue IP prior to Week 16 will have their Week 16 value censored (if not missing). Missing or censored values at Week 16 will be imputed as described in Section 5.3.
CCI [redacted]	[redacted]

4.1. Strata Misallocations

Participants who are randomized to the wrong stratum, 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

Statistical inference will be based on the primary endpoint: change from baseline in HbA1c at Week 16.

The null hypothesis is that there is no difference between PF-06882961 and placebo on the primary endpoint. The alternative hypothesis is that PF-06882961 is superior (i.e. greater reduction) to placebo on the primary endpoint.

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

Each dose of PF-06882961 will be tested separately compared to placebo.

No adjustment for multiple comparisons will be made.

Similar hypotheses will be applied to the secondary endpoints (change from baseline in fasting plasma glucose/body weight at Week 16 & response as defined by an HbA1c <7% at Week 16), where the type I error rate will also be 5% (1-sided). Note the alternative

[redacted]

hypothesis for the response endpoint is that PF-06882961 is superior to placebo as shown by an odds ratio > 1. As above, each dose of PF-06882961 will be tested separately compared to placebo with no adjustment for multiple comparisons.

5.2. General Methods

The efficacy analyses will be based on the appropriate population for analysis (see Section 4).

5.2.1. Summary Analyses for Continuous Data

Continuous variables will be presented using summary statistics: number of observations, arithmetic mean, standard deviation, median, minimum and maximum values.

5.2.2. Summary Analyses for Categorical Data

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 *will include treatment, time, strata* (defined as *metformin vs. diet and exercise alone* CCI [redacted] *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 2-sided p-values will be obtained for each treatment at each time point.

Differences in LSMeans between each dose of PF-06882961 relative to placebo at each time point, together with 90% confidence intervals, standard errors and 2-sided p-values, will also be obtained.

Example SAS code is provided in Appendix 4.

Statistical Model Diagnostics

[redacted]

The presence of outliers will be investigated for this model. 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.

5.2.4. Logistic Regression

The logistic regression model will include a term for treatment, strata (metformin vs. diet and exercise alone), and baseline will be included as a covariate. The model will be fitted to the Week 16 timepoint only.

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 dose of PF-06882961 relative to placebo and corresponding 90% confidence interval will be obtained.

Example SAS code is provided in Appendix 4.

5.2.5. Emax Model (no baseline interaction)

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

The model structure will take the form:

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

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 LSMean results from the primary MMRM model (Section 6.1.1.1) utilizing a Bayesian methodology approach with weakly informative priors as described in Appendix 3.

[redacted]

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 (5th and 95th percentiles of the relevant posterior distribution) will be reported for each target randomized dose (including Placebo) and their differences relative to placebo. 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; (3) otherwise an Emax model (no baseline interaction) 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 change from baseline (CFB) dose-response relationships, assuming an interaction between baseline HbA1c and dose.

The model structure will be the same as Section 5.2.5.

The model will be applied to the LSMean results from individual MMRM models that have been applied to each treatment separately. The MMRM models will include time and strata 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 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 LSmean results from the separate MMRM models will be calculated for the overall baseline HbA1c (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 3.

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 (5th and 95th percentiles of the relevant posterior distribution) will be reported for each target randomized dose (including Placebo)

[redacted]

and their differences relative to placebo. 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; (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 separately and will be plotted on the same graph up to Week 18 (i.e. including the follow-up visit 10). 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 will be considered.

Details of censoring are included in Section 6 and example SAS code is provided in Appendix 4.

5.3. Methods to Manage Missing Data

Details of efficacy data to be censored is described in Sections 3.2 and 3.3.

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

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).

5.3.1. Multiple Imputation

For summarizing the proportion of responses in HbA1c [redacted], all data from Estimand Set 2 [redacted] will be included (i.e. all time points up to and including Week 16). 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 (HbA1c [redacted]) and strata. Twenty sets of imputations of each missing value will be constructed from the MI method and the proportion of responses by treatment will be determined with associated standard errors utilizing a normal approximation and will be combined using standard multiple imputation techniques proposed by Rubin (2) 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.

[redacted]

For logistic regression of HbA1c [redacted], all data from Estimand Set 2 [redacted] will be included (i.e. all time points up to and including Week 16). 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 dose relative to placebo will be combined using standard multiple imputation techniques proposed by Rubin (2) 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.

6.1. Primary Endpoint(s)

6.1.1. Change from Baseline in HbA1c at Week 16

6.1.1.1. Main Analysis

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

Absolute values and changes from baseline in HbA1c will be summarised descriptively by treatment and time point as described in Section 5.2.1. Tables will present all data from the screening (visit 1, absolute tables only), beginning of the placebo run-in (visit 2, absolute tables only), baseline and post-baseline time points (including follow-up, which will be restricted to participants who completed 16 weeks of treatment and did not initiate glycemic rescue medication). Box and whisker plots of absolute values and changes from baseline will also be separately produced.

The primary analysis will be an MMRM (as described in Section 5.2.3) applied to the change from baseline at Weeks 2, 4, 6, 8, 12 and 16 that will be used to estimate the treatment effect related to the primary Estimand 1A (as described in Section 2.1.1).

The following results from the above primary analysis will be plotted:

- Profile plots of the LSMeans (including 90% confidence intervals) over time, with a separate line for each treatment
- Profile plots of the LSMean differences to Placebo (including 90% confidence intervals) over time, with a separate line for each dose of PF-06882961
- Plot of the LSMeans (including 90% confidence intervals) at Week 16, with a separate point for each treatment (applicable to analyses related to HbA1c only)

[redacted]

- Plot of the LSMean differences to Placebo (including 90% confidence intervals) at Week 16, with a separate point for each dose of PF-06882961 (applicable to analyses related to HbA1c 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.

In addition, a waterfall plot of the individual changes from baseline in HbA1c at Week 16 in the Estimand Set 1A will be created, ordered in increasing reductions from baseline and coloured by treatment.

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 HbA1c (without a baseline interaction) an Emax model will be applied to the LSmeans at Week 16 from the main MMRM analysis as described in Section 5.2.5.

- To assume a dose-response relationship with HbA1c, including an interaction between baseline HbA1c and dose, an Emax model will be applied to the LSMeans at Week 16 from separate MMRM models at the overall baseline HbA1c as described in Section 5.2.6.

CCI [redacted]

CCI [redacted]

6.2. Secondary Endpoint(s)

6.2.1. Response as defined by an HbA1c <7% at Week 16

6.2.1.1. Main Analysis

In all cases the Estimand Set 2 as specified in Section 4 will be utilised.

The proportion of responses at Week 16 will be summarised descriptively by treatment as described in Section 5.2.2, where no imputation for missing values will be conducted. The proportion of responses at Week 16 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 Week 16 that will be used to estimate the treatment effect related to the secondary Estimand 2 (as described in Section 2.1.2.1), where missing values will be imputed using multiple imputation as per Section 5.3. Baseline HbA1c 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. Change from Baseline in HbA1c at Weeks 2, 4, 6, 8 and 12

Summaries and analysis results will be reported as per the primary analysis approach (Section 6.1.1).

6.2.3. Change from Baseline in Fasting Plasma Glucose at Weeks 2, 4, 6, 8, 12 and 16

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

Absolute values and changes from baseline in fasting plasma glucose will be summarised similar to HbA1c in Section 6.1.1.1.

The analysis will be an MMRM (as described in Section 5.2.3) applied to the change from baseline at Weeks 2, 4, 6, 8, 12 and 16 that will be used to estimate the treatment effect related to the primary Estimand 3 (as described in Section 2.1.1). Results similar to HbA1c in Section 6.1.1.1 (excluding the waterfall plot) will be reported along with standard SAS output (where the latter will not be included in the CSR).

[redacted]

6.2.4. Change from Baseline in Body Weight at Weeks 2, 4, 6, 8, 12 and 16

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

Absolute values and changes from baseline in body weight will be summarised similar to HbA1c in Section 6.1.1.1.

The analysis will be an MMRM (as described in Section 5.2.3) applied to the change from baseline at Weeks 2, 4, 6, 8, 12 and 16 that will be used to estimate the treatment effect related to the primary Estimand 4 (as described in Section 2.1.1). Results similar to HbA1c in Section 6.1.1.1 (including the waterfall plot for changes from baseline at Week 16) will be reported along with standard SAS output (where the latter will not be included in the CSR).

6.2.4.1. Supplementary Analyses

The percentage change from baseline in body weight at Weeks 2, 4, 6, 8, 12 and 16 will similarly be summarised and analysed to absolute changes as per Section 6.2.4.

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 give geometric LSMeans and ratios of geometric LSMeans.

The percent change will then be calculated as follows:

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

CCI [redacted]

CCI

[redacted]

CCI

[redacted]

[redacted]

6.5. Baseline and Other Summaries and Analyses

6.5.1. Baseline Summaries

A baseline table (or separate tables, as required) summarizing the following will be produced by treatment and overall: age; gender; race; ethnicity; height; weight; body mass index; duration of T2DM; stratification factor; country; HbA1c; fasting plasma glucose; [redacted] eGFR; systolic blood pressure; diastolic blood pressure; and duration of metformin.

This table will also be re-produced limited to each strata. The table restricted to participants on a background of metformin only will not include the strata summary. The table restricted to diet and exercise will not include the metformin dose and strata summary.

6.5.2. Study Conduct and Participant Disposition

Participant evaluation groups will show participant disposition for each phase of the study (screening, placebo-run in, double-blind treatment and follow-up) and will additionally show which participants were analysed for efficacy (Estimand Set 1A, [redacted] and 2) as well as for safety. Frequency counts and percentages will be supplied for participant discontinuation(s) by treatment and overall.

Data will be reported in accordance with the sponsor reporting standards (including creation of COVID-19 related outputs).

[redacted]

CCI [redacted]

[redacted]

6.5.4. 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.

A separate table listing the use of rescue medication will be produced according to current sponsor reporting standards. The use of rescue medication (grouped by class of medication) will be summarized descriptively by treatment and overall as described in Section 5.2.2. The classes will be defined based on medications used in the study.

6.5.5. Treatment Compliance

A summary table of treatment compliance (by treatment and overall) will be produced according to current sponsor reporting standards.

6.5.6. Discontinuations

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

Data will be reported in accordance with the sponsor reporting standards.

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.

6.6. Safety Summaries and Analyses

The three AEs of interest are: diarrhoea, nausea and vomiting (based on preferred term).

6.6.1. Adverse Events

Adverse events (Tier 1, 2 and 3 AEs as described in Section 3.5.1) will be summarized by treatment 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 and overall.

[redacted]

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. Number of participants and percent will be presented, along with the risk difference between each dose of PF-06882961 and placebo. 95% confidence intervals and 2-sided 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 dose of PF-06882961 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 2-sided 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 dose of PF-06882961.

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 PF-06882961 versus placebo."

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.

6.6.1.1. Subset reporting interval

The subset adverse events, described in Section 3.5.1.1, 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'.

6.6.1.2. Hypoglycemic Adverse Events

The hypoglycemic AEs will be listed in a separate table and summarized categorically by treatment and overall as per Section 5.2.2.

[redacted]

6.6.1.3. 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, discontinue from IP or initiate glycemic rescue medication prior to the start of the AE event of interest will be censored at the discontinuation/initiation date.

The three AEs of interest are: diarrhoea, nausea and vomiting (based on preferred term). A separate plot for each AE will be produced separately.

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

6.6.1.4. Percentage of Participants with Ongoing AEs of Interest

Exploratory summaries on the percentage of participants reporting AEs of interest will be produced by week and treatment and overall as per Section 5.2.2 from Week 1 up to Week 18 (i.e. including the follow-up visit 10).

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 line plot of the percentages over each week will be produced based on the summary statistics with a separate line for each treatment.

The AEs of interest are diarrhoea, nausea and vomiting (based on preferred term); where a separate table and plot will be produced for each AE.

6.6.1.5. 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.

6.6.2. Laboratory Data

Laboratory data from will be listed and summarized by treatment and overall, in accordance with the sponsor reporting standards. Baseline is as defined in Section 3.5.3.

6.6.2.1. 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 3.5.3.1) will be summarized by treatment and time point as per Section 5.2.1 (unplanned readings will not be included in these summaries). Follow-up data will be included in the summaries with data from all participants in the safety analysis set. Tables will be paged by parameter. Baseline is as defined in Section 3.5.3.

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

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

MMRM models (as described in Section 5.2.3) will be applied to the change from baseline at Weeks 2 (only for liver function tests and eGFR), 4, 6 (only for liver function tests and eGFR), 8, 12 and 16 for each parameter separately using the safety analysis set (as defined in Section 4). From each model, the LSMeans together with 90% confidence intervals, standard errors and 2-sided p-values will be obtained for each treatment at each time point. Differences in LSMeans between each dose of PF-06882961 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 dose of PF-06882961 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.

Additionally, MMRM models will be applied to the changes from baseline in Triglycerides, ALT, AST, Alkaline Phosphatase and GGT for each parameter separately on the natural log scale at the same above relevant weeks using the safety analysis set. This parameter list may be expanded after review of unblinded tables if MMRM diagnostics reveal any other residuals that deviate substantially from a normal distribution and require log transformation for interpretation. In the models, the related baseline will be included on the natural log scale also. Similar outputs will be reported as above, where the percent change from baseline or percent difference to placebo (as applicable) will be reported as calculated using:

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

6.6.2.2. 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 3.5.3.2 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 (i.e. post-dose), as per Section 5.2.2.

6.6.3. 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 and time point, according to sponsor reporting standards. Tables will be paged by parameter. Baseline is as defined in Section 3.5.3.2.

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 with all treatments 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 (0 and 2H separately), 6, 8 (0 and 2H separately), 12 and 16 (0 and 2H separately) for supine systolic and diastolic blood pressure and pulse rate separately using the safety analysis set (as defined in Section 4). From each model, the LSMeans together with 90% confidence intervals, standard errors and 2-sided p-values will be obtained for each treatment at each time point. Differences in LSMeans between each dose of PF-06882961 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 dose of PF-06882961 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 using categories as defined in the Appendix 5. 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, according to sponsor reporting standards.

6.6.4. 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 and time point using sponsor reporting standards. Tables will be paged by parameter. Baseline is as defined in Section 3.5.5.

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 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 (0 and 2H separately), 6, 8 (0 and 2H separately), 12 and 16 (0 and 2H 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 2-sided p-values will be obtained for each treatment at each time point. Differences in LSMeans between each dose of PF-06882961 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 dose of PF-06882961 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 using categories as defined in the Appendix 5 (for QTc these correspond to the Pfizer Guidance (3) as referenced in Section 8). Numbers and percentages of participants meeting the categorical criteria will be provided. All planned and unplanned postdose 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 according to sponsor reporting standards.

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

CCI [redacted]

7. INTERIM ANALYSES

7.1. Introduction

An interim analysis will be performed at least once while the study is ongoing, after at least 25% of participants (ie, approximately 100) are randomized, with further details provided in the IRC charter. This interim analysis will assess, at a minimum, unblinded safety of the

[redacted]

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, stopping for futility, stopping for early success, or conducting a sample size reestimation. 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 standard operating procedure (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 the 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. Rubin DB. Multiple Imputation for Nonresponse in Surveys. 1987; Wiley, New York.

3. 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.

4. 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

5. 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.

6. 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.

7. 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.

9. APPENDICES

[redacted]

Appendix 1. Summary of Efficacy Analyses

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Change from baseline at Week 16 in HbA1c	Summary	Estimand Set 1A	GRM: Censor all values post first use DFI: Censor all values post discontinuation MV: Not imputed	N/A
	Primary analysis			MMRM
	Supplementary analyses			Emax
	CCI [redacted]	[redacted]	[redacted]	[redacted]
Response as defined by an HbA1c <7% at Week 16	Summary	Estimand Set 2	GRM: Censor all values post first use DFI: Censor all values post discontinuation MV: Not imputed	N/A
	Summary		GRM: Censor all values post first use DFI: Censor all values post discontinuation MV: Imputed using multiple imputation	N/A
	Secondary analysis			Logistic regression
Change from baseline in HbA1c at Weeks 2, 4, 6, 8 and 12	Summary	Estimand Set 1A	GRM: Censor all values post first use DFI: Censor all values post discontinuation MV: Not imputed	N/A
	Secondary analysis			MMRM
Change from baseline in fasting plasma glucose at Weeks 2, 4, 6, 8, 12 and 16	Summary	Estimand Set 1A	GRM: Censor all values post first use DFI: Censor all values post discontinuation MV: Not imputed	N/A
	Secondary analysis			MMRM
Change from baseline in body weight at Weeks 2, 4, 6, 8, 12 and 16	Summary	Estimand Set 1A	GRM: Censor all values post first use DFI: Censor all values post discontinuation MV: Not imputed	N/A
	Secondary analysis			MMRM

[redacted]

%Change from baseline in body weight at Weeks 2, 4, 6, 8, 12 and 16	Supplementary summary Supplementary analysis	Not applicable	GRM: Censor all values post first use DFI: Censor all values post discontinuation MV: Not imputed	N/A MMRM
CCI				

Abbreviations: DFI = discontinuation from IP; GRM = glycemic rescue medication; MV = missing values (note this includes missing values produced through censoring).

Appendix 2. Data Set Descriptions

To explore the data further and to assess the goodness-of-fit of all statistical models, separate SAS datasets and .CSV files will be provided by the clinical programmer to the statistician.

The dataset requirements for the Dataset (Snapshot PD Dataset, SNAP_PD) are specified in the table below. Note, the variable names and labels are suggested labels only and the actual names and labels and code levels should be consistent with current sponsor reporting standards:

Suggested Variable Name	Suggested Variable Label	Specifications	Order in the dataset
USUBJID	Unique Subject ID		1
Treattxt	Treatment Label		2
Leg_sort	Treatment Code		3
Dose	Dose (mg)		4
Site	Site		5
Age	Age	Age	6
Agecat	Age Category	1 = <18. 2 = 18 – 44, 3 = 45 – 64, 4 = >=65	7
Gender	Gender	1 = Female 2 = Male	8
Height	Height	Height	9
Weight_Screen	Weight (Screening)	Weight (Screening value)	10
Race	Race	Race	11
Strata	Strata	1 = Background Metformin 0 = Diet & Exercise	12
Time point	Time point (weeks)	Time point (in Weeks)	13
Rescue	Glycemic Rescue Medication	1 = Started using glycemic rescue medication 0 = Not started glycemic rescue medication	14
HbA1c	HbA1c	HbA1c value	15
HbA1c_bas	Baseline (HbA1c)	Baseline for HbA1c	16
HbA1c_Resp	Responder (HbA1c)	1 = Response 0 = Non-response (Only valid for Week 16, otherwise missing values)	17
FPG	Fasting Plasma Glucose	Fasting Plasma Glucose value	18
FPG_bas	Baseline (Fasting Plasma Glucose)	Baseline for Fasting Plasma Glucose	19
Weight	Body Weight	Body Weight value	20
Weight_bas	Baseline (Body Weight)	Baseline for Body Weight	21

Suggested Variable Name	Suggested Variable Label	Specifications	Order in the dataset
CCI [redacted]	[redacted]	[redacted]	[redacted]

The dataset(s) should contain each of the primary and secondary efficacy variables and the format can be amended as long as each variable is included along with the respective baseline variables.

Appendix 3. 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
2.5	-0.2	0.4
10	-0.4	0.5
40	-0.8	0.3
80	-1.0	0.4
120	-1.2	0.5

The residual standard deviation at Week 16 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 (4). 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*=120 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
CCI [redacted]	[redacted]

[redacted]

CCI [redacted]

In addition, the projected ED_{50} is CCI [redacted] 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: 5, 6 & 7), 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 Week 16:
CCI [redacted] Provided by programming
mmrmRes$N <- trunc((mmrm_sd/mmrmRes$StdErr)^2,0)

# Set-up priors and MCMC options:
prior_mmrm <- emaxPrior.contro CCI [redacted]
mcmc_mmrm <- mcmc.control(chains=3,thin=20,seed=169)

### Run Emax model: ###
```

[redacted]

```
emaxMMRM <-
fitEmaxB(mmrmRes$Estimate,mmrmRes$dose,prior_mmrm,modType=4,count=mmrmRes$N,msSat=mmrm_s
d^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 Week 16 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=mmrmDoseRes$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.

[redacted]

Appendix 4. 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', 'PF-06882961 2.5 mg BID', 'PF-06882961 10 mg BID',… 'PF-06882961 120 mg BID', so that when sorted in descending order 'Placebo' comes first.

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

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

* Perform multiple imputation;
proc mi data=analysisw seed=169 nimpute=20 out= analysis_mi;
        class treatment strata;
```

[redacted]

```
        fcs nbiter=10 reg(/details);
        var base treatment strata week2 week4 week6 week8 week12 week16;
run;

* Determine responders and non-responders at Week 16 for all imputed datasets;
data analysis_mi_16;
        set analysis_mi;
        value = week16 + base;
        if value < 7 and week16 ne . then resp = 1;
        if value >= 7 and week16 ne . then resp = 0;
run;

* Create datasets and combine proportions;
proc freq data=analysis_mi_16;
        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_16;
        by _imputation_ descending treatment descending resp;
proc logistic data = analysis_mi_16 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 = 'treatment 2.5 mg BID vs Placebo' then treatment = 2.5;
        if Effect = 'treatment 10 mg BID vs Placebo' then treatment = 10;
        if Effect = 'treatment 40 mg BID vs Placebo' then treatment = 40;
        if Effect = 'treatment 80 mg BID vs Placebo' then treatment = 80;
        if Effect = 'treatment 120 mg BID vs Placebo' then treatment = 120;
        if treatment = . then delete;
        estimate = log(OddsRatioEst);
        SE = (log(UpperCL) - log(LowerCL))/(2*QUANTILE('NORMAL',0.95));
run;

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

Appendix 5. 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	

[redacted]

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 6. List of Abbreviations

Abbreviation	Term
AE	adverse event
ALT	alanine aminotransferase
AST	aspartate aminotransferase
BLQ	below the limit of quantitation
BOCF	baseline observation carried forward
CFB	change from baseline
CRF	case report form
CSR	clinical study report
CV	coefficient of variation
ECG	electrocardiogram
eGFR	estimated glomerular filtration rate
Free T4	free thyroxine
GGT	gamma-glutamyl transferase
HbA1c	hemoglobin A1C
HDL	high-density lipoprotein
CCI [redacted]	[redacted]
ICD	informed consent document

[redacted]

Abbreviation	Term
IP	investigational product
LDL	low-density lipoprotein
LLQ	lower limit of quantitation
LSMean	least-squares mean
MedDRA	Medical Dictionary for Regulatory Activities
MCMC	Markov Chain Monte–Carlo
MI	multiple imputation
MMRM	mixed model repeated measures
PCC	potential clinical concern
CC	[redacted]
PR	pulse rate
PT	preferred term
QC	quality control
QQ	quantile-quantile
QTcF	corrected QT (Fridericia method)
SAE	serious adverse event
SAP	statistical analysis plan
SOA	schedule of activities
SOP	standard operating procedure
T2DM	type 2 diabetes mellitus
TEAE	treatment emergent AE
TSH	thyroid-stimulating hormone
ULN	upper limit of normal

[redacted]