Protocol C1171002

A PHASE 2A, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, DOSE-RANGING, PARALLEL GROUP STUDY TO EVALUATE SAFETY, TOLERABILITY, AND PHARMACODYNAMICS OF PF-05221304 ADMINISTERED DAILY FOR 16-WEEKS TO ADULT SUBJECTS WITH NONALCOHOLIC FATTY LIVER DISEASE

Statistical Analysis Plan (SAP)

Version: 2.0 (Amendment 1)

Author : PPD (PPD Cambridge USA)

Date: 14-MAR-2019

[redacted]

Revision History

Version	Date	Author(s)	Summary of Changes/Comments
Version 1.0	02-NOV-2017	PPD	NA
Version 2.0	14-MAR-2019	PPD	Refer to Section 1 for summary of changes

[redacted]

Distribution list of reviewers for this document prior to final sign-off.

Clinical Team* - List of reviewers	
Role Name	**Reviewer's Name**
PPD	

* : All members of the clinical team are blinded

[redacted]

Internal Review Committee† - List of reviewers	
Role Name	**Reviewer's Name**
PPD	PPD (change became effective on 01May2018)
	PPD
	PPD

† : All members of the Internal Review Committee (IRC) are unblinded to safety data, only

Unblinded Programmer† - reviewer	
Role Name	**Reviewer's Name**
PPD	

NOTE: *Italicized* text within this document has been taken verbatim from the Protocol Amendment 2 (dated 03Oct2017).

[redacted]

TABLE OF CONTENTS

LIST OF TABLES 7
LIST OF FIGURES 7
1 VERSION HISTORY 10
2 INTRODUCTION 11
2.1 Study Objectives 12
2.2 Study Design 13
3 ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS 15
3.1 Primary Endpoint(s) 15
3.2 Secondary Endpoint(s) 15
3.3 Tertiary Endpoints 15
3.4 Safety Endpoints 16
3.4.1 Adverse Events 16
3.4.2 Laboratory Data 17
3.4.3 Standard Safety Other than AEs 19
3.5 Baseline Variables 20
4 ANALYSIS SETS 20
4.1 Full Analysis Set 20
4.2 Per Protocol Analysis Set 21
4.3 Safety Analysis Set 21
4.4 Treatment Misallocations 21
4.5 Stratum Reassignment 22
4.6 Protocol Deviations 22
4.6.1 Deviations Assessed Prior to Randomization 22
4.6.2 Deviations Assessed Post-Randomization 22
4.7 Other Analysis Sets 23
4.7.1 Concentration Analysis Set 23
5 GENERAL METHODOLOGY AND CONVENTIONS 23
5.1 Hypotheses 23
5.2 General Methods 23
5.2.1 Analyses for Continuous Data 23

[redacted]

5.2.1.1 Descriptive Statistics 23
5.2.1.2 Longitudinal Analysis 24
5.2.1.3 Supportive Efficacy Analyses 25
5.2.1.4 Covariate Interaction 26
5.2.2 Analyses for Categorical Data 26
5.3 Methods to Manage Missing Data 26
6 ANALYSES AND SUMMARIES 27
6.1 Primary Endpoint 27
6.1.1 Primary Analysis on % Liver Fat assessed using MRI-PDFF 27
6.1.2 Supportive Analyses on % Liver Fat Assessed using MRI-PDFF 29
6.1.2.1 Dose-Response Model 29
6.1.2.2 Covariate-by-treatment interaction 29
6.2 Secondary Endpoint(s) 29
6.2.1 Alanine Aminotransferase (ALT) 29
6.2.2 Supportive Analyses on Alanine Amino Transferase 30
6.2.2.1 Dose-Response Model 30
6.3 Other Endpoint(s) 31
CCI [redacted]
6.3.2 Tertiary Endpoints 32
6.3.2.1 Continuous Endpoints Analyzed, Over Time, on Log Scale in Subjects With Nonalcoholic Fatty Liver Disease (Entire Study Population) 32
6.3.2.2 Continuous Endpoints Analyzed, Over Time, on Log Scale in Subjects With Diagnosed/Presumed Nonalcoholic Steatohepatitis only 32
CCI [redacted]
6.3.2.4 Continuous Endpoints Analyzed on Raw Scale in Subjects With Nonalcoholic Fatty Liver Disease (Entire Study Population) 34

[redacted]

6.3.2.5 Continuous Endpoints Analyzed, Over Time, on Raw Scale in Subjects With T2DM Only 35
CCI [redacted]
6.4 Subset Analyses 37
6.5 Other Summaries and Analyses 37
6.5.1 Standard Analyses 37
6.5.2 Baseline Summaries 38
6.5.3 Concomitant Medications and Non-Drug Treatments 38
6.6 Safety Summaries and Analyses 38
6.6.1 Adverse Events 38
6.6.2 Additional Non Standard Safety-Related Analyses to be performed: 40
6.6.2.1 Platelet count 40
6.6.2.2 Estimated Glomerular Filtration Rate 41
7 INTERIM ANALYSES 42
7.1 Introduction 42
7.2 Interim Analyses and Summaries 42
8 REFERENCES 44
9 APPENDICES 45

LIST OF TABLES

Table 1 Summary of Major Changes in SAP Amendments 10
Table 2 Clinical Laboratory Parameters of Interest 17
Table 3 Vital Signs to be Monitored 19
CCI [redacted]

LIST OF FIGURES

Figure 1. C1171002 Study Design 14

[redacted]

Figure 2 Two-Tiered Stratification in Study C1171002 37

APPENDICES

Appendix 1. SUMMARY OF EFFICACY ANALYSES 45
Appendix 2. DATA DERIVATION DETAILS 47
Appendix 2.1. Definition and Use of Visit Windows in Reporting 47
Appendix 2.2. Data Derivation for Primary, Secondary, and Tertiary Endpoints 48
Appendix 2.2.1. % Liver Fat using MRI-PDFF 48
CCI
Appendix 2.2.4. Liver Function Tests 51
CCI
Appendix 2.2.6. Lipid Panel 55
Appendix 2.2.7. Glycated Hemoglobin 56
CCI
Appendix 2.2.9. Fasting Plasma Glucose and Fasting Plasma Insulin (T2DM Subjects Only) 58
Appendix 2.2.10. Homeostatic Model Assessment for Insulin Resistance (T2DM Subjects Only) 59
CCI
Appendix 2.2.12. Serum Apolipoprotein 61
CCI
Appendix 2.2.14. Markers of Apoptotic Activity - CK18-M30 and CK18-M65 63
Appendix 2.2.15. Collagen formation-related Biomarkers - Pro-C3 and Pro-C6 65
CCI
Appendix 2.3. Data Derivation for Non Standard Safety Endpoints 71
CCI
Appendix 2.3.3. Alcohol Intake Assessed via AUDIT Questionnaire 74

[redacted]

Appendix 2.3.4. % Carbohydrate-Deficient Transferrin (%CDT) 75
Appendix 2.4. Compliance 75
Appendix 3. GRAPHICAL REPRESENTATIONS..... 77
Appendix 3.1. Box and Whisker Plot Representation for Changes/Percent Changes from Baseline at Week 16 by Randomized Arm..... 77
Appendix 3.2. Box and Whisker Plot for Changes/ Percent Changes From Baseline by Randomized Arm and Week 78
Appendix 3.3. Figure for MMRM LS Mean vs Time by Randomized Arm for Changes/ Percent Changes From Baseline Over Time 79
Appendix 3.4. Figure for MMRM LS Mean vs Time by Randomized Arm for Changes/ Percent Changes From Baseline Placebo-Adjusted Over Time 80
Appendix 3.5. Figure for MMRM LS means for Changes/ Percent Changes From Baseline by Randomized Arm at Week16 overlaid with Emax Estimates (and 80%CI) 81
Appendix 3.6. Figure for MMRM LS means for Placebo-Adjusted Changes/ Percent Changes From Baseline by Randomized Arm at Week16 82
Appendix 3.7. Additional Figure requested..... 83

[redacted]

1 VERSION HISTORY

C1171002 protocol Amendment 2 (version: 03-Oct-2017) remains effective for this amended version of the SAP.

The original SAP is amended to correct typographic errors, to reflect a number of clarifications, to aid in more clear derivation and reporting of endpoints, and types of analyses. Major changes from the original SAP are summarized in Table 1.

Table 1 Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
1	Not Applicable	Not Applicable
2	a. Section 3.4.2: Definition of permitted to be nonfasted and mandatory fasted laboratory/ pharmacodynamics parameter	a. Fasting status impacts numerical results
	b. Creation of Section 4.5: Stratum Reassignment	b. Instances have been encountered wherein subjects were randomized to incorrect strata, in error. Section is describing how to handle such cases during reporting
	c. Section 5.3: LLQ and ULQ reporting rules	**c.** Clarification on how to derive results which are either below LLQ and above ULQ – specifically for pharmacodynamic parameters
	CCI [redacted]	
	f. Section: 6.4: Revisions made to subset analysis of subjects randomized in Taiwan	**f.** Considering the limited number of subjects randomized across sites in Taiwan (n=6), generation of descriptive summary was deemed non-informative as at most each arm will have less than 2 subjects as randomization is applied at a country-level; as such, supplemental outputs restricted to subjects from Taiwan will not be generated (approach endorsed by

[redacted]

		Taiwan-FDA on 01Nov2018 based on input sought by Pfizer-Taiwan)
	g. Adding Total Bilirubin as 1 of the liver function tests	**g.** Throughout the SAP, Total Bilirubin (TBILI) added as part of the Liver Function Test group; omitted originally in error
	h. Appendix 2.4: Compliance calculation	**h.** Compliance calculation revised to utilize details from oral dosing CRF page (and not drug accountability CRF page) as the former accounted for any missed or dropped doses (where as the drug accountability page was strictly based on pill count returned) – thereby permitting a more accurate assessment of compliance via use of details captured on oral dosing CRF page for "active bottles" only
	i. Appendix 2.2.6: Additional safety figures	**i.** Addition of set of figure reporting triglycerides, ApoC3, FPI and FPG levels over time for subjects presenting at any time post-randomization with fasting serum triglycerides level ≥ 600 mg/dL
	j. HOMA-IR, CCI [redacted] are unit less	**j.** Appendices 2.2.10, CCI [redacted]: Indicate respectively that HOMA-IR, CCI [redacted] re unit-less parameters
	k. Adding summary statistics for %CDT and AUDIT questionnaire	k. Appendix 2.3.3 and 2.3.4: Both parameters added because they were omitted originally in error.
	l. Appendix 2.31: Additional safety figures	l. Appendix 2.3.1: Addition of set of figure reporting platelet count, INR and aPTT over time for subject presenting at any time post-randomization platelet count lower than 100,000/mm^3 (ie, 100 x 10^3/mm^3)

2 INTRODUCTION

Acetyl-CoA carboxylase (ACC) is a biotin carboxylase that catalyzes the adenosine triphosphate (ATP) dependent condensation of acetyl-CoA and carbonate to form malonyl-CoA. The malonyl-CoA produced by ACC serves two major physiologic functions. It is an essential and rate-limiting substrate for de novo lipogenesis (DNL) and it acts as an allosteric inhibitor of the enzyme carnitine-palmitoyltransferase 1 (CPT-1). CPT-1 is responsible for the transport of long-chain fatty acyl-CoAs across the mitochondrial membrane into the mitochondria where they become available for fatty acid oxidation. The transport step is rate-determining for this process. Thus, ACC is positioned as a key physiologic switch regulating the transition from oxidative to lipogenic metabolism. Reducing malonyl-CoA production via ACC inhibition is expected to inhibit simultaneously the de novo synthesis of fatty acids and to increase flux through CPT-1

[redacted]

leading to increased β-oxidation of long-chain fatty acids. There are two closely related isoforms of ACC, ACC1, and ACC2, encoded by separate gene products that differ in tissue and cellular distribution.

PF-05221304 is a potent, selective, orally bioavailable, and reversible dual ACC1/2 inhibitor designed to have asymmetric distribution to the liver, with ≥ 100 fold asymmetric hepatic distribution demonstrated in both rats and monkeys; as such PF-05221304 is expected to inhibit DNL and stimulate fatty acid oxidation in the liver to a greater extent than in peripheral tissues. In humans, administration of PF-05221304 has been shown to suppress hepatic DNL in a Phase 1 trial in healthy subjects; in addition, the drug is expected to stimulate hepatic fatty acid oxidation, and consequently reduce fat accumulation in the liver. This inhibition of hepatic DNL is postulated to result in a decrease and normalization of the excessive DNL observed in nonalcoholic fatty liver disease (NAFLD). In addition, the inhibition of ACC via administration of PF-05221304 also has the potential for anti-inflammatory effects in nonalcoholic steatohepatitis (NASH).

The current study is the first clinical trial proposed with PF-05221304 in adult subjects with NAFLD. The investigational product will be administered orally, once-daily for up to 16 weeks. It is designed as a dose-ranging trial with placebo and 4 active doses of PF-05221304 to assess the safety, tolerability and the effect of PF-05221304 on liver fat. In addition, assessment of the effect of PF-05221304 on other pharmacodynamics/exploratory parameters is planned.

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study C1171002. 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

Primary Objective:
• *To evaluate the dose-response for the effect on **liver fat***, as assessed via MRI-PDFF, *with a range of PF-05221304 doses administered daily in adults **with nonalcoholic fatty liver disease (entire study population)***
Secondary Objective(s):
• *To evaluate the dose-response for the effect on **ALT** with a range of PF-05221304 doses administered daily in the 1st tier stratification comprising of adults **with diagnosed/presumed nonalcoholic steatohepatitis, only***
• *To evaluate the safety and tolerability of a range of PF-05221304 doses administered daily in adults with nonalcoholic fatty liver disease (entire study population)*
Tertiary Objective(s):
• *To evaluate the dose-response for the effect on liver fat, **over time**, with a range of PF-05221304 doses administered daily in adults **with nonalcoholic fatty liver disease (entire study population)***

[redacted]

CCI

- *To evaluate the dose-response for the effect on key liver function tests and NASH-related biomarkers **over time**, with a range of PF-05221304 doses administered daily in the 1st tier stratification comprising of adults **with diagnosed/presumed nonalcoholic steatohepatitis, only***

CCI

- *To evaluate the dose-response for the effect on other potentially mechanism-related parameters and metabolic parameters, **over time**, with a range of PF-05221304 doses administered daily in adults with **nonalcoholic fatty liver disease (entire study population)***

- *To evaluate the dose-response for the effect on glycemic parameters, **over time**, with a range of PF-05221304 doses administered daily in the 2nd tier stratification comprising of **adults with T2DM, only***

CCI

2.2 Study Design

This is a randomized, double-blind, placebo-controlled, 5-arm (placebo, plus 4 active doses of PF-05221304), parallel-group study.

Determination of eligibility for this study will occur via a sequential, 3-step process – starting with pre-qualification. Subjects identified to be eligible based on pre-qualification will then undergo Visit 1 (Screen 1) with only those confirmed to continue to be eligible progressing to Visit 2 (Screen 2). For a given subject, this 3-step eligibility assessment (from pre-qualification to Screen 2, inclusive) may take up to 2-months.

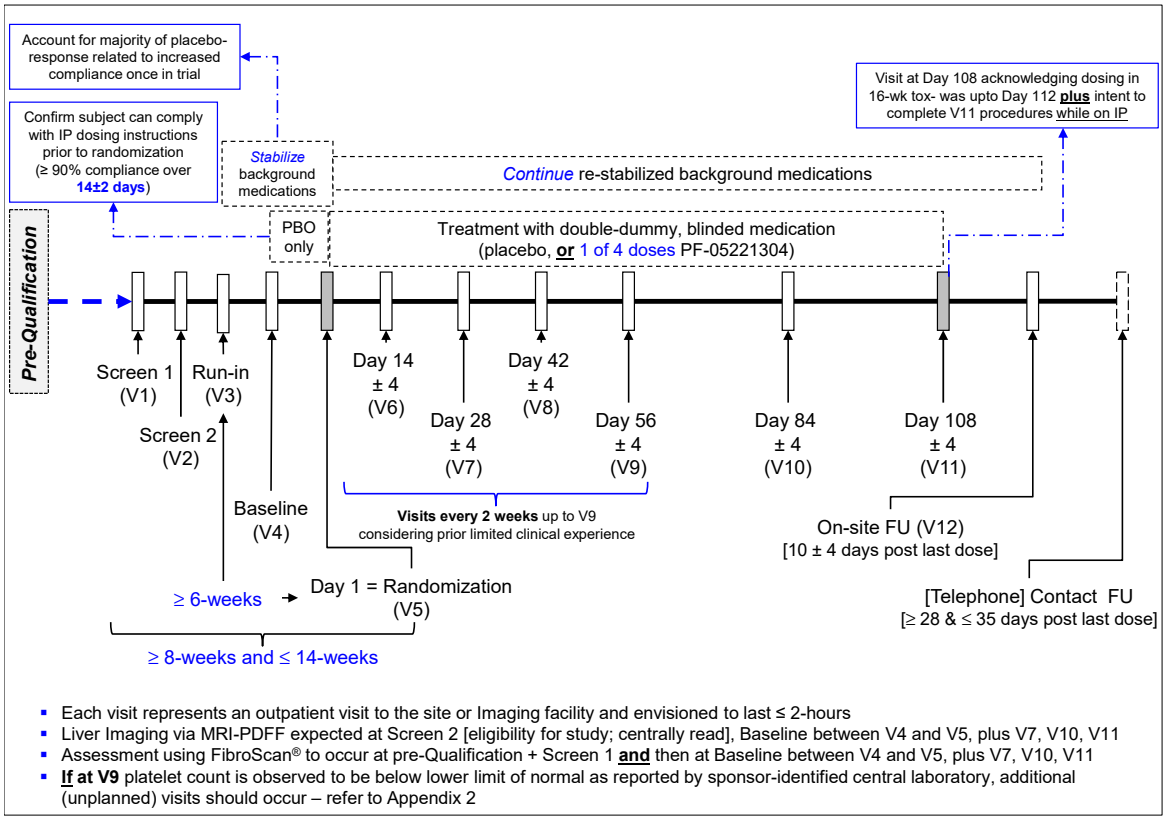
*Once confirmed to be eligible based on Screen 2 procedure(s), subjects will progress to a run-in phase when concomitant treatment(s) will be stabilized, and baseline-related visits will occur. At Visit 5 (Day 1), subjects will be randomized to receive 1 of 5 blinded regimens for a duration of up to 16 weeks (ie, 108±4 days). Excluding the pre-qualification visit, this study includes a total of 12 scheduled outpatient visits **in the***

[redacted]

morning, *3 random visits for PK blood draws,* ***and*** *a potential for up to at least 3 visits for repeat assessments (for procedures/tests at Visit 1, Visit 2, or 2nd Follow-up visit). The total participation, from Visit 1 (Screen 1) to 2nd Follow-up visit, will be 27-weeks (minimum) to 35-weeks (maximum) – refer to Figure 1.*

At least 360 subjects (72 per arm) will be randomized at approximately 120 sites to ensure a minimum of 300 subjects (60 per arm) complete the study. ***In each arm****, ≥ 50% of subjects randomized will be deemed to be eligible for 1st tier stratification of diagnosed/presumed NASH; with the remaining subjects per arm eligible for 1st tier stratification of NAFLD with likely minimal inflammation and fibrosis* (*refer to Section 4.3* of the protocol for more details).

Figure 1. C1171002 Study Design



Subjects who are noted to have non-evaluable baseline MRI-PDFF, as determined by the sponsor-identified central imaging vendor, may be replaced, at sponsor discretion (*refer to Section 7.1.4* of the protocol for more details); *otherwise, there are no plans to replace subjects who are prematurely withdrawn. To ensure that the required minimum number of subjects complete the study, the plan is to overenroll by approximate 20% to account for premature early withdrawal.*

[redacted]

3 ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1 Primary Endpoint(s)

- *Percent change from baseline in liver fat, as assessed using MRI-PDFF, at Week 16 in subjects* ***with nonalcoholic fatty liver disease (entire study population)***

3.2 Secondary Endpoint(s)

- *Percent change from baseline in ALT, at Week 16 in adults* ***with diagnosed/presumed nonalcoholic steatohepatitis, only***

- *Assessment of TEAEs, safety-related clinical laboratory tests, vital signs, and 12-lead ECGs in adults* ***with nonalcoholic fatty liver disease (entire study population)*** – refer to Section 3.4, below

3.3 Tertiary Endpoints

- *Percent change from baseline in liver fat, as assessed using MRI-PDFF, over time up to Week 16 in adults* ***with nonalcoholic fatty liver disease (entire study population)***

CCI [redacted]

- *In adults* ***with diagnosed/presumed nonalcoholic steatohepatitis only***, p*ercent change from baseline, over time up to Week 16, for*:
  - ✓ *ALT, AST, Alkaline Phosphatase (ALKP), GGT*, Total Bilirubin (TBILI)
  - ✓ *CK18-M30 and CK18-M65*

    CCI [redacted]
  - ✓ *Pro-C3 & Pro-C6*

CCI [redacted]

[redacted]

CCI [redacted]

- *In adults **with nonalcoholic fatty liver disease (entire study population)**, percent change from baseline, over time up to Week 16, for:*
  - Potentially mechanism-related parameters –
    - ✓ *Serum apolipoprotein A1, B (total), C3, E*

    CCI [redacted]
  - *Metabolic parameters*
    - ✓ *Fasting lipid panel (fasting total cholesterol, direct LDL-C, HDL-C, triglycerides, VLDL)*

    CCI [redacted]
  - Change from baseline (**not percent change from baseline as mentioned in protocol**), *over time up to Week 16 for* Glycated hemoglobin (*HbA1c*)

CCI [redacted]

- *In adults **with T2DM, only***
  - *Change from baseline, over time up to Week 16 for*
    - ✓ Glycated hemoglobin (*HbA1c*)
    - ✓ Fasting plasma glucose (*FPG*)
    - ✓ Fasting plasma insulin (*FPI*)
    - ✓ *HOMA-IR*

CCI [redacted]

3.4 Safety Endpoints

3.4.1 Adverse Events

An adverse event will be considered ***treatment emergent*** (ie, TEAEs) relative to a given treatment if:

- event occurs for the first time during the effective duration of treatment and was not seen prior to 1st dose of double-blinded study medication on Day 1 [Visit 5], for example, during the internal between subjects offered informed consent for pre-qualification up to day prior to 1st dose, or
- event was seen prior to the start of dosing on Day 1 [Visit 5] but increased in severity during treatment.
- ***NOTE:*** Adverse events deemed non-treatment emergent observed during the interval between subjects offering informed consent for pre-qualification up to 1st dose on Day 1 (Visit 5) but resolved before 1st dose on Day 1 or not increasing in

[redacted]

severity post dosing on Day 1 will be listed and summarized under the category of "Run-In".

The effective duration of treatment is determined by the lag time defined as the Pfizer Standard of 999 days post last dose of double-blinded study medication. Any **treatment-emergent** event occurring within the lag time, whether it occurs during a break in treatment or at the end of treatment, will be attributed to the assigned randomized regimen.

A 3-tier approach will be used to summarize TEAEs. Under this approach, TEAEs will be classified into 1 of 3 tiers. Different analyses will be performed for different tiers (See Section 6.6.1).

Tier-1 events: These are pre-specified events of clinical importance (TMEs, DMEs, and CPTs) and are maintained in a list in the product's Safety Review Plan (version dated 01Sep2017 [GDMS link - http://gdms.pfizer.com/gdms/drl/objectId/090177e18cd3d394])

Tier-2 events: These are events that are not tier-1 but are "common". A MedDRA preferred term (PT) is defined as a tier-2 event if there are at least 4 subjects with at least one occurrence in any treatment group, per Pfizer Standard to distinguish Tier-2 events from Tier-3 events.

Tier-3 events: These are events that do not meet criteria for either tier-1 nor tier-2 events

3.4.2 Laboratory Data

For the specific laboratory parameters listed in Table 2 below, the following endpoints will be evaluated using CaPS:

- Absolute value and change from baseline at week 2, week 4, week 6, week 8, week 12, week 16, and Follow-up.
- Number of occurrences of these abnormalities defined as "Flag Level" or "Alert Level"
- Number of subjects with these abnormalities defined as "Flag Level" or "Alert Level"

Table 2 Clinical Laboratory Parameters of Interest

Parameter	***Flag Level***	***Alert Level***	***Conventional Units***
Fasting Serum Triglycerides	*≥ 400*	*≥ 600*	*mg/dL*
	--	*≥ 800*	*mg/dL*
Platelet Count	*< 100*	*< 75*	*x 10^3 /mm^3*
Fasting Plasma Glucose	*< 70*	*≤ 49*	*mg/dL*
	≥ 140	>270	mg/dL
Alanine aminotransferase	*≥ 2x ULN*	*Pfizer std flag for PCC*	*IU/L*
Aspartate aminotransferase	*≥ 2x ULN*	*Pfizer std flag for PCC*	*IU/L*

[redacted]

Alkaline Phosphatase	*≥ 2x ULN*	*Pfizer std flag for PCC*	*IU/L*
Gamma Glutamyl Transferase	*> ULN*	*Pfizer std flag for PCC*	*IU/L*
Total Bilirubin	*> 1.5x ULN*	*Pfizer std flag for PCC*	*mg/dL*
Direct (Conjugated) Bilirubin	*> ULN*	*Pfizer std flag for PCC*	*mg/dL*

**All flag level changes are cumulative from baseline (defined as result closest prior to dosing at Visit 5 (Day 1); reflect either threshold for entry criteria into study or clinically significant thresholds*
^All alert changes are cumulative from baseline (defined as result closest prior to dosing at Visit 5 (Day 1); values when noted during the study, by central laboratory necessitates rapid notification (via fax/e-mail) to site and Study Clinician
PCC – potential clinical concern
ULN – upper limit of normal as determined by the central laboratory

The safety events of special clinical concern as defined in Table 2 will be summarized and the frequency of laboratory abnormalities (both number of occurrence **and** number of subjects).

Across all the laboratory-related analytes (for safety and pharmacodynamics/biomarker), inclusion of results when assessed at scheduled nominal visits meet the following criteria:-

1. Collection must be prior to AM dose of double-blinded study medication
2. Collection permitted to be *nonfasted for the following analytes and will be permitted to be grouped with fasted results* given that nonfasted state is not known to impact results

Chemistry panel	**Hematology**	**Other**	**Urinalysis**
Total bilirubin	Hemoglobin	aPTT	pH
Direct bilirubin	Hematocrit	PT	Protein
Indirect bilirubin	RBC	INR	Glucose
Alk Phos	MCV	Urine-drug-testing	Ketones
ALT	MCH	Serology testing	Bilirubin
AST	MCHC	[hepatitis B,	Urobilinogen
GGT	WBC	hepatitis C, HIV]	Blood
Urea Nitrogen	Neutrophils (abs)	Alpha-1-antitrypsin	Nitrite
Creatinine	Lymphocytes (abs)	Ceruloplasmin	Leukocyte esterase
Uric acid	Monocytes (abs)	HbA1C	Microscopic UA
Calcium	Eosinophils (abs)	CCI	
Total protein	Basophils (abs)	β-hCG	
Albumin	Platelets	FSH	
Creatine kinase	Reticulocyte count%		
Sodium	Reticulocyte count		
Potassium			
Bicarbonate			
Chloride			

3. ***But,*** the *following analytes will only be included in summary outputs if collections are confirmed to be following an overnight fast of at least 8 hours* as it is well known that fasting status impacts numerical results (***or*** there is insufficient data regarding effect of meals on results)

Biomarkers	**Lipid-related**	**Other Blood-related**

[redacted]

CCI Plasma Glucose Plasma Insulin CCI CK18-M30 CK18-M65 ProC3 ProC6	Total cholesterol HDL-C Direct LDL-C Triglycerides VLDL Apolipoprotein A1 Apolipoprotein B Apolipoprotein C3 Apolipoprotein E	CCI

3.4.3 Standard Safety Other than AEs

Individual subjects' vital signs outlined in Table 3 of special clinical concern, which align with the Pfizer standard criteria, will be summarized by randomized arm, as part of standard safety-related outputs.

Table 3 Vital Signs to be Monitored

Parameter	***Flag Level****	***Conventional Units***
Systolic Blood Pressure (seated)	*< 90*	*mm Hg*
	≥ 30 change from baseline	*mm Hg*
Diastolic Blood Pressure (seated)	*< 50*	*mm Hg*
	≥ 20 change from baseline	*mm Hg*
Pulse rate (seated)	*<40*	*bpm*
	> 120	*bpm*

**All flag level changes are cumulative from baseline (defined as result closest prior to dosing at Visit 5 (Day 1)*
bpm – beats per minute

In addition Pfizer standard criteria for reporting of cardiac conduction intervals will be summarized by randomized arm.

In addition, changes in background medications related to glycemic control, lipid control, or blood pressure will be summarized as follows, at any time after Day -14 and up to 1st Follow-up Visit (or Early Termination), whichever is first –

- Dose increase and decrease in concomitant medication (see list in Protocol section 5.8.1) for glycemic control,
- Dose increase and decrease in concomitant lipid-modifying medications (see list in Protocol section 5.8.2).
- Dose increase and decrease in concomitant blood pressure management medications (see list in Protocol section 5.8.3).

[redacted]

3.5 Baseline Variables

For each subject, for each of the following parameters, baseline will be defined as:

- Liver fat content, via MRI-PDFF, defined as the value of total mean liver fat obtained between Visit 4 and Visit 5 (ie, labeled as baseline)

CCI [redacted]

- Result closest prior to dosing on Day 1 [Visit 5] for the remaining pharmacodynamic endpoints **and all** safety-related continuous endpoints
  - As such would be results as reported on Day 1 [Visit 5], and **if** these are not reported would be the results reported on Day -14 [Visit 4]
  - If results on Day 1 [Visit 5] and on Day -14 [Visit 4] are missing the baseline will be considered as missing.

CCI [redacted]

- For Adverse events, refer to definition of treatment-emergent outlined in Section 3.4.1, above

Change from baseline (CBL), relative change from baseline (RCBL) and percent change from baseline (PBL) will be calculated as follows:

- CBL = Observed Value – Baseline Value
- RCBL = Observed Value / Baseline Value
- PBL = 100 × (Observed Value – Baseline Value) / Baseline Value

(NOTE: RCBL and PBL will not be considered for endpoints that can have a value of zero at baseline)

4 ANALYSIS SETS

Data for all subjects will be assessed to determine if subjects meet the criteria for inclusion in each analysis population set ***close (but prior to)*** database limit-write- access is achieved ***but prior to*** unblinding and releasing the database. Such classifications will be documented per standard operating procedures.

4.1 Full Analysis Set

This is the most inclusive subset of randomized subjects in a trial and is therefore compatible with Intention-to-Treat (ITT) analyses. This set includes all randomized

[redacted]

subjects who receive at least one dose of randomized study treatment (ie, on Day 1 [Visit 5]).

4.2 Per Protocol Analysis Set

The Per Protocol Analysis Set (PPAS) is a subset of the Full Analysis Set (FAS) excluding the following subjects **for the MMRM analysis on the primary efficacy endpoint**:

- Subjects with compliance, with randomized regimen administered (ie, relevant of the 3 bottles dispensed) over the entire treatment period, of less than 80%.
- Subjects with protocol deviation(s) deemed as compromising primary objective of the study (the list of subjects will be finalized ***close (but prior to)*** database limit-write- access is achieved ***but prior to*** unblinding and releasing the database).

For the Dose-Response analysis on the primary efficacy endpoint, PPAS is a subset of the FAS excluding the following subjects:

- Subjects with compliance, with randomized regimen administered (ie, relevant of the 3 bottles dispensed) over the entire treatment period, of less than 80%.
- Subjects with protocol deviation(s) deemed as compromising primary objective of the study (the list of subjects will be finalized ***close (but prior to)*** database limit-write- access is achieved ***but prior to*** unblinding and releasing the database).
- Subjects who withdraw during the randomized treatment period of the trial (post first dose on Day 1 [Visit 5] and before Week 16 [Visit 11]) and/or do not offer MRI (ie, not done or non-evaluable data) at Week 16 [Visit 11].

4.3 Safety Analysis Set

The safety analysis set is all subjects who receive at least one dose of randomized study treatment (ie, on Day 1 [Visit 5]).

4.4 Treatment Misallocations

In order to assess treatment misallocations drug dosing information will be collected. Treatment misallocations will be handled as follows. If a subject is:

Randomized but not treated, then the subject will be excluded from all the statistical analyses.

Randomized but took incorrect treatment, then the subject will be reported under his/her randomized group for all analyses that utilize the FAS and excluded from all statistical analyses that utilize the PPAS. The subject will be reported for safety and PK analysis under his/her actual randomized arm.

Treated but not randomized, then by definition they will be excluded from the efficacy analyses since randomized arm is missing, but will be reported under the dosing regimen they actually received for all safety [redacted] analyses.

[redacted]

4.5 Stratum Reassignment

Subjects who are randomized to the wrong stratum, in error, will have the incorrect stratum assignment remain in IMPALA but database (OCRDC) corrected based on following handling conventions:

Randomized but not treated, then the subject will be excluded from all the statistical analyses.

Randomized, treated, and took correct treatment, then the subject will be reassigned to and reported under the stratum he/she was supposed to be randomized to (and for which eligibility criteria are met) for all analyses (ie safety, CCI and PD analysis) that utilize the FAS and the PPAS.

Randomized, treated but took incorrect treatment, then the subject will be reassigned to and reported under the stratum he/she was supposed to be randomized to for all analyses (ie safety, CCI and PD analysis) that utilize the FAS and excluded from all statistical analyses that utilize the PPAS.

4.6 Protocol Deviations

The following sections describe any protocol deviations that relate to the statistical analyses.

It is possible that unexpected deviations will arise, becoming known only after the study has been active; hence more deviations may be added. A full list of protocol deviations for the study report will be compiled prior to database release.

4.6.1 Deviations Assessed Prior to Randomization

Granted exceptions to the inclusion or exclusion criteria are not expected to occur. Any subject who enters the study when the inclusion or exclusion criteria should have prevented entry will be considered to have had a protocol deviation. Whether such subjects will be excluded from FAS or PPAS will be determined as outlined in Sections 4.1 to 4.3, above.

4.6.2 Deviations Assessed Post-Randomization

All protocol deviations post-randomization will be reviewed and those that are thought to affect the primary objective/endpoint will be considered to define FAS and PPAS.

Examples of the protocol deviations assessed post randomization to assess for potential impact on the primary objective/endpoint include:

- subjects who receive excluded concomitant medications
- subjects whose background medications are changed in violation of protocol

- subjects who were randomized but took incorrect dosing regimen
- subjects whose overall (ie, over 16-weeks of dosing) compliance is below 80%

4.7 Other Analysis Sets

CCI [redacted]

5 GENERAL METHODOLOGY AND CONVENTIONS

5.1 Hypotheses

Statistical inference will be made on the primary endpoint: percent change from baseline at Week 16 (Visit 11) in % liver fat assessed using MRI-PDFF. The null hypothesis is that there is no difference between PF-05221304 and placebo on the primary endpoint. The alternative hypothesis is that PF-05221304 is superior to placebo on the primary endpoint.

Mathematically, the null and alternative hypotheses are expressed as follows.

$$H_0 : \mu_i - \mu_0 = 0$$

vs.

$$Ha : \mu_i - \mu_0 < 0 \ ,$$

where μ_i is the mean percent change from baseline for PF-05221304 doses of 2 mg QD, 10 mg QD, 25 mg QD and 50 mg QD, and μ_0 is the mean percent change from baseline for placebo.

Statistical inferences for the primary endpoint will be done at a one-sided significance level of 10%. A dose is considered efficacious if the lower 90% lower confidence limit of $\mu_1 - \mu_0$ exceeds 0.

No Adjustment will be made for multiple comparisons.

5.2 General Methods

5.2.1 Analyses for Continuous Data

5.2.1.1 Descriptive Statistics

Each endpoint will be summarized accordingly to Appendix 2.

[redacted]

5.2.1.2 Longitudinal Analysis

The over time effect of PF-05221304 on primary, secondary, and tertiary endpoints (see list in section 3) will be analyzed using the MMRM approach with randomized arm, time and randomized arm by-time interaction as fixed effects, subject as random effect using unstructured (UN) covariance structure. Baseline (in raw scale or log scale depending on the endpoint) will be included as covariate in the model.
Other covariance structures (eg Compound Symmetry (CS), Autoregressive(1) [AR(1)], Heterogeneous CS (CSH), and Heterogeneous AR(1) [ARH(1)]) will be considered if convergence in the previous model is not met.

The effect of PF-05221304 doses at all **post-randomization** visits will be estimated by the least-square (LS) means as well as LS mean differences against placebo respectively from the model described above.

For all analyses performed on log-transformed relative change from baseline (RCBL), the LS mean estimates and their corresponding confidence intervals (CIs) will be obtained for each randomized arm and week. Differences (and corresponding confidence intervals) between LS mean estimates will be obtained comparing randomized arms for each week. All LS means and LS mean differences (including CI's) will be back transformed to provide adjusted geometric LS means and ratios of adjusted geometric LS means. Then, percent change from baseline will be calculated as follows:

Percent change = 100* (RCBL - 1)

where RCBL is either the adjusted geometric LS means estimate or the ratio of adjusted geometric means coming from the statistical model. The corresponding 80% CIs will be calculated as well.

For all analyses performed on raw scale, the LS means estimates and their corresponding confidence intervals (CIs) will be obtained for each randomized arm and week. Differences (and corresponding confidence intervals) between LS means will be obtained comparing randomized arms for each week.

Regardless if the data are analyzed in raw scale or after log-transformation, the inter- and intra-subject variability from the statistical model will be available and reported.

For % liver fat assessed by MRI-PDFF analyzed on log-scale, 80% confidence intervals will be presented for the LS geometric mean ratio.

For endpoints which are on raw scale, 80% confidence intervals will be presented for the LS mean differences along with the corresponding two-sided *p*-values when changes from baseline is the primary parameter.

[redacted]

For other endpoints that, similar to % liver fat assessed by MRI-PDFF, are log-transformed, 80% confidence intervals will be presented for the LS geometric mean ratio.

The LS means and corresponding standard errors will be plotted over time by randomized arms.

If, for any endpoint, there are major deviations from statistical assumptions underlying the MMRM model then alternative transformations or non-parametric analyses may be presented. Justification for any alternative to the planned analysis will be given in the study report.

5.2.1.3 Supportive Efficacy Analyses

The effect of PF-05221304 doses on % Liver Fat assessed with MRI-PDFF [redacted] will be performed **post-database release** using the following dose-response model on the individual natural log transformed of relative change from baseline observed at Week 16.

Non-linear Dose-Response Model

The 3 parameter Emax model will be fit to the response at Week 16 on all doses. If the model converges, model estimated parameters will be presented along with their standard error (SE), 80% (Liver Fat [redacted]) confidence intervals as well as p-values (one-sided). The model projected placebo adjusted treatment effect for each PF-05221304 dose will also be presented along with their SE and confidence intervals. The Delta-Method will be employed to estimate standard errors and confidence intervals.

If the model fails to converge, the message "Algorithm fails to converge" will be displayed.

The 3-parameter Emax model is expressed as follows:

$$y_{ij} = E_0 + E_{\max} \frac{d_i}{d_i + ED_{50}} + \varepsilon_{ij}$$

, where y_{ij} denotes the response for subject j on dose i, d_i denotes dose i, and $\varepsilon_{ij} \sim$ i.i.d $N(0, \sigma^2)$.

Alternative Dose-Response Model in case of no convergence

Selection of alternative dose-response model (eg linear, quadratic) may be explored after visual inspection of the response at Week 16. The rationale of the model selection and the subsequent analyses may be included in the CSR.

[redacted]

5.2.1.4 Covariate Interaction

The impact of the covariates will be examined to identify potential parameters that could influence the primary endpoint (ie % liver fat using MRI-PDFF).

CCI [redacted]

For the categorical covariate, an ANCOVA model which includes the randomized arm, covariate and the randomized arm-by-covariate interaction as independent factors, and log(baseline) value as an additional covariate will be used. If the *p*-value for the interaction is less than 0.05, the LS means and LS mean differences will be displayed by each level (absence = 0 or presence = 1) of the categorical covariate.

For the continuous variable (eg weight, baseline), an ANCOVA model which includes the randomized arm, covariate and the randomized arm-by-covariate interaction as independent variables, and log(baseline) value as an additional covariate will be used If the *p*-value for the interaction is less than 0.05, the LS means and LS mean differences will be displayed at Q1, median, and Q3 levels of the continuous covariate.

5.2.2 Analyses for Categorical Data

Categorical data will be summarized by presenting the number of subjects and the percentage.

The Wilcoxon rank sum statistic will be used to assess the randomized arm difference in the distribution (median) of the number of hypoglycemic events per subject.

The Wilcoxon rank sum statistic will be used to assess the randomized arm difference in the distribution (median) of the number of thrombocytopenia events per subject.

The Wilcoxon rank sum statistic will be used to assess the randomized arm difference in the distribution (median) of the number of hypertriglyceridemia events per subject.

Hypoglycemic risk ratio, thrombocytopenia risk ratio, and hypertriglyceridemia risk ratio may be estimated by analysis methods based on a counting process approach (Andersen and Gill, 1982) for analyzing recurrent events (using the robust sandwich estimate of the covariance matrix by Lin and Wei (1989) in order to take into account the non-independence of multiple events for a single subject).

5.3 Methods to Manage Missing Data

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

CCI [redacted]

[redacted]

Pharmacodynamic parameters

- Missing data for ***all*** the pharmacodynamic parameters will be treated as such and no imputed values will be derived when presenting descriptive statistics at scheduled assessments.
- Concentrations below Lower Limit of Quantification (LLQ) will be set to half-LLQ (ie LLQ/2) for summary outputs; in listings BLQ values will be reported as "<LLQ". Derived data by calculation (eg indirect bilirubin, choriogonadotropin Beta) will not follow the rule described above and will be reported as it is (ie reported as "< LLQ" in listing and missing if summary statistics are required).
- Derived data by calculation CCI will not follow the rule described above and will be reported as <LLQ (in listings) with a derived value of (<negative value> minus 0.0001 = positive value) when generating summary statistics, as required.
- Any values above the Upper Limit of Quantification (ULQ) will be assigned to the ULQ for display purposes in Figures and for computation of summary statistics; in listings ULQ values will be reported as ">ULQ".
- Observed cases will be used when analyzing continuous variables using Mixed-Model Repeated Measures (MMRM)[1], analysis of covariance (ANCOVA) and nonlinear models.

6 ANALYSES AND SUMMARIES

6.1 Primary Endpoint

Primary endpoint will be analyzed on the FAS and PPAS. These analyses will be based on observed cases.

6.1.1 Primary Analysis on % Liver Fat assessed using MRI-PDFF

Population: Entire Study Population

Endpoint: Relative change from baseline

Parameters: % Liver fat assessed with MRI-PDFF:

- Analysis time points: **all** applicable post Day 1/Visit 5 time points where scheduled assessment of specific continuous parameter is outlined in protocol
- Analysis population**:** FAS and PPAS (as sensitivity analysis).
- Analysis methodology**:** Log transformed relative change from baseline will be analyzed using MMRM model with log(baseline) as a covariate (specified in section 5.2.1.2).

Reporting results:

- Raw data and percent changes from baseline: Randomized subject, number of subject contributing to summary, sample size, mean, standard deviation, median, minimum, maximum, Q1 and Q3 at baseline and post-baseline visits will be presented for each randomized arm (see Appendix 2 for further detail).
- MMRM Outputs:
  - ✓ Adjusted geometric LS mean estimates with 80% confidence interval for each randomized arm along with the translation in percent change over time,
  - ✓ Ratio of adjusted geometric LS mean estimates for each active randomized arm against placebo and the corresponding 80% confidence interval along with the translation in percent difference over time.

Figures from MMRM outputs:

1. Line plot of percent change from baseline after back transformation of adjusted geometric LS means with 80% confidence interval over time (see Appendix 3.3),
2. Line plot of percent change from baseline placebo-adjusted after back transformation of adjusted geometric LS means ratios with 80% confidence interval over time (see Appendix 3.4),
3. Percent change from baseline after back transformation of adjusted geometric LS means with 80% confidence interval at Week 16 (see Appendix 3.5),
4. Percent change from baseline placebo-adjusted after back transformation of adjusted geometric LS means ratios with 80% confidence at Week 16 (see Appendix 3.6).

Note regarding the figures;

Instead of producing 4 independent figures, and in order to having a better understanding of the placebo effect, 1 figure containing side-by-side two panels presenting bullet 1 and 2 outputs (right panel for response over time and left panel for placebo adjusted response over time), and 1 figure with two panels presenting side-by-side bullet 3 and 4 outputs (right panel for response at Week 16 and left panel for placebo adjusted response at Week 16) should be produced.

Note regarding MRI windowing;

Subjects with MRI assessment up to a maximum of 48Hrs post last dose on Week 16 instead of 24Hrs post last dose on Week 16 as currently stated will be included in FAS and PPAS. This is to permit assessment of the primary endpoint while on drug.

[redacted]

6.1.2 Supportive Analyses on % Liver Fat Assessed using MRI-PDFF

6.1.2.1 Dose-Response Model

For % Liver Fat assessed by MRI-PDFF at Week 16, the analysis will be performed on the FAS without covariates in subject with nonalcoholic fatty liver disease (entire study population). The primary endpoint is the log transformed relative change from baseline in liver fat at Week 16 using the dose response model as described in section 5.2.1.3.

Endpoint: Relative change from baseline

Parameter: % Liver fat assessed by MRI-PDFF

- Analysis time points Week 16
- Analysis population**:** FAS in subject with nonalcoholic fatty liver disease (entire study population)
- Analysis methodology**:** Natural log-transformed relative change from baseline in % liver fat will be analyzed using the Dose-Response model (specified in section 5.2.1.3).
- Supporting objective: Primary Objective

Reporting results for model:

- The three parameters of the Emax model (E0, ED50, and Emax) along with their respective 80% confidence interval if non-linear model. .

6.1.2.2 Covariate-by-treatment interaction

Covariate-by-Treatment interaction will be performed on % Liver Fat at Week 16 on FAS and PPAS (observed cases) as described in section 5.2.1.4.

6.2 Secondary Endpoint(s)

The secondary endpoint will be analyzed on the FAS. All analyses will be based on observed cases.

6.2.1 Alanine Aminotransferase (ALT)

Population: Subjects with diagnosed/presumed nonalcoholic steatohepatitis, only

Endpoint: Relative change from baseline

Parameters: ALT:

- Analysis time points: **all** applicable post Day 1/Visit 5 time points where scheduled assessment of specific continuous parameter is outlined in protocol)
- Analysis population**:** FAS.

[redacted]

- Analysis methodology**:** Log transformed relative change from baseline will be analyzed using MMRM model with log(baseline) as a covariate (specified in section 5.2.1.2).

Reporting results:

- Raw data and percent changes from baseline: Randomized subject, number of subject contributing to summary, sample size, mean, standard deviation, median, minimum, maximum, Q1 and Q3 at baseline and post-baseline visits will be presented for each randomized arm (see Appendix 2 for further detail).
- MMRM Outputs:
  - ✓ Adjusted geometric LS mean estimates with 80% confidence interval for each randomized arm along with the translation in percent change over time,
  - ✓ Ratio of adjusted geometric LS mean estimates for each active randomized arm against placebo and the corresponding 80% confidence interval along with the translation in percent difference over time.

Figures from MMRM outputs:

1. Line plot of percent change from baseline after back transformation of adjusted geometric LS means with 80% confidence interval over time (see Appendix 3.3).
2. Line plot of percent change from baseline placebo-adjusted after back transformation of adjusted geometric LS means ratios with 80% confidence interval over time (see Appendix 3.4).
3. Percent change from baseline after back transformation of adjusted geometric LS means with 80% confidence interval at Week 16 (see Appendix 3.5).
4. Percent change from baseline placebo-adjusted after back transformation of adjusted geometric LS means ratios with 80% confidence at Week 16 (see Appendix 3.6).

Note regarding the figures;

Instead of producing 4 independent figures, and in order to having a better understanding of the placebo effect, 1 figure containing side-by-side two panels presenting bullet 1 and 2 outputs (right panel for response over time and left panel for placebo adjusted response over time), and 1 figure with two panels presenting side-by-side bullet 3 and 4 outputs (right panel for response at Week 16 and left panel for placebo adjusted response at Week 16) should be produced. Both sets of figures should start X-axis at Day 1/Baseline.

CCI

CCI

CCI

CCI

CCI [redacted]

6.3.2 Tertiary Endpoints

All tertiary endpoints will be analyzed **on FAS**. These analyses will be based on observed cases.

6.3.2.1 Continuous Endpoints Analyzed, Over Time, on Log Scale in Subjects With Nonalcoholic Fatty Liver Disease (Entire Study Population)

Endpoint: Relative change from baseline

Parameters: CCI [redacted] fasting lipid panel (ie total cholesterol, LDL-C, HDL-C, VLDL, and TG), serum apolipoprotein (ApoA1, ApoB, ApoC3, and ApoE), CCI [redacted]:

- Analysis time points: **all** applicable post Day 1/Visit 5 time points where scheduled assessment of specific continuous parameter is outlined in protocol)
- Analysis population**:** FAS
- Analysis methodology**:** Log transformed relative change from baseline will be analyzed using MMRM model with log(baseline) as a covariate (specified in section 5.2.1.2).

Reporting results:

- Raw data and percent changes from baseline: Randomized subject, number of subject contributing to summary, sample size, mean, standard deviation, median, minimum, maximum, Q1 and Q3 at baseline and post-baseline visits will be presented for each randomized arm (see Appendix 2 for further detail).
- MMRM Outputs:
  - ✓ Adjusted geometric LS mean estimates with confidence interval for each randomized arm along with the translation in percent change over time,
  - ✓ Ratio of adjusted geometric LS mean estimates for each active randomized arm against placebo and the corresponding confidence interval along with the translation in percent difference over time.

6.3.2.2 Continuous Endpoints Analyzed, Over Time, on Log Scale in Subjects With Diagnosed/Presumed Nonalcoholic Steatohepatitis only

Endpoint: Relative change from baseline

Parameters: CCI [redacted] fasting liver function tests (ie AST, Alkaline Phosphatase, GGT, Total Bilirubin, CCI [redacted] CCI [redacted] cell death biomarkers (ie CK18-M30, and CK18-M65), fibrogenesis biomarkers (ie Pro-C3, and Pro-C6), CCI [redacted]:

[redacted].

- Analysis time points: **all** applicable post Day 1/Visit 5 time points where scheduled assessment of specific continuous parameter is outlined in protocol)
- Analysis population**:** FAS in subjects with diagnosed/presumed nonalcoholic steatohepatitis only
- Analysis methodology**:** Log transformed relative change from baseline will be analyzed using MMRM model with log(baseline) as a covariate (specified in section 5.2.1.2).

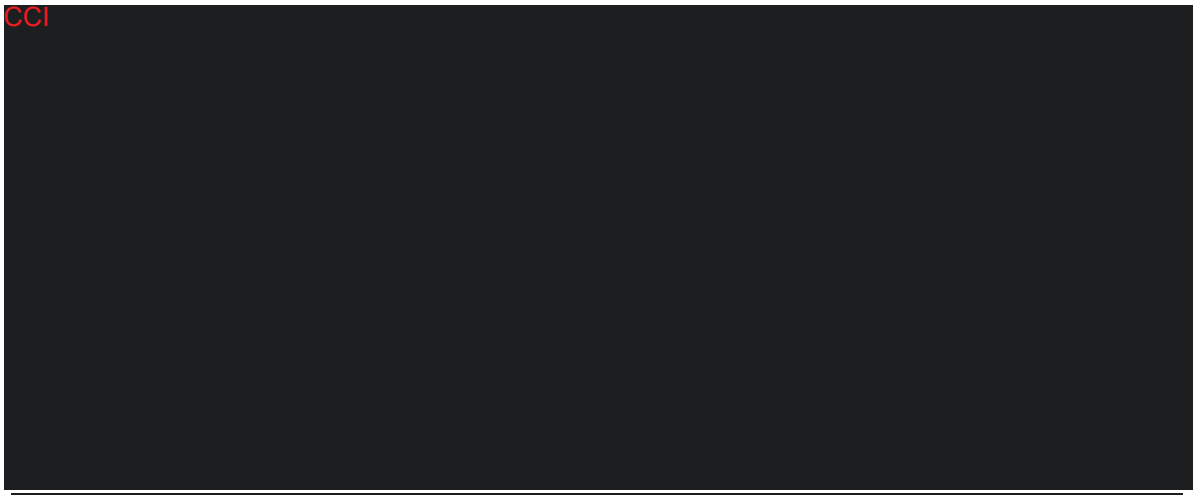
Reporting results:

- Raw data and percent changes from baseline: Randomized subject, number of subject contributing to summary, sample size, mean, standard deviation, median, minimum, maximum, Q1 and Q3 at baseline and post-baseline visits will be presented for each randomized arm (see Appendix 2 for further detail).
- MMRM Outputs:
  - ✓ Adjusted geometric LS mean estimates with confidence interval for each randomized arm along with the translation in percent change,
  - ✓ Ratio of adjusted geometric LS mean estimates for each active randomized arm against placebo and the corresponding confidence interval along with the translation in percent difference over time.

Figures from MMRM outputs

- Line plot of percent change from baseline after back transformation of adjusted geometric LS means with 80% confidence interval over time (see Appendix 3.3),
- Line plot of percent change from baseline placebo-adjusted after back transformation of adjusted geometric LS means ratios with 80% confidence interval over time (see Appendix 3.4),
- Vertical Bar chart of geometric LS means and 80% confidence interval at Week 16 (see Appendix 3.5).
- Vertical bar chart of placebo adjusted geometric LS means and 80% confidence interval at Week 16 (see Appendix 3.6).

CCI



[redacted].

CCI

[redacted]

CCI

[redacted]

6.3.2.4 Continuous Endpoints Analyzed on Raw Scale in Subjects With Nonalcoholic Fatty Liver Disease (Entire Study Population)

Endpoint: Change from baseline

Parameter: Glycated hemoglobin (HbA1c):

- Analysis time points : **all** applicable post Day 1/Visit 5 time points where scheduled assessment of specific continuous parameter is outlined in protocol),
- Analysis population**:** FAS in subjects with nonalcoholic fatty liver disease (entire study population),
- Analysis methodology**:** change from baseline will be analyzed using MMRM model (specified in section 5.2.1.2).

Reporting results:

- Raw data and changes from baseline: Randomized subject, number of subject contributing to summary, sample size, mean, standard deviation, median, minimum, maximum, Q1 and Q3 at baseline and post-baseline visits will be presented for each randomized arm (see Appendix 2 for further detail).
- MMRM Outputs:
  - ✓ LS mean estimates, 80% confidence interval for each randomized arm over time,
  - ✓ Differences between LS mean estimates for each active randomized arm against placebo and the corresponding 80% confidence interval over time.

[redacted].

6.3.2.5 Continuous Endpoints Analyzed, Over Time, on Raw Scale in Subjects With T2DM Only

Endpoint: Change from baseline

Parameters: Glycated hemoglobin (HbA1c), fasting plasma glucose (FPG), fasting plasma insulin (FPI), and HOMA-IR:

- Analysis time points: **all** applicable *post* Day 1/Visit 5 time points where scheduled assessment of specific continuous parameter is outlined in protocol)
- Analysis population**:** FAS in Subjects With T2DM only
- Analysis methodology**:** change from baseline will be analyzed using MMRM model (specified in section 5.2.1.2).

Reporting results:

- Raw data, and changes from baseline: randomized subject, number of subject contributing to summary, sample size, mean, standard deviation, median, minimum, maximum, Q1 and Q3 at baseline and post-baseline visits will be presented for each randomized arm randomized arm (see Appendix 2 for further detail).
- MMRM Outputs:
  - ✓ LS mean estimates, 80% confidence interval for each randomized arm over time,
  - ✓ Differences between LS mean estimates for each active randomized arm against placebo and the corresponding 80% confidence interval over time.

CCI

[redacted]

[redacted].

CCI

CCI

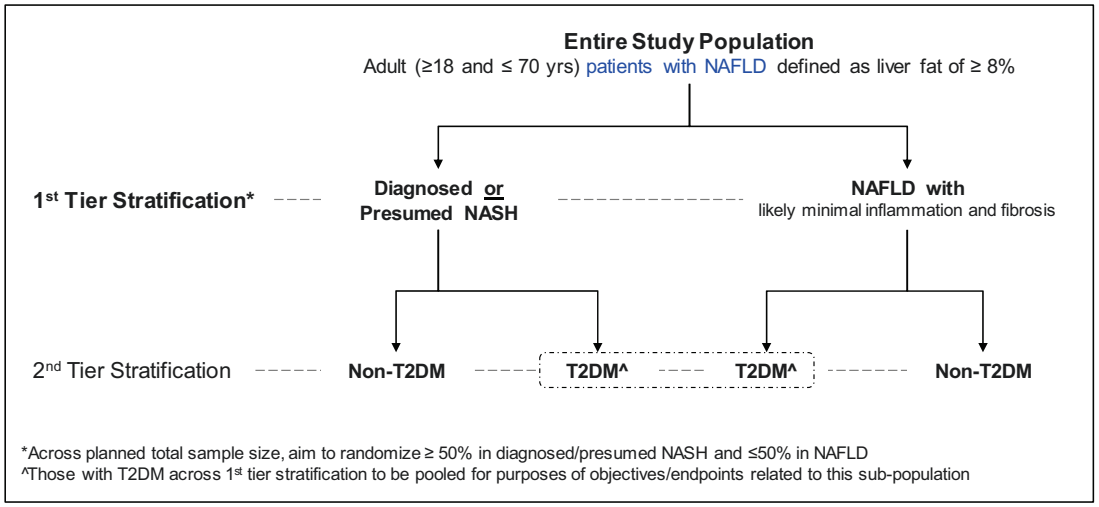
CCI

CCI

6.4 Subset Analyses

Beyond the analyses planned by 1st and 2nd tier strata (Figure 2), additional subset analyses are not planned.

Figure 2 Two-Tiered Stratification in Study C1171002



6.5 Other Summaries and Analyses

6.5.1 Standard Analyses

- Listings that account for all subjects who offered at least consent for pre-qualification and whether screen failed (at any point ***after*** the consent offered at the prequalification visit and ***before*** randomization on Day 1/Week0/Visit 5 will be presented.
- Study Conduct and Subject Disposition: The number of subjects randomized, treated, completing and discontinuing from the study, as well as the number of subjects in each analysis population will be summarized by randomized arm. The number of subjects randomized, treated, completing and discontinuing from the study, as well as the number of subjects in each analysis population will be summarized in each category of pre-defined medical history terms (ie, T2DM, Hyperglycemia, HDL increased, TG increased, hypertension, waist circumference

CCI

increased) by randomized arm. For subjects who did not complete the study, the reasons for withdrawal from the study will be presented.

- Baseline Characteristics: Demographic characteristics such as subject age, gender, height, weight, body mass index (BMI), race, by 1st tier stratification and total per arm, by treatment group will be tabulated.
- For each PD endpoint (ie primary, secondary or tertiary endpoint) and, as detailed in Appendix 2, baseline values will be listed and summarized by randomized arm and/or stratum and/or metabolic status (ie Type 2 Diabetes subjects vs non Type 2 Diabetes subjects). Descriptive statistics for these endpoints and derived endpoints will be also tabulated as detailed in Appendix 2.
- Overall compliance (Day 1 to Week 16 or premature termination during treatment phase) will be summarized by frequency (%) in the following two categories by randomized arm. Formula to compute compliance is defined in Appendix 2.4.
    - $< 80\%$
    - $\geq 80\%$ (and up to 110%)

6.5.2 Baseline Summaries

For each endpoint (ie primary, secondary or tertiary endpoint) and, as detailed in Appendix 2, baseline values will be listed and summarized by randomized arm and/or stratum and/or metabolic status (ie subjects with versus those without Type 2 Diabetes Mellitus).

6.5.3 Concomitant Medications and Non-Drug Treatments

Number of subjects taking, at any time post randomization, medication affecting % liver fat (eg Vitamin E) will be tabulated by randomized arm.

Number of subjects under T2DM medication (T2DM), lipid medication, hypertension medication (HTN) will be tabulated by randomized arm.

For each category of medication (ie T2DM, lipid and HTN), the number of subjects receiving 0, 1, 2, or 3 medications will also be tabulated by randomized arm.

List of remaining concomitant drugs not part of the above list will only be listed.

6.6 Safety Summaries and Analyses

6.6.1 Adverse Events

Unless otherwise noted, all safety data, including the following, will be summarized descriptively by randomized dose group through appropriate data tabulations and descriptive statistics:

[redacted]

- Tier-3 adverse events will be summarized according to Pfizer standards. Tier-1 and tier 2 events will also be included in these standard summary.
- Safety laboratory tests will be summarized according to Pfizer standards (note that actual visit window will be applied as defined in Appendix 2.1 in relation to randomized treatment).
- ECGs will be summarized by visits. Categorical summary of ECGs will be presented according to Pfizer standards (note that baseline is defined as in section 3.4 in relation to randomized treatment).
- Vital signs will be summarized according to Pfizer standards (note that baseline is defined as in Appendix 2. in relation to randomized treatment).
- For the safety endpoints of special clinical concern other than AEs as defined in section 3.4.2, they will be presented with descriptive statistics as appropriate for continuous or categorical data as defined in section 5.2.
- Hypoglycemic Adverse Event (HAE) for T2DM subjects will be characterized for each randomized arm of subjects with T2DM ***only***, as follows:
    - The number and percent of subjects reporting at least one hypoglycemic event
    - The hypoglycemic event rate, defined as the number of events reported per subject-month of treatment (per 100 subject months for severe events). Risk ratio of each active randomized arm against placebo will be displayed along with 80% CI and p-value.
    - The number of subjects with 0 event, 1 event, 2 events, 3 events, 4 events, and >=5 events.
    - The median number of events per subject. P-value of each active randomized arm as compared to placebo using rank-sum test will be displayed.
- Analysis of Tier-1 and Tier-2 TEAEs:
    - ✓ Unconditional exact test will be utilized. Risk difference will be used for the comparison.
    - ✓ No multiplicity adjustment will be made.
    - ✓ Tables and graphs will be generated. For tier-1 events, point estimates, 95% confidence intervals for the risk difference and p-values will be presented; for tier-2 events, point estimates and 95% confidence intervals for the risk difference will be presented.

[redacted]

- ✓ TEAEs, by preferred term, will be sorted in descending point estimate of risk difference.
- ✓ Pair-wise comparison will be made for each active randomized arm (each dose of PF-05221304) against placebo.
- ✓ For table/graphic output, include the following footnotes to provide proper interpretation of p-values and/or confidence intervals, describe how comparison is 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-05221304 xx mg 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. CCI [redacted]

6.6.2 Additional Non Standard Safety-Related Analyses to be performed:

CCI [redacted]

[redacted]

CCI

CCI

CCI

7 INTERIM ANALYSES

7.1 Introduction

*Interim analysis (IA) will be performed **at least once** while the study is on-going. This review will include an assessment of the observed safety, and observed (versus assumed) variability for the primary endpoint (liver fat via MRI-PDFF) as well as key secondary endpoint (ALT). The 1 planned IA is proposed after **at least one-third** of the total planned randomized subjects (ie, approximately 120 subjects), complete through **at least one-third** of the total dosing duration (ie, **Visit 8 [Week 6]**) of the study. Interim analysis results may be used for conducting a sample size re-estimation, stopping a dose-level for observed safety, CCI and internal business decisions regarding planning of future trials with PF-05221304.*

***In addition**, during study conduct, after approximately **every** 20% of the total planned randomizations [for example: 72 subjects (20%), 145 subjects (40%), 215 subjects (60%), and 290 subjects (81%) of planned 360 subjects], based on blinded safety review by the selected members of the Sponsor's study team, unblinded review of the safety data by the IRC will be triggered and a determination of the safety to continue the study documented if post randomization **either** of the below conditions are met –*

- *More than 20% of subjects, develop a moderate or severe AE in the same system organ class (SOC);*
- *More than 15 % of subjects meet the individual permanent discontinuation of dosing rules* as described in sections 3.2 of the protocol.

7.2 Interim Analyses and Summaries

For planned interim analysis and for triggered unblinded reviews, IRC members will receive the following data summarized/analyzed by randomized arm along with the relevant subject data listings.

Study conduct and subject characteristics:

- Study conduct issues (e.g., enrollment status, currentness of the database, time between randomization and first treatment, eligibility violations)
- Subject evaluation groups
- Subject stratification factors
- Demographic and other baseline characteristics

Safety Data:

The Committee will monitor certain specific safety-related laboratory parameters, vital sign parameters, and treatment-emergent adverse events (TEAEs) as well as serious adverse events (SAEs including death) listings in an unblinded manner for overall

[redacted]

frequency of occurrence of events. To do so, here are the listings which should be provided:

- TEAEs (and SAEs) leading to temporary discontinuation of dosing in subjects;
- TEAEs (and SAEs) leading to permanent discontinuation and withdrawal of subjects;
- The listing of individual subject-level TEAEs and SAEs;
- All instances of hypoglycemia, or hyperglycemia, or platelet reduction, or hypertriglyceridemia leading to permanent withdrawal;
- Listing of Treatment emergent adverse events (TEAEs) under the 3-tier approach as defined in section 6.6 of this SAP;
- Listings of SAEs/deaths and narratives requested from the safety database (ie, ARGUS or similar)

CCI laboratory parameters outlined in Table 2 of this SAP represent the minimum list of parameters which will be listed and provided to the committee. This list may be expanded depending on team blinded review.

The Committee will receive the distribution of subjects over time whose values cross the boundaries (ie Flag and Alert levels separately) for each of the parameters included in Table 2 of this SAP.

Percent change from baseline in other fasting lipid parameters including total cholesterol, direct low density lipoprotein cholesterol (LDL-C), very low density lipoprotein (VLDL) and high density lipoprotein cholesterol (HDL-C) will be also provided for review.

The following graphical representations of the parameters in Table 2, CCI other fasting lipid parameters (see list above) will be provided:

- Individual values: median response with Q1 and Q3 per randomized arm over time, Box and whiskers plots as depicted in Appendix 3.2,
- Changes from baseline: median response with Q1 and Q3 per randomized arm over time, Box and whiskers plots as depicted in Appendix 3.2,
- Percent change from baseline: median response with Q1 and Q3 per randomized arm over time, Box and whiskers plots as depicted in Appendix 3.2

Additionally, for T2DM subjects only, graphical representation of individual values and changes from baseline in fasting plasma glucose and fasting plasma insulin over time will be provided using median response with Q1 and Q3 over time and Appendix 3.2 template.

Individual subjects' vital signs (blood pressure and pulse rate) via a listing of the raw data as well as values meeting the criteria outlined in Table 3 of this SAP will be provided.

[redacted]

No formal statistical analyses (eg MMRM, rank test) will be performed on safety data.

Intra- and Inter-Subject Variability Determination of % Liver Fat Using MRI-PDFF and ALT variability determination:

At the time of a planned interim analysis, MRI-PDFF data up to and including Day 1 (Visit 5) for all randomized subjects will be analyzed using a mixed model with week as fixed factor and subject as random factor to permit an assessment of the observed intra- and inter-subject variability in order to benchmark the results against the assumed standard deviations utilized to derive the sample size for the study.

At the time of a planned interim analysis, all data available for ALT parameter in all randomized subjects in placebo arm will be included as part of the formal internal analysis and analyzed using a mixed model with subject as random factor to permit an assessment of the observed percent change from baseline variability in the placebo arm in order to permit an assessment of whether the observed standard deviation utilized to derive the sample size for the study are in line with the assumed standard deviation for this parameter.

8 REFERENCES

1. Siddiqui, Hung and O'Neil, "MMRM vs. LOCF: A Comprehensive Comparison Based on Simulation Study and 25 NDA Datasets", Journal of Biopharm. Stats, 19: 227-246, 2009.

2. Westfall and Krishen, "Optimally weighted, fixed sequence, and gatekeeping multiple testing procedures", Journal of Statistical Planning and Inference. 99, 25-40, 2001.

3. Hoeting, J.A, Davis R.A Merton A.A and Thompson S. E, "Model selection for Geostatistical Models", Ecological Applications, 16(1), 87–98, 2006.

4. Littell, R.C, Milliken, G A., Stroup, WW, and Wolfinger, R D., "SAS System for Mixed Models (second edition)" , Cary, NC: SAS Institute Inc, 2006.

5. Wagenknecht L.E., Palmer N.D., Bowden D.W., Rotter J.I., Norris J.M., Ziegler J., Chen Yii-Der I., Haffner S., Scherzinger A., and Langefeld C.D., "Association of PNPLA3 with non-alcoholic fatty liver disease in a minority cohort: the Insulin Resistance Atherosclerosis Family Study", Liver Int., 31(3), 412-416, 2011.

[redacted]

9 APPENDICES

Appendix 1. SUMMARY OF EFFICACY ANALYSES

Endpoint	Analysis Set	Statistical Method	Model/ Covariates/ Strata	Missing Data	Interpretation
Percent change from Baseline in % Liver Fat assessed by MRI-PDFF Over Time	FAS and PPAS	MMRM on log-transformed relative change	Treatment, time and treatment-by-week interaction as fixed effects, subject as random effect with log(baseline) as covariate.	No Imputation	Primary Analysis
Percent Change from Baseline in % Liver Fat assessed by MRI-PDFF at Week 16	FAS and PPAS	ANCOVA on log-transformed relative change	Treatment, Weight and treatment-by-Weight interaction as fixed effects, subject as random effect with log(baseline) as additional covariate.	No Imputation	Sensitivity analysis with two covariates on primary endpoint.
Percent Change from Baseline in % Liver Fat assessed by MRI-PDFF at Week 16	FAS and PPAS	ANCOVA on log-transformed relative change	Treatment, concomitant drug and treatment-by-concomitant drug interaction as fixed effects, subject as random effect with log(baseline) as additional covariate.	No Imputation	Sensitivity analysis with two covariates on primary endpoint.
Percent Change from Baseline in % Liver Fat assessed by MRI-PDFF at Week 16	FAS	Dose-Response model	See section 5.2.1.3	No Imputation	Supportive analysis on primary endpoint
Percent change from Baseline in ALT Over Time	FAS	MMRM on log-transformed relative change	Treatment, time and treatment-by-week interaction as fixed effects, subject as random effect with log(baseline) as covariate.	No Imputation	Secondary analysis

[redacted]

Endpoint	**Analysis Set**	**Statistical Method**	**Model/ Covariates/ Strata**	**Missing Data**	**Interpretation**
Percent Change from Baseline in ALT at Week 16	FAS	Dose-Response model	See section 5.2.1.3	No Imputation	Supportive Analysis on secondary endpoint
Tertiary endpoints: Change from Baseline over time	FAS	MMRM on change from baseline	Treatment, time and treatment-by-week interaction as fixed effects, subject as random effect and baseline as covariate	No Imputation	Tertiary analysis
Tertiary endpoints: Percent change from Baseline over time other than % Liver Fat using MRI-PDFF	FAS	MMRM on log-transformed relative change	Treatment, time and treatment-by-week interaction as fixed effects, subject as random effect and log(baseline) as covariate.	No Imputation	Tertiary Analysis

[redacted]

Appendix 2. DATA DERIVATION DETAILS

Appendix 2.1. Definition and Use of Visit Windows in Reporting

Each collected assessment will be mapped to their actual visit window according the following algorithm described in Table A2.1-1.

Table A2.1-1. Algorithm to Map Collected assessment to Actual Visit

Visit label	Protocol defined Window	Visit	Visit window as part of final reporting*
			For efficacy, PD, PK and safety endpoints
Pre-Qualification	NA	1	NA
Screen 1	NA	2	NA
Screen 2	NA	3	NA
Run-In	NA	3	NA
Week -2	-14 ± 2	4	• Between Week 0 and Week -2, data are used to define baseline • Intent is for baseline to be result closest ***prior to*** dosing on randomized regimen on Day 1 [Visit 5] • Window for baseline measurements is from Day -20 to Day 1 [Visit 5] pre-dose
Week 0	1	5	
Week 2	14 ± 4	6	'Date of assessment' – 'Baseline date' + 1 = [8; 20]; ie, ±6 days
Week 4	28 ± 4	7	'Date of assessment' – 'Baseline date' + 1 = [22 ; 34]; ie, ±6 days
Week 6	42 ± 4	8	'Date of assessment' – 'Baseline date' + 1 = [36 ; 48]; ie, ±6 days
Week 8	56 ± 4	9	'Date of assessment' – 'Baseline date' + 1 = [50 ; 62]; ie, ±6 days
Week 12	84 ± 4	10	'Date of assessment' – 'Baseline date' + 1 = [77 ; 91]; ie, ±7 days
Week 16	108 ± 4	11	'Date of assessment' – 'Baseline date' + 1 = [101; 115]; ie, ±7 days
1st Follow-up visit	10 ± 4 days post last dose	12	At least 5 days post last dose and up to days 27 post last dose
2nd Follow-up visit	28 + 7 days post last dose	13	

Note: * permitted to be wider than window defined in protocol for purposes of reporting data by nominal visit.

Special consideration:

- For subject withdrawn early, the follow-up visit window (F/U) should be excluded from windowing algorithm, and the visit label is defined in accordance with the nominal label as recorded on the "E_TERM/FOLLOW_UP" Administration CRF.
- If 2 or more observations fall within the visit window for Week N, the observation used will be the one closest to Day 7N (randomization day is Day 1). If the 2 closest observations are equidistant from Day 7N, the earlier observation will be used.

[redacted].

An assessment is considered 'on treatment' if the following is true:

- 'Date (time when databased) of assessment' > 'Baseline date (time)' and ≤'Last dose date (time)'

Appendix 2.2. Data Derivation for Primary, Secondary, and Tertiary Endpoints

Appendix 2.2.1. % Liver Fat using MRI-PDFF

For each subject with nonalcoholic fatty liver disease (entire study population), the Whole Liver PDFF (WLPDFF in %) will be calculated from the pre-defined individual segmental PDFF (SPDFF in %) measured in Segment I, II, III, IVa, IVb, V, VI, VII and VIII as follows:

$$WLPDFF = (SPDFF_{Segment\ I} + SPDFF_{Segment\ II} + SPDFF_{Segment\ III} + SPDFF_{Segment\ IVa} + SPDFF_{Segment\ IVb} + SPDFF_{Segment\ V} + SPDFF_{Segment\ VI} + SPDFF_{Segment\ VII} + SPDFF_{Segment\ VIII}) / (\text{number of segments assessed and no missing/mapping at Baseline, and on Week 4, Week 12, and Week 16})$$

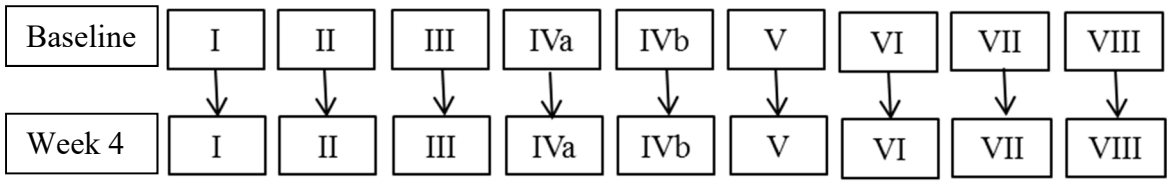
Note: If any negative value is reported by the central vendor for an individual segment, the value of this segment will be imputed to "0" before inclusion into derivation of whole liver PDFF. However, the negative value, as offered by the central vendor, will be included in the listing.

<u>Calculation Rule for % Liver Fat when Assessed with MRI-PDFF:</u>

WLPDFF will be calculated on mapping no missing segments at Baseline, and on Week 4, Week 12, and Week 16.

For instance,

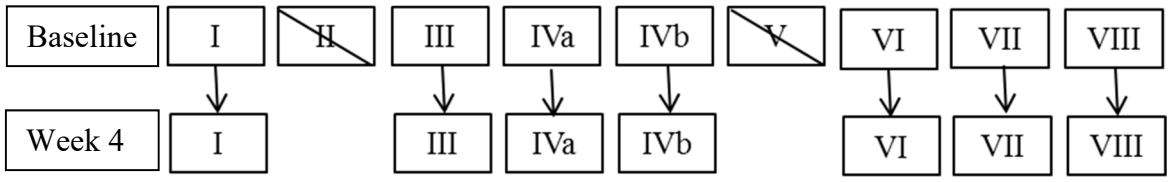
if results are reported for all segments at Baseline, and on Week 4, Week 12, and Week 16 (see below), The number of segments assessed is equal to 9 .



If at Baseline, all segment data are available but, on Week 4, or Week 12, or Week 16, only 7 segments have results reported, WLPDFF will be calculated at Baseline, and on Week 4, Week 12, and Week 16 and using the matching individual segments providing values on ***<u>all</u>*** weeks. For instance, based on the schema below, segment II and Segment V will not be used to calculate

[redacted]

WLPDFF on any of the time points reported. In this case, WLPDFF will be calculated using data in the 7 segments reported at **all** visits.



WLPDFF and individual % liver fat measured in each individual segment will be listed. WLPDFF will be summarized by randomized arm, and Week as described in Table A2.2-1.

Note: If baseline at the individual segment level is zero by itself or after imputation, % change from baseline at individual segment level will be reported as "." (because not calculable) with a footnote explaining that the value at baseline was zero or imputed to zero.

Table A2.2-1: Summary Statistics for % Liver Fat (WLPDFF) assessed by MRI-PDFF

Weeks	Summary Statistics
Screen2, Baseline, and 4, 12, 16	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3

Derived Parameters

Individual change from baseline, percent change from baseline and relative change from baseline for WLPDFF will be calculated and listed. Percent change from baseline will be summarized by randomized arm as described in Table A2.2-2.

Table A2.2-2: Summary Statistics for Percent Change From Baseline in % Liver Fat (WLPDFF) using MRI-PDFF

Weeks	Derived parameter	Summary Statistics
4, 12, 16	Percent Change from Baseline	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3

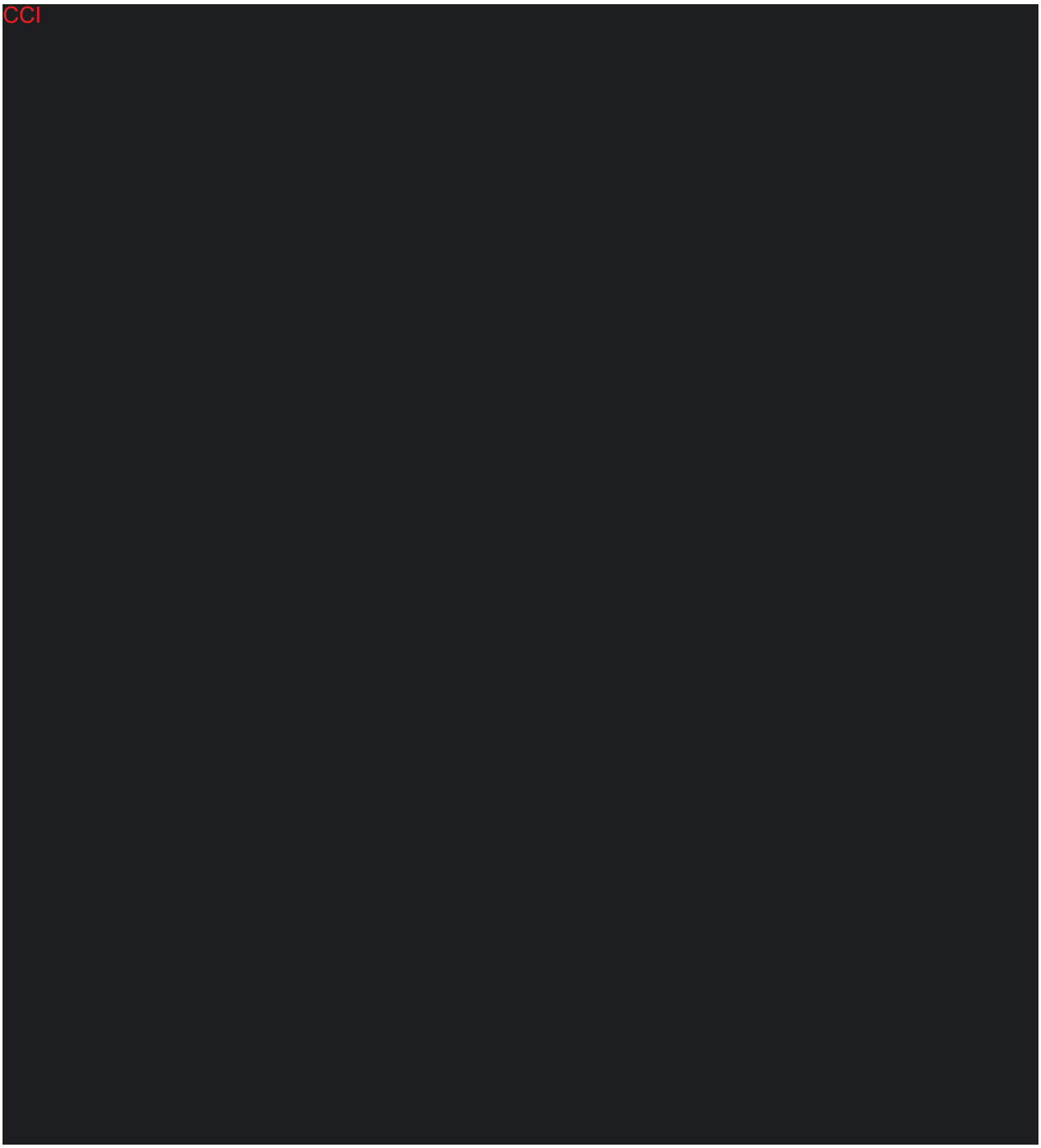
Figures:

- Median response versus week will be presented against randomized arm. Q1 and Q3 values will be added.

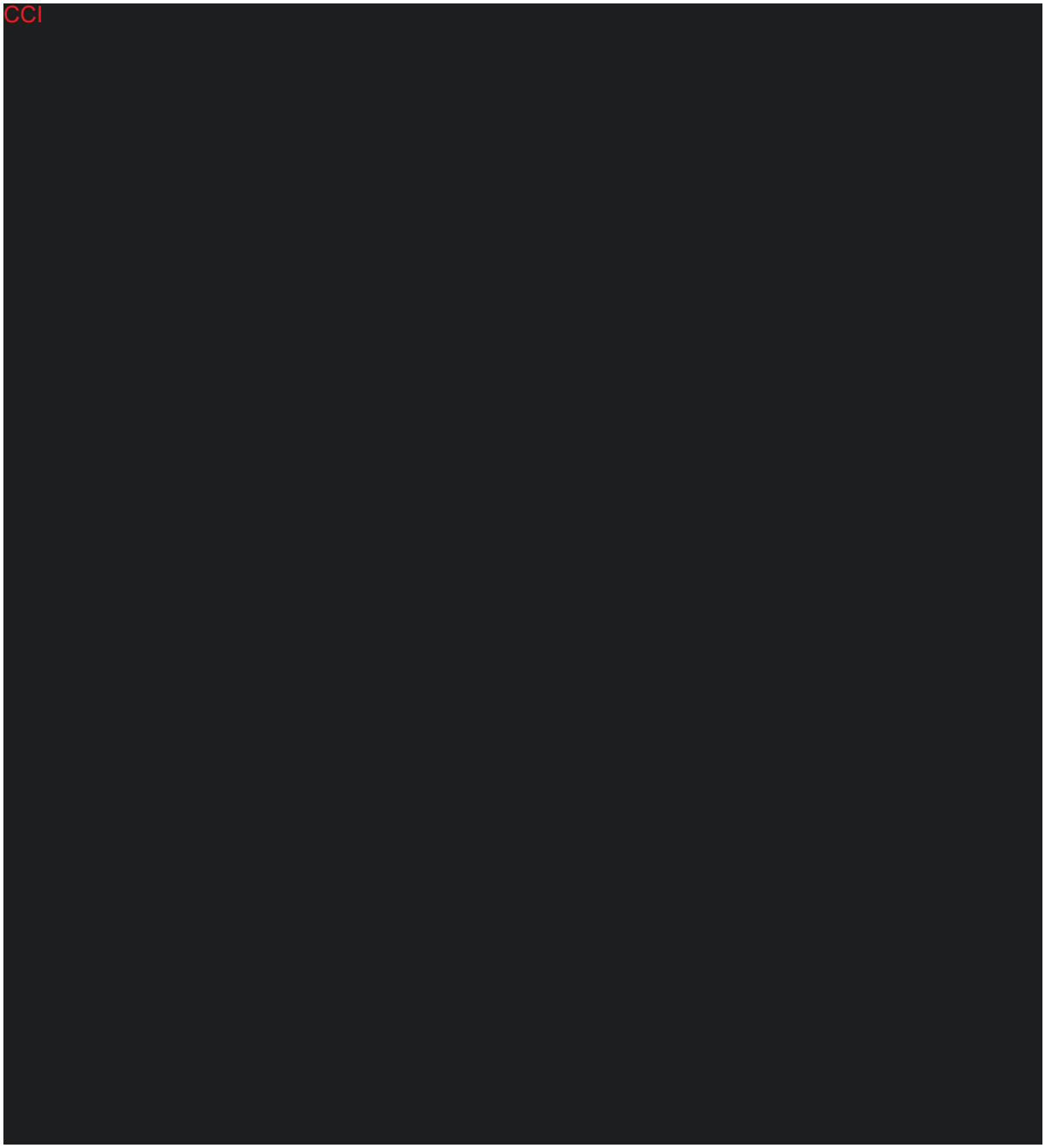
[redacted]

- Median percent change from baseline versus week will be presented against randomized arm. A reference line cutting the Y-Axis at 0 will be overlaid. Q1 and Q3 values will be added.
- Box and whisker plots for individual percent change from baseline at week 16 versus total daily dose will be presented and overlaid with arithmetic means. Only box, whiskers, median, arithmetic mean, and outliers should be reported in this graphic. Total daily dose will be considered as categorical (see Appendix 3.1).

CCI



[redacted].



Appendix 2.2.4. Liver Function Tests

Alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALKP), gamma-glutamyl transferase (GGT), and total bilirubin (TBILI) samples will be taken according to the Schedule of Activities given in the protocol.

CCI [redacted]

Table A2.2-7: Summary Statistics for ALT, AST, ALKP, GGT, and TBILI Over Time

Weeks	Parameters	Summary Statistics
Baseline, 2, 4, 6, 8, 12, 16, FU	ALT, AST, ALKP, GGT, TBILI in "diagnosed/presumed nonalcoholic steatohepatitis subjects"	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3
CCI [redacted]		

Note: FU = 1st Follow-up as defined in Tables 1 and 2 of the protocol

Derived Parameters

Individual percent change from baseline and relative change from baseline over time for each liver function test concentration will be calculated. CCI [redacted] percent change from baseline will be summarized by randomized arm, and week as follow:

Table A2.2-8: Summary Statistics for Percent Change From Baseline in Liver Function Test (ALT, AST, ALKP, GGT, and TBILI) Concentrations Over Time

Weeks	Parameter	Population	Summary Statistics
Baseline, 2, 4, 6, 8, 12, 16, FU	Percent Change from baseline for each liver function test (ie ALT, AST, ALKP, GGT, TBILI)	Subjects with "diagnosed / presumed nonalcoholic steatohepatitis"	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3
CCI [redacted]			

Note: FU = 1st Follow-up as defined in Tables 1 and 2 of the protocol

Figure (1 set for each liver function test parameter CCI [redacted]):

[redacted]

- Median response versus week will be presented against randomized arm. Q1 and Q3 values will be added.
- Median percent change from baseline versus week will be presented against randomized arm. A reference line cutting the Y-Axis at 0 will be overlaid. Q1 and Q3 values will be added.

Additional Figure for ALT concentrations only in Subjects with "diagnosed / presumed nonalcoholic steatohepatitis"

- Box and whisker plots for individual percent change from baseline at week 16 versus total daily dose will be presented and overlaid with arithmetic means. Only box, whiskers, median, arithmetic mean and outliers should be reported in this graphic. Total daily dose will be considered as categorical (see Appendix 3.1).

CCI

[redacted].

CCI

[redacted]

CCI

[redacted]

Appendix 2.2.6. Lipid Panel

All lipids will be assessed on the fasting state using serum matrix. Triglycerides(TG), direct low density lipoprotein (LDL-C), very low-density lipoprotein (VLDL), high-density lipoprotein (HDL-C), and total cholesterol samples in subjects with nonalcoholic fatty liver disease (ie entire study population) will be taken according to the Schedule of Activities given in the protocol in subjects with nonalcoholic fatty liver disease.

TG, direct LDL-C, VLDL, HDL-C, and total cholesterol will be listed and summarized by randomized arm, and week as described in Table A2.2-12.

Table A2.2-12: Summary Statistics for TG, Direct LDL-C, HDL-C and Total Cholesterol Over Time

Weeks	Parameters	Summary Statistics
Baseline, 2, 4, 6, 8, 12, 16, FU	TG, direct LDL-C, VLDL, HDL-C, Total Cholesterol	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3

Note: FU = 1st Follow-up as defined in Tables 1 and 2 of the protocol

Derived Parameters

For each lipid parameter, individual relative change and percent change from baseline over time will be calculated. Percent change from baseline will be summarized by randomized arm, and week as follow:

Table A2.2-13: Summary Statistics for Percent Change From Baseline in Lipids (ie TG, direct LDL-C, VLDL, HDL-C, and Total Cholesterol) Over Time

Weeks	Derived parameter	Summary Statistics
Baseline, 2, 4, 6, 8, 12, 16, FU	Percent Change from Baseline for each lipid parameter	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3

Note: FU = 1st Follow-up as defined in Tables 1 and 2 of the protocol

Figure (1 set for each lipid parameter):

- Median response versus week will be presented against randomized arm. Q1 and Q3 values will be added.
- Median percent change from baseline versus week will be presented against randomized arm. A reference line cutting the Y-Axis at 0 will be overlaid. Q1 and Q3 values will be added.

[redacted].

Box and whisker plots by randomized arm and by week will also be displayed to characterize the distribution of percent changes from baseline over time (see Appendix 3.2). Only box, whiskers, arithmetic mean, and outliers should be reported in this graphic.

Figure for triglycerides only (should be included in the safety set of figures)

- For subjects with any **after** Day 1 (ie post-randomization) where fasting TG ≥ 600 mg/dL and up to Week 16 (**include** unplanned in chronological/continuous order – not categorical), graph fasting TG, FPI, ApoC3, and FPG data together. To do so, real date reported in eCRFe-data loads received from Covance will be used

- If possible, this figure containing 4x4 panels (1 panel per subject) by treatment will be created with LEFT Y-axis reporting TG and FPG levels and RIGHT Y-axis reporting FPI and ApoC3 levels.

Appendix 2.2.7. Glycated Hemoglobin

Glycated hemoglobin (HbA1c) values will be assessed at timepoints described in the Schedule of Activity of the protocol.
HbA1c values will be listed and summarized by randomized arm, and week as described in Table A2.2-14

Table A2.2-14: Summary Statistics for HbA1c Over Time

Weeks	Parameter	Population	Summary Statistics
Baseline, 2, 4, 6, 8, 12, 16, FU	HbA1c	Subjects with nonalcoholic fatty liver disease (ie entire study population)	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3
Baseline, 2, 4, 6, 8, 12, 16, FU	HbA1c	T2DM subjects	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3

Note: FU = 1st Follow-up as defined in Tables 1 and 2 of the protocol

Derived Parameter

Individual change from baseline over time on HbA1c will be calculated, listed, and summarized by randomized arm, and week as described in Table A2.2-15.

Table A2.2-15: Summary Statistics for Change From Baseline in HbA1c Over Time

Weeks	Derived parameter	Population	Summary Statistics
Baseline, 2, 4, 6,	Change from Baseline	Subjects with nonalcoholic fatty liver disease (ie entire	N, arithmetic mean, median, standard deviation,

[redacted]

Weeks	Derived parameter	Population	Summary Statistics
8, 12, 16, FU		randomized study population)	minimum, maximum, Q1, Q3
2, 4, 6, 8, 12, 16, FU	Change from Baseline	T2DM Randomized Subjects	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3

Note: FU = 1st Follow-up as defined in Tables 1 and 2 of the protocol

Figure (1 set for entire randomized study population and 1 set for T2DM subjects):

- Median response versus week will be presented against randomized arm. Q1 and Q3 values will be added.
- Median change from baseline versus week will be presented against randomized arm. A reference line cutting the Y-Axis at 0 will be overlaid. Q1 and Q3 values will be added.

CCI

[redacted]

CCI

[redacted]

Appendix 2.2.9. Fasting Plasma Glucose and Fasting Plasma Insulin (T2DM Subjects Only)

Fasting plasma glucose (FPG) and fasting plasma insulin (FPI) values will be assessed at timepoints described in the Schedule of Activity of the protocol.
FPG and FPI values will be listed and summarized by randomized arm, and week as described in Table A2.2-18.

Table A2.2-18: Summary Statistics for FPG And FPI Over Time

Weeks	Parameter	Summary Statistics
Baseline, 2, 4, 6, 8, 12, 16, FU	FPG in subject with T2DM only	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3
Baseline, 2, 4, 6, 8, 12, 16	FPI in subject with T2DM only	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3

Note: FU = 1st Follow-up as defined in Tables 1 and 2 of the protocol

Note: if plasma glucose concentrations are reported in mmol/L, transform the value in mg/dL using the following equation $Glucose_{(mg/dL)} = 18 * Glucose_{(mmol/L)}$

Derived Parameters

Individual change from baseline over time on FPG and FPI will be calculated, and listed. Changes from baseline will be summarized by randomized arm, and week as described in Table A2.2-19.

[redacted]

Table A2.2-19: Summary Statistics for Change From Baseline in Fasting Plasma Glucose and Fasting Plasma Insulin Over Time

Weeks	Derived parameter	Summary Statistics
2, 4, 6, 8, 12, 16, FU	FPG - change from baseline in subject with T2DM only	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3
2, 4, 6, 8, 12, 16	FPI - change from baseline in subject with T2DM only	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3

Note: FU = 1st Follow-up as defined in Tables 1 and 2 of the protocol

Figure (1 set for each parameter):

- Median response versus week will be presented against randomized arm. Q1 and Q3 values will be added.
- Median change from baseline versus week will be presented against randomized arm. A reference line cutting the Y-Axis at 0 will be overlaid. Q1 and Q3 values will be added.

Appendix 2.2.10. Homeostatic Model Assessment for Insulin Resistance (T2DM Subjects Only)

The Homeostatic Model Assessment for Insulin Resistance (HOMA-IR) will be derived using fasting plasma glucose values and fasting plasma insulin values at timepoints described in the Schedule of Activity of the protocol.

The formula used to derive HOMA-IR is as follow:

$$\text{HOMA-IR} = (\text{Glucose}_{\text{Fasting Concentration}} \times \text{Insulin}_{\text{Fasting Concentration}}) / 405$$

where plasma glucose concentrations is reported in mg/dL and plasma insulin is reported in mU/L. As such, HOMA-IR is unit-less.
HOMA-IR values will be listed and summarized by randomized arm, and week as described in Table A2.2-20.

Table A2.2-20: Summary Statistics for HOMA-IR Over Time

Weeks	Parameter	Summary Statistics
Baseline, 2, 4, 6, 8, 12, 16	HOMA-IR in subjects with T2DM only	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3

Derived Parameters

[redacted]

Individual change from baseline over time on HOMA-IR will be calculated, and listed. Changes from baseline will be summarized by randomized arm, and week as described in Table A2.2-21.

Table A2.2-21: Summary Statistics for Change From Baseline in HOMA-IR Over Time

Weeks	Derived parameter	Summary Statistics
2, 4, 6, 8, 12, 16	Change from Baseline in in subjects with T2DM only	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3

Figure:

- Median response versus week will be presented against randomized arm. Q1 and Q3 values will be added.
- Median change from baseline versus week will be presented against randomized arm. A reference line cutting the Y-Axis at 0 will be overlaid. Q1 and Q3 values will be added.

CCI

CCI

[redacted]

Appendix 2.2.12. Serum Apolipoprotein

Serum apolipoprotein (ie ApoA1, ApoB, ApoC3 and ApoE) in subjects with nonalcoholic fatty liver disease (ie entire study population) will be assessed at timepoints described in the Schedule of Activity of the protocol
Each serum apolipoprotein will be listed and summarized by randomized arm, and week as described in Table A2.2-24.

Table A2.2-24: Summary Statistics for Serum Apoliprotein Over Time

Weeks	Parameter	Summary Statistics
Baseline, 2, 4, 6, 8, 12, 16	ApoA1 in subject with nonalcoholic fatty liver disease	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3
Baseline, 2, 4, 6, 8, 12, 16	ApoB in subject with nonalcoholic fatty liver disease	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3
Baseline, 2, 4, 6, 8, 12, 16	ApoC3 in subject with nonalcoholic fatty liver disease	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3
Baseline, 2, 4, 6, 8, 12, 16	ApoE in subject with nonalcoholic fatty liver disease	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3

Derived Parameters

For each serum apolipoprotein, individual relative from baseline and percent change from baseline over time will be calculated. Percent change from baseline will be summarized by randomized arm, and week as follow:

[redacted]

Table A2.2-25: Summary Statistics for Percent Change From Baseline in Serum Apoliprotein Over Time

Weeks	Parameter	Summary Statistics
2, 4, 6, 8, 12, 16	ApoA1 - Percent Change from Baseline in subject with nonalcoholic fatty liver disease	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3
2, 4, 6, 8, 12, 16	ApoB - Percent Change from Baseline in subject with nonalcoholic fatty liver disease	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3
2, 4, 6, 8, 12, 16	ApoC3 - Percent Change from Baseline in subject with nonalcoholic fatty liver disease	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3
2, 4, 6, 8, 12, 16	ApoE - Percent Change from Baseline in subject with nonalcoholic fatty liver disease	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3

Figure (1 set for each serum apolipoprotein) :

- Median response versus week will be presented against randomized arm. Q1 and Q3 values will be added.
- Median percent change from baseline versus week will be presented against randomized arm. A reference line cutting the Y-Axis at 0 will be overlaid. Q1 and Q3 values will be added.
- Box and whisker plots by randomized arm and by week will also be displayed to characterize the distribution of percent changes from baseline over time (see Appendix 3.2). Only box, whiskers, median, arithmetic mean, and outliers should be reported in this graphic.

CCI

[redacted].

CCI

[redacted]

Appendix 2.2.14. Markers of Apoptotic Activity - CK18-M30 and CK18-M65

Markers of apoptotic activity will be assessed at timepoints described in the Schedule of Activity of the protocol using two circulating biomarkers measuring either the caspase-cleaved cytokeratin-18-M30 (ie CK18-M30) produced during apoptosis or the caspase-cleaved and intact cytokeratin-18-M65 (ie CK18-M65) usually released from cells undergoing necrosis.

CK18-M30 and CK18-M65 levels will be listed and summarized CCI [redacted] randomized arm, and week as described in Table A2.2-28.

Table A2.2-28: Summary Statistics for CK18-M30 and CK18-M65 Over Time

Weeks	Parameters	Population	Summary Statistics
CCI [redacted]			

[redacted].

Baseline, 4, 8, 12, 16	CK18-M30	subject with diagnosed/presumed nonalcoholic steatohepatitis	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3
Baseline, 4, 8, 12, 16	CK18-M65	subject with diagnosed/presumed nonalcoholic steatohepatitis	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3

Derived Parameters

Individual relative change and percent change from baseline over time for each cell death biomarker concentration will be calculated. Percent changes from baseline will be summarized by stratum, randomized arm, and week as follow:

Table A2.2-29: Summary Statistics for Percent Change From Baseline in Cell Death Biomarkers (ie CK18-M30, and CK18-M65) Levels Over Time

Weeks	Parameters	Population	Summary Statistics
CCI [redacted]			
4, 8, 12, 16	CK18-M30 Percent Change from Baseline	subject with diagnosed/presumed nonalcoholic steatohepatitis	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3
4, 8, 12, 16	CK18-M65 Percent Change from Baseline	subject with diagnosed/presumed nonalcoholic steatohepatitis	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3

Figure (1 set for each cell death biomarker CCI [redacted]):

- Median response versus week will be presented against randomized arm. Q1 and Q3 values will be added.
- Median percent change from baseline versus week will be presented against randomized arm. A reference line cutting the Y-Axis at 0 will be overlaid. Q1 and Q3 values will be added.

[redacted].

Appendix 2.2.15. Collagen formation-related Biomarkers - Pro-C3 and Pro-C6

Collagen formation will be assessed at timepoints described in the Schedule of Activity of the protocol using two circulating biomarkers measuring either N-terminal pro-peptide of type III procollagen (ie Pro-C3) or C-terminal fragment of the α3 chain of procollagen type VI (ie Pro-C6).

Pro-C3 and Pro-C6 levels will be listed and summarized CCI [redacted] randomized arm, and week as described in Table A2.2-30.

Table A2.2-30: Summary Statistics for Pro-C3 and Pro-C6 Levels Over Time

Weeks	Parameters	Summary Statistics
CCI [redacted]		
Baseline, 4, 8, 12, 16	Pro-C3 in subjects with diagnosed / presumed nonalcoholic steatohepatitis	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3
CCI [redacted]		
Baseline, 4, 8, 12, 16	Pro-C6 in subjects with diagnosed / presumed nonalcoholic steatohepatitis	N, arithmetic mean, median, %CV, standard deviation, minimum, maximum, Q1, Q3

Derived Parameters

Individual relative change and percent change from baseline over time for each fibrogenesis biomarker will be calculated. Percent changes from baseline will be summarized by stratum, randomized arm, and week as follow:

Table A2.2-31: Summary Statistics for Percent Change From Baseline in Fibrogenesis Biomarkers (ie Pro-C3, and Pro-C6) Levels Over Time

Weeks	Parameters	Population	Summary Statistics
CCI [redacted]			

[redacted]

4, 8, 12, 16	Pro-C3 Percent Change from Baseline	Subjects with diagnosed / presumed nonalcoholic steatohepatitis	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3
CCI [redacted]			
4, 8, 12, 16	Pro-C6 Percent Change from Baseline	Subjects with diagnosed / presumed nonalcoholic steatohepatitis	N, arithmetic mean, median, standard deviation, minimum, maximum, Q1, Q3

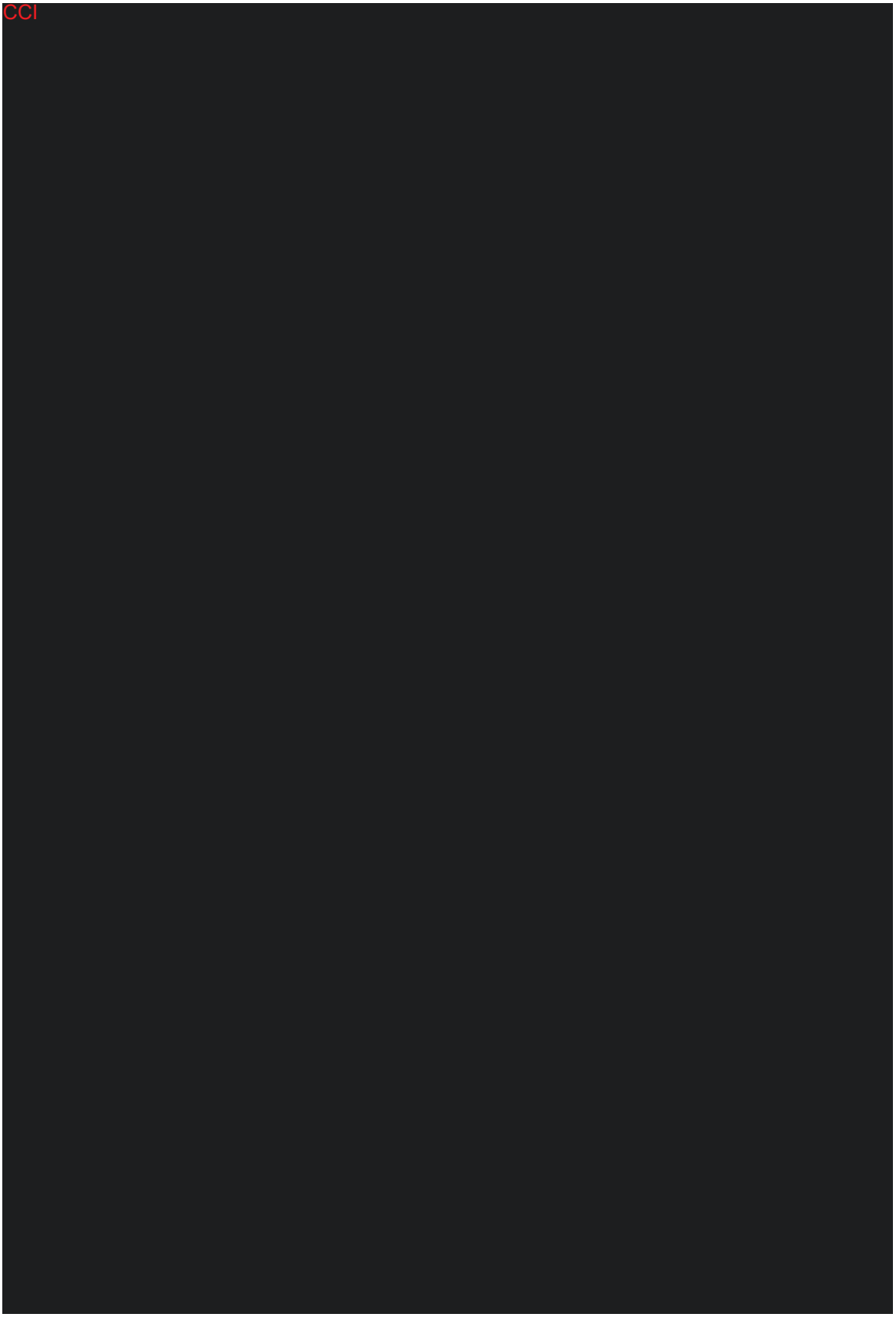
Figure (1 set for each fibrogenesis biomarker CCI [redacted]):

- Median response versus week will be presented against randomized arm. Q1 and Q3 values will be added.
- Median percent change from baseline versus week will be presented against randomized arm. A reference line cutting the Y-Axis at 0 will be overlaid. Q1 and Q3 values will be added.

CCI [redacted]

[redacted]

CCI



.

CCI

CCI

CCI

[redacted]

[redacted]

CCI

[redacted]

[redacted].

CCI

[redacted]

Appendix 2.3. Data Derivation for Non Standard Safety Endpoints

CCI

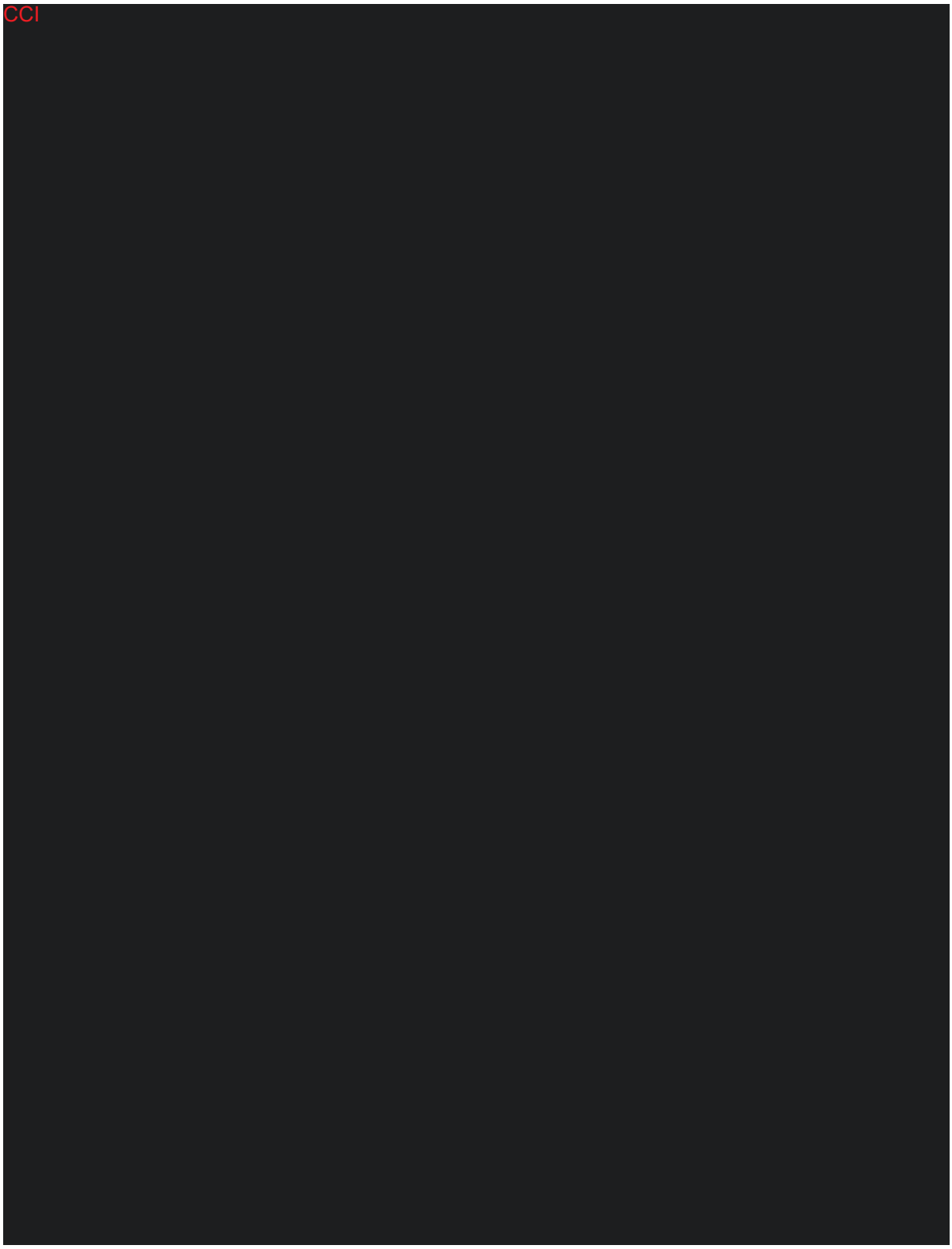
[redacted]

[redacted].

CCI

CCI

.


CCI

CCI

.

CCI

[redacted]

Appendix 2.4. Compliance

Compliance with double-blinded study medication will be calculated using data entry exclusively from the **oral dosing** CRF pages.

For PPAS determination, compliance will be calculated using ACTIVE bottles only (ie not using all 3 bottles).

The following formula are only indicative when the three bottles are containing active drug and should therefore be adapted for C1171002 study
On Week 4, Week 8, Week 12 and Week 16, the number of tablets used by the subject between each visit will be calculated as below (assuming that each bottle – B1, B2, and B3 - has 38 tablets):

$$\#TabUsed_{Week1\text{-}Week4} = (38.- (EndDate - StartDate)_{Week4Page1/3} +1) + (38.-. (EndDate - StartDate)_{Week4Page2/3} +1)+(38.-. (EndDate - StartDate)_{Week4Page3/3} +1)$$

$$\#TabUsed_{Week4\text{-}Week8} = (38.- (EndDate - StartDate)_{Week8Page1/3}+1) + (38.-. (EndDate - StartDate)_{Week8Page2/3}+1) + (38.-. (EndDate - StartDate)_{Week8Page3/3}+1)$$

$$\#TabUsed_{Week8\text{-}Week12} = (38.- (EndDate - StartDate)_{Week12Page1/3}+1) + (38.-. (EndDate - StartDate)_{Week12Page2/3}+1) + (38.-. (EndDate - StartDate)_{Week12Page3/3}+1)$$

$$\#TabUsed_{Week12\text{-}Week16} = (38.- (EndDate - StartDate)_{Week16Page1/3}+1) + (38.-. (EndDate - StartDate)_{Week16Page2/3}+1) + (38.-. (EndDate - StartDate)_{Week16Page3/3}+1)$$

$$\#TabUsedDuringStudy = \#TabUsed_{Week1\text{-}Week4} + \#PillsUsed_{Week4\text{-}Week8} + \#TabUsed_{Week8\text{-}Week12} + \#PillsUsed_{Week12\text{-}Week16}$$

Where $\#Tab_{ReturnedBi}$ = Remaining tablets counted in each bottle and reported by site pharmacist in each CRF page as returned amount.

For each subject, the total expected number of tablets used during the study is computed as

[redacted]

#ExpectedTabUsedOverall = 3*(last dose date – first dose date + 1)

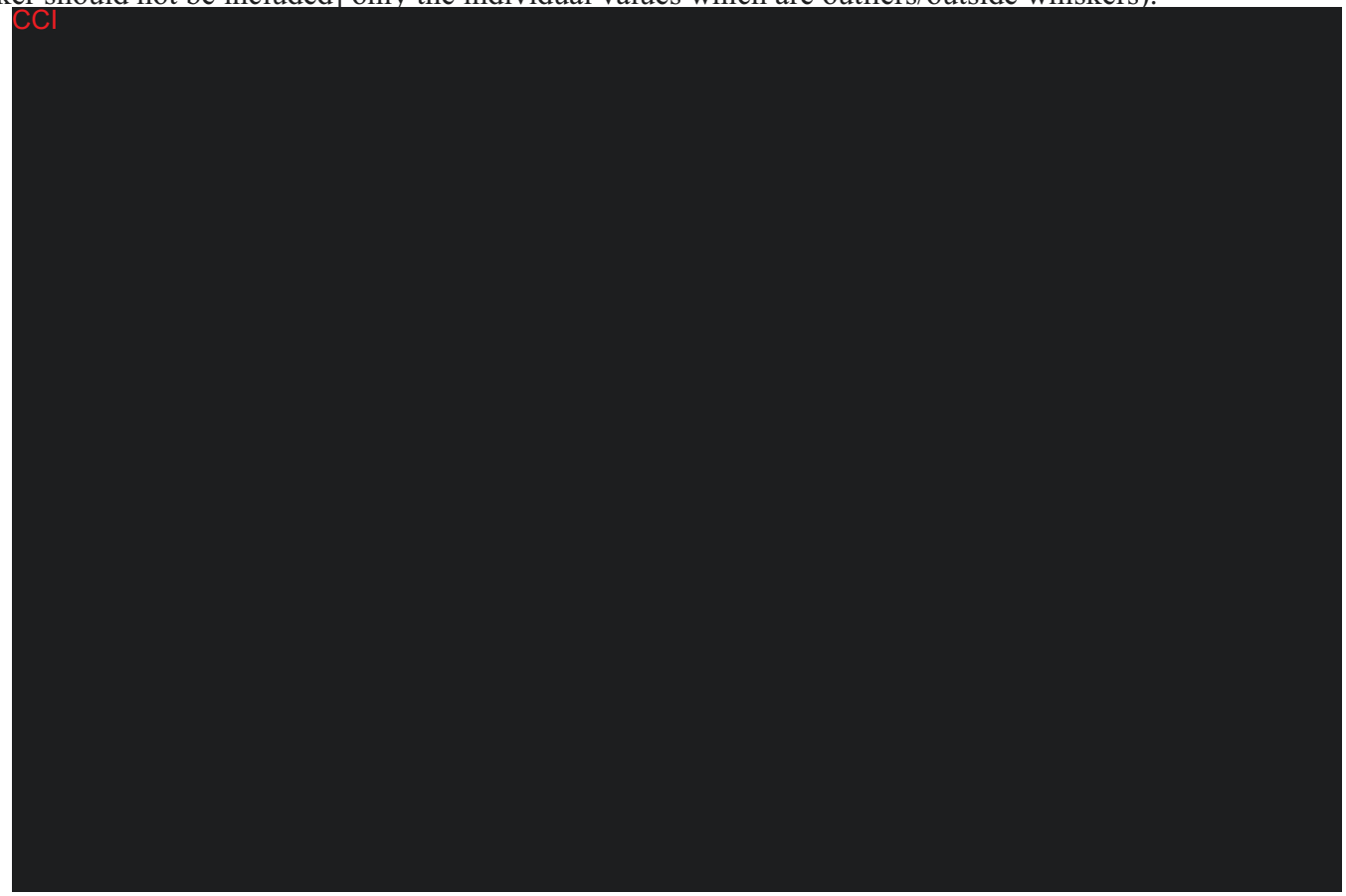
NOTE: Data entered on **Drug Accountability** CRF page (which only captures details regarding pill count returned) will not be used to calculate and report out subject-level compliance.

[redacted]

Appendix 3. GRAPHICAL REPRESENTATIONS

Appendix 3.1. Box and Whisker Plot Representation for Changes/Percent Changes from Baseline at Week 16 by Randomized Arm

Note: To be adapted for C1171002 design (eg weeks as legend) with correct total daily dose. Only box, whiskers, median, arithmetic mean and outliers should be reported in this graphic (individual subject data [open circles within 5th and 95th whisker should not be included] only the individual values which are outliers/outside whiskers).



[redacted]

Appendix 3.2. Box and Whisker Plot for Changes/ Percent Changes From Baseline by Randomized Arm and Week

Note: To be adapted for C1171002 design (eg weeks as legend) with correct total daily dose. Only box, whiskers, median, arithmetic mean and outliers should be reported in this graphic (individual subject data [open circles within 5th and 95th whisker should not be included] only the individual values which are outliers/outside whiskers).

[redacted]

Appendix 3.3. Figure for MMRM LS Mean vs Time by Randomized Arm for Changes/ Percent Changes From Baseline Over Time

Note: To be adapted for C1171002 design (eg correct total daily dose in legend, Weeks on X-axis, and Y-axis label).

[redacted]

Appendix 3.4. Figure for MMRM LS Mean vs Time by Randomized Arm for Changes/ Percent Changes From Baseline Placebo-Adjusted Over Time

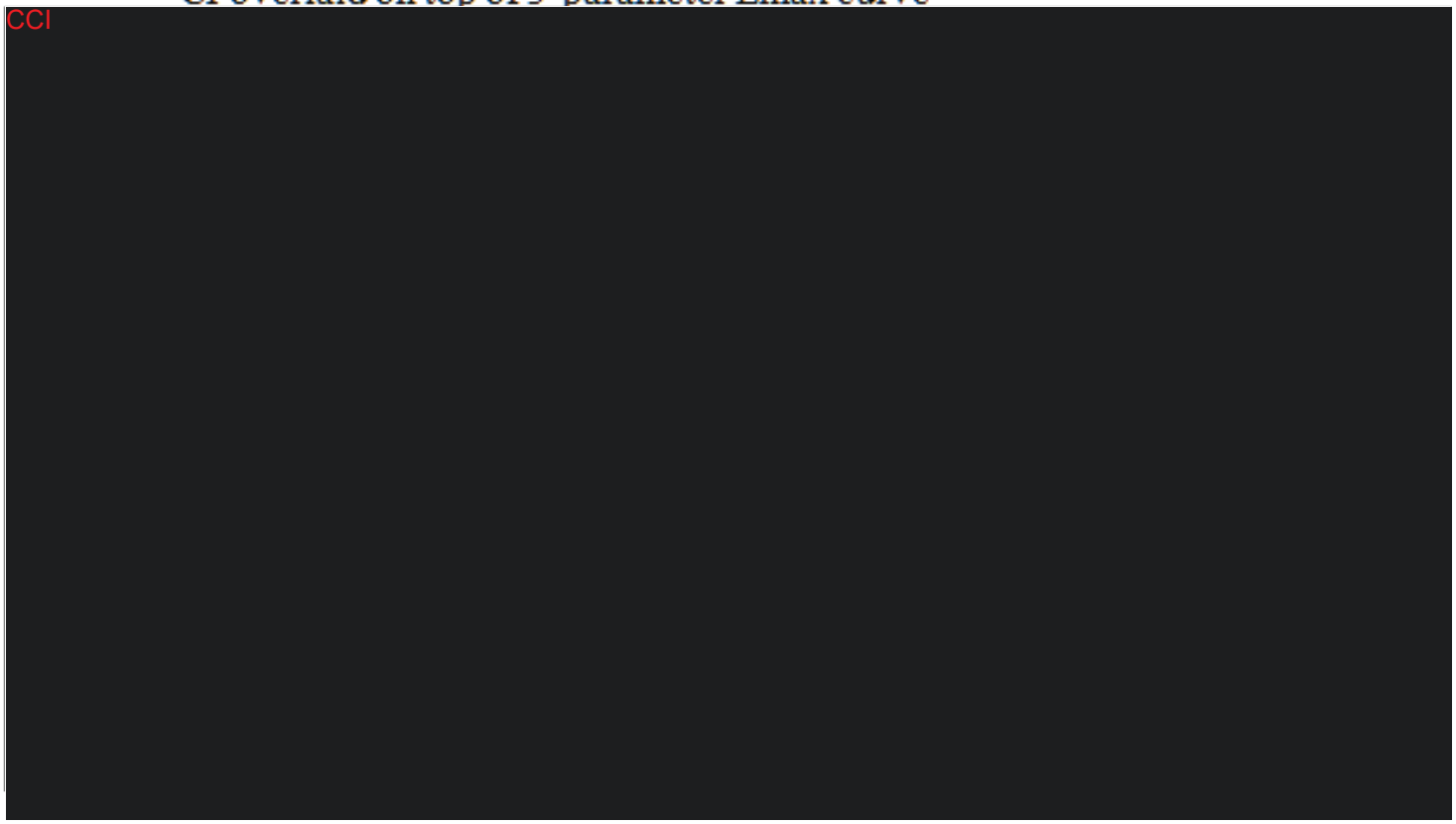
Note: To be adapted for C1171002 design (eg correct total daily dose in legend, Weeks on X-axis, and Y-axis label).

[redacted]

Appendix 3.5. Figure for MMRM LS means for Changes/ Percent Changes From Baseline by Randomized Arm at Week16 overlaid with Emax Estimates (and 80%CI)

Note: To be adapted for C1171002 design (eg correct total daily dose on X-axis, Y-axis label, inclusion of "y=0" line).

Figure 1: HbA1C (%) Change from Baseline at Week 12 – Placebo-Adjusted LS Mean and 80% CI overlaid on top of 3-parameter Emax curve



[redacted].

Appendix 3.6. Figure for MMRM LS means for Placebo-Adjusted Changes/ Percent Changes From Baseline by Randomized Arm at Week16

Note: To be adapted for C1171002 design (eg correct total daily dose on X-axis, Y-axis label, inclusion of "y=0" line).

[redacted]

Appendix 3.7. Additional Figure requested

In order to apprehend the dose –effect relationship between different endpoints, please provide the following additional figures:

- % change in liver fat at Wk16 (X-axis) versus % change in fasting TG (Y-axis) with 5 colored dots/squares/circles representing 5 arms (red = placebo) to permit either an assessment of whether increase in TG is related to dose or individual-subject sensitivity to drug.
- Change in fasting TG at Wk16 (X-axis) versus change in fasting ApoC3 (Y-axis) with 5 colored dots/squares/circles representing 5 arms (red = placebo) to permit either an assessment of whether increase in ApoC3 is related to Dose or individual-subject sensitivity to drug.

[redacted].