Protocol C2541013

A PHASE 2, RANDOMIZED, DOUBLE-BLIND, DOUBLE-DUMMY, PLACEBO CONTROLLED, DOSE RANGING, DOSE-FINDING, PARALLEL GROUP STUDY TO ASSESS EFFICACY AND SAFETY OF PF 06865571 (DGAT2I) ALONE AND WHEN COADMINISTERED WITH PF 05221304 (ACCI) IN ADULT PARTICIPANTS WITH BIOPSY CONFIRMED NONALCOHOLIC STEATOHEPATITIS AND FIBROSIS STAGE 2 OR 3

Statistical Analysis Plan
(SAP)

Version: **3**

Date: 03 May 2024

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

TABLE OF CONTENTS

LIST OF TABLES 6
LIST OF FIGURES 6
APPENDICES 7
1. VERSION HISTORY 8
2. INTRODUCTION 14
2.1. Study Objectives, Endpoints, and Estimands 14
2.1.1. Primary Estimand(s) 14
2.1.2. Secondary Estimand(s) 15
2.1.3. Additional Estimand(s) 16
2.2. Study Design 18
3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS 20
3.1. Primary Endpoint(s) 20
3.2. Secondary Endpoint(s) 22
3.3. Other Endpoint(s) 22
3.4. Baseline Variables 25
3.5. Safety Endpoints 25
3.5.1. Adverse Events 25
3.5.2. Laboratory Data 26
3.5.3. Other Standard Safety other than AEs 29
4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS) 31
4.1. Treatment Misallocations 33
5. GENERAL METHODOLOGY AND CONVENTIONS 33
5.1. Hypotheses and Decision Rules 33
5.2. General Methods 34
5.2.1. Analyses for Binary Endpoints 34
5.2.2. Analyses for Continuous Endpoints 36
5.2.3. Analyses for Categorical Endpoints 37
5.2.4. Analyses for Time-to-Event Endpoints 37
5.3. Methods to Manage Missing Data 37
6. ANALYSES AND SUMMARIES 39

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

6.1. Primary Endpoint(s) 39
6.1.1. Proportion of participants achieving centrally adjudicated resolution of NASH without worsening of fibrosis or improvement in fibrosis by ≥1 stage without worsening of NASH, or both, at Week 48 39
6.1.1.1. Main Analysis 39
6.1.1.2. Sensitivity/Supplementary Analyses 40
6.2. Secondary Endpoint(s) 41
6.2.1. Proportion of participants achieving centrally adjudicated resolution of NASH without worsening of fibrosis, at Week 48 41
6.2.1.1. Main Analysis 41
6.2.1.2. Sensitivity/Supplementary Analysis 42
6.2.2. Proportion of participants achieving centrally adjudicated improvement in fibrosis by ≥1 stage without worsening of NASH, at Week 48 42
6.2.2.1. Main Analysis 42
6.2.2.2. Sensitivity/Supplementary Analysis 43
6.2.3. Proportion of participants achieving centrally adjudicated improvement in fibrosis by ≥2 stages without worsening of NASH, at Week 48 43
6.2.3.1. Main Analysis 43
6.2.3.2. Sensitivity/Supplementary Analysis 43
6.2.4. Proportion of participants achieving centrally adjudicated improvement of ≥2 points in Total NAFLD Activity Score (NAS) without worsening of fibrosis, at Week 48 44
6.2.4.1. Main Analysis 44
6.2.4.2. Sensitivity/Supplementary Analysis 44
6.2.5. Percentage change from baseline in liver fat (assessed via MRI-PDFF) at Week 48 44
6.2.5.1. Main Analysis 44
6.2.5.2. Sensitivity/Supplementary Analysis 45
6.3. Other Endpoint(s) 46
6.3.1. Proportion of participants achieving improvement in fibrosis by ≥1 stage independent of changes in NAS, based on assessment by sponsor-identified central pathologist(s), at Week 48 compared to baseline 46
6.3.1.1. Main Analysis 46

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

6.3.2. Proportion of participants achieving resolution of NASH without worsening of fibrosis and improvement in fibrosis by ≥1 stage without worsening of NASH, based on assessment by sponsor-identified central pathologist(s), at Week 48 compared to baseline 46
6.3.2.1. Main Analysis 46
6.3.3. Proportion of participants with progression of fibrosis by ≥1 stage (defined as an increase of at least 1 stage in the Brunt-Kleiner scale compared to baseline) independent of changes in NAS, at Week 48 47
6.3.3.1. Main Analysis 47
6.3.4. Proportion of participants with progression of fibrosis by ≥1 stage and worsening of ≥2 points in Total NAS at Week 48 48
6.3.4.1. Main Analysis 48
6.3.5. Proportion of participants with Worsening of ≥2 points in Total NAS (defined as an increase of at least 2 points in Total NAS compared to baseline) independent of changes in fibrosis at Week 48 49
6.3.5.1. Main Analysis 49
6.3.6. Other Continuous Tertiary/Exploratory Endpoints 50
6.3.6.1. Main Analysis 51
6.3.7. Event-based Tertiary/Exploratory Endpoints 52
6.3.7.1. Main Analysis 52
6.3.8. Pharmacokinetic (PK) Endpoints 53
6.3.8.1. Main Analysis 53
6.4. Subset Analyses 54
6.5. Baseline and Other Summaries and Analyses 54
6.5.1. Baseline Summaries 54
6.5.2. Study Conduct and Participant Disposition 54
6.5.3. Study Treatment Exposure 54
6.5.4. Medical History 54
6.5.5. Concomitant Medications and Nondrug Treatments 55
6.6. Safety Summaries and Analyses 55
6.6.1. Adverse Events 55
6.6.2. Laboratory Data 56
6.6.3. Vital Signs 56
6.6.4. Electrocardiograms 56

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

7. INTERIM ANALYSES 56
7.1. Introduction 56
7.2. Interim Analyses and Summaries 57
8. REFERENCES 58
9. APPENDICES 59

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

LIST OF TABLES

Table 1. Summary of Changes 8
Table 2. Biopsy results to be reported for each participant 21
Table 3. Clinical Laboratory Parameters [redacted] 26
Table 4. Laboratory Assessments (non-fasted) 28
Table 5. Laboratory Assessments (fasted) 28
Table 6. Vital Signs to be Monitored 29
Table 7. ECG Categories 29
Table 8. Analysis Set Definitions 31
Table 9. Summary of Analysis Methods 33
Table 10. Handling non-evaluable histological change-from-baseline data 38

LIST OF FIGURES

Figure 1. Study Design 18

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

APPENDICES

Appendix 1. Summary of Efficacy Analyses 59
Appendix 2. Data Derivation Details 67
Appendix 2.1. Definition and Use of Visit Windows in Reporting 67
Appendix 2.2. Endpoint Derivations 70
Appendix 2.2.1. % Liver Fat using MRI-PDFF and CCI 70
Appendix 2.2.2. MRI-derived Liver Volume 71
Appendix 2.2.3. MRI-derived CCI 71
Appendix 2.2.4. Fasting Plasma Glucose and Fasting Plasma Insulin (T2DM Participants Only) 71
Appendix 2.2.5. Homeostatic Model Assessment for Insulin Resistance (T2DM Participants Only) 71
Appendix 2.3. Data Derivation for Patient Reported Outcomes 72
Appendix 2.3.1. NASH-CHECK 72
Appendix 2.3.2. PROMIS Fatigue Custom 9-Item Version 76
Appendix 2.3.3. NASH Symptom Diary 76
Appendix 2.4. Data Derivation for Non Standard Safety Endpoints 77
Appendix 2.4.1. ELF Test 77
Appendix 2.4.2. MELD score 78
Appendix 2.4.3. Estimated Glomerular Filtration Rate 79
Appendix 2.4.4. AUDIT Total Score 79
Appendix 2.5. Definition of Protocol Deviations That Relate to Statistical Analyses/Populations 81
Appendix 2.5.1. Compliance 81
Appendix 3. Data Set Descriptions 81
Appendix 4. Statistical Methodology Details 81
Appendix 5. List of Abbreviations 82

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

1. VERSION HISTORY

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
Version 3/ 03 May 2024	N/A	Clarification of analyses or additional details added following Draft Table Reviews (DTRs/BDRs) and prior to Database Release	Changed 'participants' to 'patients' in definition of estimands when describing the wider patient population and clarified disease confirmation will be based on on eligibility assessment of SCR2 biopsy (Sections 2.1.1, 2.1.2, 2.13). Added clarification how to handle endpoint derivation for participants who withdraw treatment for other reasons but have evaluable biopsy data at withdrawal or Week 48 and participants with no Week 48 biopsy data (Sections 2.1.2, 2.1.3) Added exploratory Estimands 3.2 and 3.3, and their corresponding endpoint definition and analysis (Sections 2.1.3, 3.3, 6.3, Appendix 1) Added the drug names of the investigational products and the corresponding total daily dose of BID dose arms (Section 2.2) Added clarification for primary endpoint derivation (Section 3.1) Added clarification for secondary endpoints derivation (Section 3.2) Added clarification for exploratory endpoints (Section 3.3) Added allowable analysis windows for FibroScan® and lab results (Section 3.4, Appendix 4.1)

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
			Added Flag Level categories for eGFR in Table 3 Clincial Laboratory Parameters [redacted] (Section 3.5.2) Removed derived non-HDLC in hematology analytes as direct VLDL is being collected instead and added other Biomarkers (or components) that should be based on samples collected in a fasted state(Section 3.5.2) Added categories for heart rate in ECG criteria of clinical concern (Section 3.5.3) Updated missing data handling for hypothetical estimands (Sections 4, 5.3, Appendix 1) Changed 'CI' to 'credible interval' for results obtained from Bayesian Emax model (Sections 5.2.1, 5.2.2, 6.1.1, 6.2.1, 6.2.2, 6.2.3, 6.2.4, 6.3.3, 6.3.5) Changed '95%' to 90% CIs/credible intervals except for the 3-tier approach for the analysis of adverse events (Sections 5.2.1, 5.2.2, 6.1.1, 6.2.1 to 6.2.5, 6.3.1 to 6.3.6) Added clarification that Total Daily Dose will be used in Bayesian Emax modeling (Section 5.2.1) Added estimation of risk difference in treatment comparisons (Sections 5.2.1, 6.1.1, 6.2.1 to 6.2.4, 6.3.1 to 6.3.5) Added specification on Bayesian Emax modeling for liver fat (Sections 5.2.2, 6.2.5)

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
			Added instructions on how to handle non-evaluable histological change from baseline data (Section 5.3) Added specifications for supplementary analyses of the primary endpoint (Section 6.1.1.2) Added prior estimates for the Bayesian Emax modeling of liver fat (Section 6.2.5) Added specifications for the analyses of other continuous tertiary endpoints (Section 6.3.6) Added a criterion for exclusion of samples in PK summaries (Section 6.3.8) Removed text describing the summary of participant disposition by pre-defined medical history terms (Section 6.5.2) Increased the number of categories for which the number of participants differing numbers with prespecified medical histories will be summarized (Section 6.5.4) Added specification for individual plots to be produced for selected laboratory parameters (Section 6.6.2) Added references to estimands and additional exploratory endpoints to the Summary of Effiacy Analyses (Appendix 1)

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
			Added clarifications in analysis visit windows (Appendix 2.1) Added clarification for % liver fat derivation (Appendix 2.2.1) Added specification for the derivation of NASH Symptom Diary Score (Section 2.3.3) Added specification for the derivation of MELD score (Appendix 2.4.2) Added specification for the derivation of AUDIT Total Score (Appendix 2.4.4)
Version 2/ 06 Apr 2022	Amendment 1 30 August 2021	Clarification of analyses or additional details added to prior version and updates to reflect Protocol Amendment 1	Updated the definition of estimands by removing the phrase "**the study population of randomized and treated**" in the description of the target inferential population (Sections 2.1.1, 2.1.2, 2.1.3, 3.1) Throughout the document, ensured that the term **participants** has been used when referring to subjects consented into this study Added exploratory estimand to assess endpoints related to worsening in liver biopsy components as well as capture qualitative assessment on each of 4 domains of histological assessments (Sections 2.1.3, 6.3.1, 6.3.2, Appendix 1) Applied changes made in Protocol Amendment 1 related to the discontinued evaluation of 2 arms with administration of DGAT2i alone (150

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
			mg QD and 300 mg QD) (Sections 2.2, 4, 4.1, 5.1, 5.2, 5.2.1) Updated the hypothetical estimand definition to include information on censored data (Sections 2.1.2, 2.1.3, 4, Appendix 1) Added additional details about the definition and derivation of the primary endpoint, and biopsy assessments (Section 3.1) Added CCI [redacted] and CCI [redacted] as exploratory endpoints (Section 3.3, 6.3.3, Appendix 2.2.1, Appendix 2.2.3) Added derivation of baseline vital signs (Section 3.4) Added clarification on the definition of treatment emergent adverse events (Section 3.5.1) Added clarification on laboratory parameters (Section 3.5.2 and Appendix 2.4) Updated the formula for QTcF, definition of changes in background medications of special interest, data to be summarized for AUDIT, and analysis of DILI (Section 3.5.3) Added exclusion of participants from Evaluable and Safety Analysis Sets for issues related to quality of protocol execution leading to a Sponsor decision to prematurely withdraw randomized participants and close site(s) (Section 4)

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
			Updated the definition of Enrolled, Evaluable and Safety Set to explain how participants originally randomized to the 2 DGAT2i QD groups will be handled (Section 4) Added defined analysis data set for biopsy-based composite estimands related to worsening in liver biopsy components and updated the definition of defined analysis data set for hypothetical estimands (Section 4) Corrected the derivation of percent change from baseline in the analysis of log-transformed continuous endpoints to first determine the relative change from baseline (Section 5.2.2, 6.2.5.1, 6.3.3.1) Updated the description of MMRM analysis (Sections 5.2.2, 6.3.3.1) Added derivation of date of birth (Section 5.3) Added supplementary analyses to assess quality of biopsy assessments (6.1.1.2) Updated the list of event-based exploratory endpoints that will be summarized based on the pre-specified events undergoing adjudication (Section 6.3.4.1) Removed subset analysis on LFT endpoints in participants with ALT>ULN at baseline (Section 6.4) Added demographic and baseline characteristics variables to be summarized (Section 6.5.1)

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
			Updated description of baseline and other summaries (Section 6.5) Changed the percentages of randomized participants to be used to trigger the interim analyses (Section 7.1) Clarified data inclusion and handling of intercurrent events and missing data in efficacy analyses (Appendix 1) Updated definition and use of visit windows (Appendix 2.1) Incorporated additional clarification to endpoint derivations (Appendix 2.2)
Version 1/ 31 Mar 2020	Original 22 Jan 2020	N/A	N/A

2. INTRODUCTION

This Statistical Analysis Plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study C2541013 (MIRNA). This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

2.1. Study Objectives, Endpoints, and Estimands

2.1.1. Primary Estimand(s)

- **Estimand 1.1**: *Using a composite estimand strategy in* patients *with biopsy-confirmed NASH and fibrosis* on eligibility assessment of SCR2 biopsy, *estimate the treatment effect of* Diacylglycerol acyltransferase 2 inhibitor (*DGAT2i) doses administered alone and coadministration of DGAT2i* + Acetyl-CoA carboxylase inhibitor (*ACCi) relative to placebo, and coadministration of DGAT2i + ACCi, relative to DGAT2i alone, in terms of the proportion of 'clinical responders', defined as participants achieving centrally adjudicated: resolution of* Nonalcoholic steatohepatitis *(NASH) without worsening of fibrosis or improvement in fibrosis by ≥1 stage without worsening of NASH, or both, at Week 48 and treating all cases of withdrawal from treatment for lack of efficacy or toleration as non- responders.*

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

The main intercurrent event of withdrawal from treatment for lack of efficacy or toleration is addressed as part of the responder variable definition, therefore this is not handled separately. Participants who withdraw from treatment for other reasons but have evaluable biopsy data at withdrawal or at Week 48 will have their biopsy data assessed to determine whether they are responders or not. Participants with no Week 48 biopsy data will be considered to be non-responders. All treatment effect contrasts will be obtained through fitting the Bayesian dose-response (DR) and/or exposure-response (ER) models with a logit transformation, or logistic regression models.

2.1.2. Secondary Estimand(s)

Secondary estimands to assess secondary objectives will be:

- **Estimand 2.1***: Using a composite estimand strategy, in* patients *with biopsy-confirmed NASH and fibrosis,* on eligibility assessment of SCR2 biopsy *as assessed by the sponsor-identified central pathologist(s), estimate the treatment effect of DGAT2i doses administered alone and coadministration of DGAT2i + ACCi relative to placebo, and coadministration of DGAT2i + ACCi, relative to DGAT2i alone, in terms of the proportion of 'clinical responders', defined as participants achieving centrally adjudicated resolution of NASH without worsening of fibrosis, at Week 48 and treating all cases of withdrawal from treatment for lack of efficacy or toleration as non-responders.* Participants who withdraw from treatment for other reasons but have evaluable biopsy data at withdrawal or at Week 48 will have their biopsy data assessed to determine whether they are responders or not. Participants with no Week 48 biopsy data will be considered to be non-responders. *All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models with a logit transformation, or logistic regression models.*

- **Estimand 2.2**: *Using a composite estimand strategy, in* patients *with biopsy-confirmed NASH and fibrosis,* on eligibility assessment of SCR2 biopsy *as assessed by the sponsor-identified central pathologist(s), estimate the treatment effect of DGAT2i doses administered alone and coadministration of DGAT2i + ACCi relative to placebo, and coadministration of DGAT2i + ACCi, relative to DGAT2i alone, in terms of the proportion of 'clinical responders', defined as participants achieving centrally adjudicated improvement in fibrosis by ≥1 stage without worsening of NASH, at Week 48 and treating all cases of withdrawal from treatment for lack of efficacy or toleration as non-responders.* Participants who withdraw from treatment for other reasons but have evaluable biopsy data at withdrawal or at Week 48 will have their biopsy data assessed to determine whether they are responders or not. Participants with no Week 48 biopsy data will be considered to be non-responders. *All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models with a logit transformation, or logistic regression models.*

- **Estimand 2.3**: *Using a composite estimand strategy, in* patients *with biopsy-confirmed NASH and fibrosis,* on eligibility assessment of SCR2 biopsy *as assessed by the sponsor-identified central pathologist(s), estimate the treatment effect of DGAT2i doses administered alone and coadministration of DGAT2i + ACCi relative to placebo, and*

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

coadministration of DGAT2i + ACCi, relative to DGAT2i alone, in terms of the proportion of 'clinical responders', defined as participants achieving centrally adjudicated improvement in fibrosis by ≥2 stages without worsening of NASH, at Week 48 and treating all cases of withdrawal from treatment for lack of efficacy or toleration as non-responders. Participants who withdraw from treatment for other reasons but have evaluable biopsy data at withdrawal or at Week 48 will have their biopsy data assessed to determine whether they are responders or not. Participants with no Week 48 biopsy data will be considered to be non-responders. *All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models with a logit transformation, or logistic regression models.*

- **Estimand 2.4**: *Using a composite estimand strategy, in* patients *with biopsy-confirmed NASH and fibrosis,* on eligibility assessment of SCR2 biopsy *as assessed by the sponsor-identified central pathologist(s), estimate the treatment effect of DGAT2i doses administered alone and coadministration of DGAT2i + ACCi relative to placebo, and coadministration of DGAT2i + ACCi, relative to DGAT2i alone, in terms of the proportion of 'clinical responders', defined as participants achieving centrally adjudicated improvement of ≥2 points in Total* Nonalcoholic fatty liver disease *(NAFLD) Activity Score (NAS)* without worsening of fibrosis *at Week 48 and treating all cases of withdrawal from treatment for lack of efficacy or toleration as non-responders.* Participants who withdraw from treatment for other reasons but have evaluable biopsy data at withdrawal or at Week 48 will have their biopsy data assessed to determine whether they are responders or not. Participants with no Week 48 biopsy data will be considered to be non-responders. *All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models with a logit transformation, or logistic regression models.*

- **Estimand 2.5**: *Using a hypothetical estimand strategy, in* patients *with biopsy-confirmed NASH and fibrosis,* on eligibility assessment of SCR2 biopsy *as assessed by the sponsor-identified central pathologist(s), estimate the treatment effect of DGAT2i doses administered alone and coadministration of DGAT2i + ACCi relative to placebo, and coadministration of DGAT2i + ACCi, relative to DGAT2i alone, in terms of population level average percentage change from baseline in liver fat (assessed via* Magnetic resonance imaging using proton density fat fraction acquisition *(MRI-PDFF)) at Week 48, assuming all participants had remained in the trial and received treatment as planned without withdrawal up to 48 weeks. Endpoint data collected after treatment withdrawal will be censored. All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models, or* Analysis of covariance (*ANCOVA*) *models.*

2.1.3. Additional Estimand(s)

A secondary estimand to assess the primary objective will be:

- **Estimand 1.2**: Using a hypothetical estimand strategy, *in* patients *with biopsy-confirmed NASH and fibrosis,* on eligibility assessment of SCR2 biopsy *as assessed by the sponsor-identified central pathologist(s), estimate the treatment effect of DGAT2i doses*

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

administered alone and coadministration of DGAT2i + ACCi relative to placebo, and coadministration of DGAT2i + ACCi, relative to DGAT2i alone, in terms of estimating the proportion of 'clinical responders', defined as participants achieving centrally adjudicated: resolution of NASH without worsening of fibrosis or improvement in fibrosis by ≥1 stage without worsening of NASH, or both, at Week 48, assuming all patients had remained in the trial and received treatment as planned without withdrawal up to 48 weeks. Endpoint data collected after treatment withdrawal will be censored. All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models with a logit transformation, or logistic regression models.

This estimand assesses the treatment effect in an alternative, hypothetical setting of adherence to their treatment as planned.

An exploratory estimand to assess endpoints related to worsening in liver biopsy endpoint components at Week 48 is also defined as follows:

- **Estimand 3.1:** Using a composite estimand strategy, in patients with biopsy-confirmed NASH and fibrosis, on eligibility assessment of SCR2 biopsy as assessed by the sponsor-identified central pathologist(s), estimate the treatment effect of DGAT2i doses administered alone and coadministration of DGAT2i + ACCi relative to placebo, and coadministration of DGAT2i + ACCi, relative to DGAT2i alone, in terms of the proportion of 'clinical responders' (for the purposes of analysis), defined as participants with worsening in liver biopsy endpoint components at Week 48 and treating all cases of withdrawal from treatment for lack of efficacy or toleration also as 'responders'. Participants who withdraw from treatment for other reasons but have evaluable biopsy data at withdrawal or at Week 48 will have their biopsy data assessed to determine whether they are 'responders' or not. Participants with no Week 48 biopsy data will be considered to be 'responders'. Participants who improve or remain the same compared to baseline in their liver biopsy endpoint components are treated as 'non-responders', as response for these endpoints is defined as a worsening in liver biopsy endpoint components. All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models with a logit transformation, or logistic regression models.

Finally, exploratory estimands to assess endpoints related to additional definitions of 'clinical response' are defined as follows:

- **Estimand 3.2**: Using a composite estimand strategy, in patients with biopsy-confirmed NASH and fibrosis, on eligibility assessment of SCR2 biopsy as assessed by the sponsor-identified central pathologist(s), estimate the treatment effect of DGAT2i doses administered alone and coadministration of DGAT2i + ACCi relative to placebo, and coadministration of DGAT2i + ACCi, relative to DGAT2i alone, in terms of the proportion of 'clinical responders', defined as participants achieving improvement in fibrosis by ≥1 stage independent of changes in NAS, at Week 48 and treating all cases of withdrawal from treatment for lack of efficacy or toleration as non-responders.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

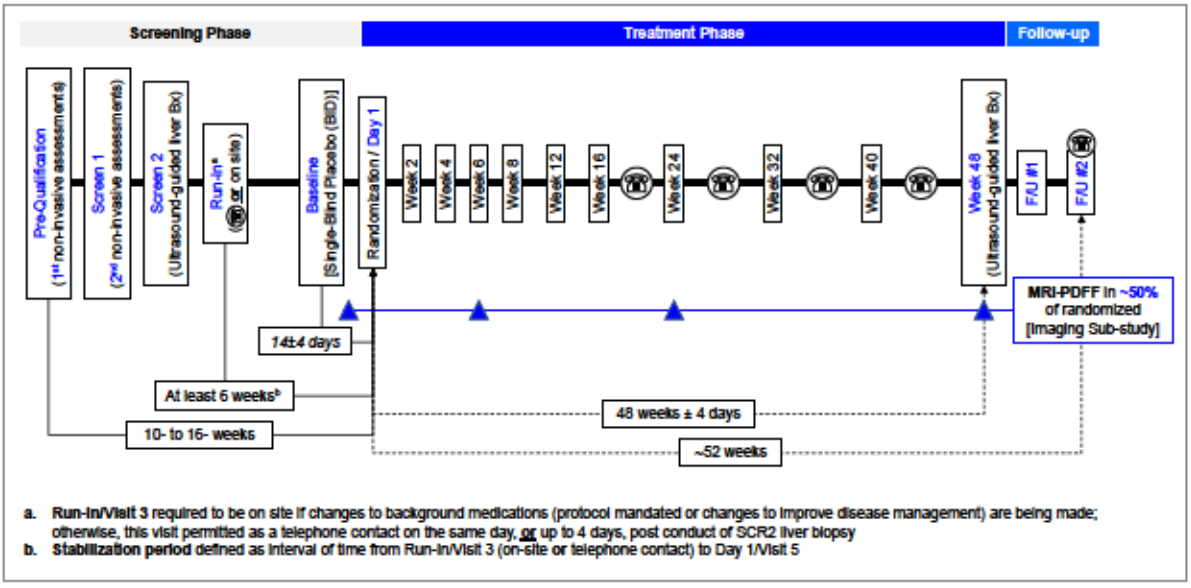
Participants who withdraw from treatment for other reasons but have evaluable biopsy data at withdrawal or at Week 48 will have their biopsy data assessed to determine whether they are responders or not. Participants with no Week 48 biopsy data will be considered to be non-responders. All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models with a logit transformation, or logistic regression models.

- **Estimand 3.3**: Using a composite estimand strategy, in patients with biopsy-confirmed NASH and fibrosis, on eligibility assessment of SCR2 biopsy as assessed by the sponsor-identified central pathologist(s), estimate the treatment effect of DGAT2i doses administered alone and coadministration of DGAT2i + ACCi relative to placebo, and coadministration of DGAT2i + ACCi, relative to DGAT2i alone, in terms of the proportion of 'clinical responders', defined as participants achieving resolution of NASH without worsening of fibrosis and improvement in fibrosis by ≥1 stage without worsening of NASH, at Week 48 and treating all cases of withdrawal from treatment for lack of efficacy or toleration as non-responders. Participants who withdraw from treatment for other reasons but have evaluable biopsy data at withdrawal or at Week 48 will have their biopsy data assessed to determine whether they are responders or not. Participants with no Week 48 biopsy data will be considered to be non-responders. All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models with a logit transformation, or logistic regression models.

2.2. Study Design

The detailed set-up of the study is summarized below –

Figure 1. Study Design



090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

This is a multicenter, multinational, randomized, double-blind, double-dummy, placebo-controlled, parallel-group, evaluation of DGAT2i alone, and DGAT2i+ACCi. Approximately 350 participants (50 per arm) with biopsy confirmed NASH with F2-F3 fibrosis, as assessed by sponsor-identified central pathologist(s), will be randomly assigned to the study intervention to ensure approximately 280 evaluable participants (40 per arm) offer evaluable data for the primary objective.

Participants will be randomized in a balanced ratio, using a computer-generated randomization code, to receive either double-blind, double-dummy placebo, 1 of 4 doses/dosing regimens of DGAT2i alone, or 1 of 2 dose-levels of DGAT2i+ACCi for a treatment duration of up to 48-weeks.

For participants randomized before Protocol Amendment 1 is enacted, those randomized to either of the 2 QD (DGAT2i 150 mg QD and DGAT2i 300 mg QD) *dosing regimens of DGAT2i alone will remain in the study but will be switched to the corresponding BID dosing regimen of DGAT2i alone, maintaining the assigned total daily dose and the double-blind, double-dummy design.* New randomization to the QD treatment arms will cease, after Health Authority (where applicable) and IRB/EC approvals of Protocol Amendment 1. Therefore, following Protocol Amendment 1, participants will be randomized to one of the following treatment arms:

- Placebo

DGAT2i (Ervogastat) BID dose arms (for characterization of dose-response):

- DGAT2i 25 mg BID, total daily dose of 50 mg
- DGAT2i 75 mg BID, total daily dose of 150 mg
- DGAT2i 150 mg BID, total daily dose of 300 mg
- DGAT2i 300 mg BID, total daily dose of 600 mg

DGAT2i (Ervogastat) + ACCi (Clesacostat) coadministration dose arms:

- DGAT2i 150 mg BID + ACCi 5 mg BID
- DGAT2i 300 mg BID + ACCi 10 mg BID

The treatment group labels in programmed outputs will be as follows:

Placebo	PF-5571 25 mg BID	PF-5571 75 mg BID	PF-5571 150 mg BID	PF-5571 150 mg BID + PF1304 5 mg BID	PF-5571 300 mg BID	PF-5571 300 mg BID + PF1304 10 mg BID	Total

Following footnote to be included on all tables: PF5571 = Ervogastat and PF1304 = Clesacostat
Note: A 'Total' column is required when summarizing baseline of population randomized

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

The 'PF5571 75 mg BID' group will also include those randomized to DGAT2i 150 mg QD and 'PF5571 150 mg BID' group will also include those randomized to DGAT2i 300 mg QD, therefore, there will be no reference to the QD treatment arms in any of the programmed outputs (including dose at onset for AEs and treatment duration – these will only refer to the BID groups as described above).

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint(s)

The Primary Endpoint is the *proportion of participants achieving resolution of NASH without worsening of fibrosis or improvement in fibrosis by ≥1 stage without worsening of NASH or both, based on assessment by sponsor-identified central pathologist(s), at Week 48,* compared to baseline.

In order to derive this endpoint, the components are defined as follows:

- Resolution of NASH: disappearance of ballooning (NAS ballooning score=0) and residual or no lobular inflammation (NAS lobular inflammation score of 0 or 1) and NAS steatosis score of 0, 1, 2 or 3.
- No worsening of fibrosis: no change or a decrease of at least 1 stage in the Brunt-Kleiner scale compared to baseline

and

- Improvement in fibrosis by ≥1 stage: a decrease of at least 1 stage in the Brunt-Kleiner scale compared to baseline
- No worsening of NASH: no change or no increase in NAS for ballooning, inflammation, or steatosis compared to baseline

For all biopsy endpoints evaluating change from baseline, assessment will be based on the paired double pathologist review of digitized slides. As a result of the final paired double pathologist review, if a participant is found not to be eligible for the study (e.g. if they are re-categorized as F1 or F4), they will still be included in the analysis. If there has been a different assignment of biopsy grading (e.g. if the final assignment of the baseline biopsy is F3 but at the time of eligibility determination this same biopsy was assessed as F2), then participants will be included in any analyses based on the final paired assessment, rather than the one at the time of eligibility determination, however, the fibrosis grade at eligibility determination (SCR2 Eligibility) will be used as a covariate, as this is what will be used for stratification. Planned biopsy assessments are summarized in Table 2.

- There will be two types of biopsy assessments conducted by the pathologists:
  - Quantitative: NAS grading and fibrosis staging
  - Qualitative: In the blinded side-by-side review, an assessment of which biopsy (SCR2 or Week 48/DC) is considered to be better, worse or are both the same

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 2. Biopsy results to be reported for each participant

		Blinded side-by-side (Change from Baseline) review of SCR2 and On-treatment biopsies **Limited to only those with on-treatment biopsy**	
	SCR2 (for Eligibility: **quantitative assessment only**)	SCR2 (for Analysis: **quantitative and qualitative assessments**)	On treatment - Wk 48 or D/C (for Analysis: **quantitative and qualitative assessments**)
Pathologist 1 parallel review	All participants	All participants with on-treatment biopsy	All participants with on-treatment biopsy
Pathologist 2 parallel review	All participants	All participants with on-treatment biopsy	All participants with on-treatment biopsy
Consensus review	If needed[1]	If needed, for both quantitative and qualitative comparisons[2]	If needed, for both quantitative and qualitative comparisons[2]

1. Consensus review to reach agreement is *not* required if both central pathologists agree that (a) either the NAS or Fibrosis grade renders participant ineligible; (b) total NAS score ≥4, with ≥1 for steatosis, inflammation, and ballooning, is met, even if alignment of score for each component of NAS [steatosis, inflammation, ballooning] is absent; **and/or** (c) fibrosis stage is aligned for higher group [F3 or F4] but may not be aligned for sub-group [F3a, F3b, F4a, F4b]

2. Consensus review to reach agreement ***required*** for all cases of misalignment on parallel review – ie, misalignment on scoring of each component of NAS score, sub-group of fibrosis stage [F3a, F3b, F4a or F4b], qualitative assessment of better, same, worse

Note that Trichrome Stain should only be used for the Fibrosis results and the Hemotoxylin and Eosin Stain should be used for the NAS results, only. Additionally, in change-from-baseline assessment, if there is a Consensus result (i.e. 'PATHOLOGIST 1 AND PATHOLOGIST 2') then this is the one that should be used, and if a Consensus result is not reported, the PATHOLOGIST 1 and PATHOLOGIST 2 results should be identical for all 4 domains (3 related to NAS + 1 related to fibrosis), so either PATHOLOGIST 1 or PATHOLOGIST 2 results can be used.

- Mapping of Fibrosis Grades for analysis:

Acceptable Values	Labcorp Result	Fibrosis Grade to use for analysis
0	0 - No fibrosis	0
1a	1a - Zone 3 mild perisin fib	1
1b	1b - Zone 3 mod perisin fib	1
1c	1c - Periportal/portal fib	1
2	2 - Zone 3+ periport/port fib	2

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Acceptable Values	Labcorp Result	Fibrosis Grade to use for analysis
3a	3a - Bridging fibrosis (few)	3
3b	3b - Bridging fibrosis (num)	3
4a	4a - Cirrhosis (incomplete)	4
4b	4b - Cirrhosis (established)	4
NON EVALUABLE	Not evaluable	Not evaluable

3.2. Secondary Endpoint(s)

Secondary Endpoints are:

- *Proportion of participants achieving improvement in different responder definitions based on assessment by sponsor-identified central pathologist(s) at Week 48* compared to baseline (each defined as above for the primary endpoint):
    - *Resolution of NASH, without worsening of fibrosis*
    - *Improvement in fibrosis by ≥1 stage, without worsening of NASH*
    - *Improvement in fibrosis by ≥2 stages* (defined as a decrease of at least 2 stages in the Brunt-Kleiner scale compared to baseline), *without worsening of NASH*
    - *Improvement of ≥2 points in Total NAS* (defined as a decrease of at least 2 points in Total NAS compared to baseline), without progression of fibrosis

      Total NAS can range from 0 to 8 and is calculated as the sum of scores of steatosis (0 to 3), lobular inflammation (0 to 3) and ballooning (0 to 2). If any of the sub-scale scores are non-evaluable/missing, then the total score should be derived as missing.
- *Percent change in liver fat (assessed via MRI-PDFF in substudy population), at Week 48*
- *Assessment of* Treatment Emergent Adverse Events *(TEAEs), safety-related clinical laboratory tests, vital signs, and 12-lead* Electrocardiograms *(ECGs), over time up to Week 48*

3.3. Other Endpoint(s)

- Proportion of participants achieving improvement in fibrosis by ≥1 stage independent of changes in NAS, based on assessment by sponsor-identified central pathologist(s), at Week 48 compared to baseline
- Proportion of participants achieving resolution of NASH without worsening of fibrosis and improvement in fibrosis by ≥1 stage without worsening of NASH, based on assessment by sponsor-identified central pathologist(s), at Week 48 compared to baseline

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

- *Proportion of participants with worsening in liver biopsy endpoint components based on assessment by sponsor-identified central pathologist(s) at Week 48:*
    - *Progression of fibrosis by ≥1 stage* (defined as an increase of at least 1 stage in the Brunt-Kleiner scale compared to baseline) independent of changes in NAS
    - Progression of fibrosis by ≥1 stage and worsening of ≥2 points in Total NAS
    - *Worsening of ≥2 points in Total NAS* (defined as an increase of at least 2 points in Total NAS compared to baseline) independent of changes in fibrosis

Other tertiary/exploratory endpoints are as shown below. The analyses of these endpoints are assessed for efficacy up to Week 48, but the analyses of blood based biomarkers will also include data up to and including Week 50 to allow an assessment of off-treatment effects::

- Percent change in liver fat (assessed via MRI-PDFF in Imaging substudy population), over time up to Week 48
- Percent change in liver volume (assessed via Magnetic resonance imaging (MRI) in Imaging substudy population), over time up to Week 48
- Percent change in CCI [redacted] (assessed via MRI in Imaging substudy population), over time up to Week 48
- Percent change in CCI [redacted] (assessed via MRI in Imaging substudy population), over time up to Week 48
- Percent change from baseline in liver fat as assessed via Controlled attenuation parameter (CAP™) measure by FibroScan®, (entire study population), over time up to Week 48
- Percent change from baseline in liver stiffness (assessed via Vibration-Controlled Transient Elastography (VCTE™) using FibroScan® (entire study population)), over time up to Week 48
- Percent change from baseline in Liver Function Tests – Alanine aminotransferase (ALT), Aspartate aminotransferase (AST), Alkaline phosphatase (ALP), Gamma-glutamyl transferase (GGT), Total Bilirubin, Direct Bilirubin, and Total Bile Acids over time up to Week 50
- Percent change from baseline, overtime up to Week 50 in Mechanism-Related and Metabolic Parameters – Proprotein convertase subtilisin/kexin type 9 (PCSK9), Adiponectin, Platelet count, eGFR-CKD-EPI-Cystatin-C, and Body Weight
- Percent change from baseline, overtime up to Week 50 in Cytokeratin-18-M30 fragment (CK18-M30) and Cytokeratin-18-M65 fragment (CK18-M65)

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

- Percent change from baseline, overtime up to Week 50 in N-terminal propeptide of type III procollagen (ProC3) and C-terminal fragment of α3 chain of procollagen type VI (ProC6)

- Percent change from baseline, overtime up to Week 50 in Enhanced Liver Fibrosis (ELF) score (including the individual parameters assayed - Serum hyaluronic acid (HA), Serum amino-terminal propeptide of type III procollagen (PIIINP) and Serum tissue inhibitor of metalloproteinases 1 (TIMP-1))

- Percent change from baseline, overtime up to Week 50 in High-sensitive C-reactive protein (hs-CRP)

- Percent change from baseline, overtime up to Week 50 Apolipoprotein A1, B (total), B100, B48, C3, E (Potentially mechanism-related parameters)

- Percent change from baseline, overtime up to Week 50 in Fasting lipid panel (total cholesterol, direct Low density lipoprotein-cholesterol (LDL-C), High density lipoprotein cholesterol (HDL-C), triglycerides, direct Very low-density lipoprotein (VLDL)),

- Change from baseline, overtime up to Week 50 in HbA1C, Fasting plasma glucose (FPG), Fasting plasma insulin (FPI), Homeostatic Model Assessment for Insulin Resistance (HOMA-IR)

- Change from baseline, (adults with T2DM only), over time up to Week 50, in HbA1C, Fasting plasma glucose (FPG), Fasting plasma insulin (FPI), Homeostatic Model Assessment for Insulin Resistance (HOMA-IR)

- Change from baseline, over time to Week 48 in abdominal pain and bloating questions from the NASH Symptom Diary

- Change from baseline, over time to Week 48 in fatigue, sleep disturbance, and daytime sleepiness questions from the NASH Symptom Diary

- Change from baseline, over time to Week 48 in fatigue as measured by Patient-Reported Measurement Outcome Information System (PROMIS) Fatigue questionnaire

- Change from baseline, over time to Week 48 in Health-related quality of life (HRQoL) sub-scale scores as measured by NASH-Check

- Change from baseline, over time to Week 48 in PGI-S

- Absolute score over time in Patient Global Impression of Severity (PGI-S)

- Absolute score over time in Patient Global Impression of Change (PGI-C)

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

3.4. Baseline Variables

The following baseline variable will be included as a factor:

- Baseline fibrosis stage (F2/F3)

For all biopsy based efficacy endpoints, baseline will be defined as the biopsy results obtained at SCR2 or a historical biopsy within 12-weeks prior to SCR2 (though within 24-weeks of Day 1 (Visit 5).

For Liver Fat as measured by MRI-PDFF, baseline is defined as the value of total mean liver fat obtained between Visit 4 and Visit 5 (ie, labeled as baseline).

For Liver Volume as measured by MRI-PDFF, baseline is defined as the value of total mean liver volume obtained between Visit 4 and Visit 5 (ie, labeled as baseline).

Baseline for FibroScan® based endpoints will also be defined as the evaluable values obtained between Visit 4 and Visit 5 **or** SCR1 if there isn't an evaluable value between Visit 4 and Visit 5 (ie, labeled as baseline).

For all Patient Reported Outcomes, any other continuous efficacy endpoints and all safety-related continuous endpoints, baseline will be defined as the closest results obtained prior to dosing on Day 1 (Visit 5)

For laboratory results, the baseline value to be used will be either Day 1 or Day -14 (plus any allowable analysis windows – see Appendix 2.1), but not any further back than this. If laboratory collections have been taken in the correct fasted state (as indicated in the CRF), Day 1 should be used, but if not, then Day -14 should be used (again, if in the correct fasted state for the specific laboratory parameter).

For vital signs, the baseline value to be used will be either Day 1 or Day -14, but not any further back than this. If possible, Day 1 should be used, but if this is missing, then Day -14 should be used. If both Day -14 and Day 1 is missing for a particular parameter, then the baseline should be considered as "Missing".

3.5. Safety Endpoints

3.5.1. Adverse Events

An adverse event will be considered ***treatment emergent*** relative to a given treatment if:

- The event starts on or after the first dosing day of double-blinded study medication on Day 1 [Visit 5], but before the last dose plus the lag time. The algorithm will not consider any events that started prior to the first dose date. If an AE starts on the same day as the first dose date, it will be considered treatment emergent.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

- Any AEs that occur after signing Informed Consent and before randomization and the first dose of double-blind study medication should be recorded in Medical History; with ***only*** treatment emergent AEs reported on the AE CRF page.

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 (i.e. an infinite lag). 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 AEs. Under this approach, AEs are classified into 1 of 3 tiers. Different analyses will be performed for different tiers (see Section 6.6.1).

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

Tier 2 events: These are events that are not tier 1 but are "common." A Medical Dictionary for Regulatory Activities (MedDRA) preferred term (PT) is defined as a Tier 2 event if there are at least 4 participants with at least one occurrence in any treatment group.

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

3.5.2. Laboratory Data

For the specific laboratory parameters listed in Table 3 below, the following endpoints will be evaluated using CDISC and Pfizer standards (CaPS):

- Absolute value and change from baseline at each visit (Week 2, 4, 6, 8, 12, 16, 24, 32, 40, 48 and Follow-up).
- Number of occurrences of these abnormalities defined as "Flag Level" or "Alert Level"
- Number of participants with these abnormalities defined as "Flag Level" or "Alert Level"

Table 3. Clinical Laboratory Parameters CCI [redacted]

Parameter	Flag Level*	Alert Level^	Conventional Units
Fasting Serum Triglycerides	≥400	≥600	mg/dL
	--	≥ 800	mg/dL
Platelet Count	< 100	< 75	10^9/L
Fasting Plasma Glucose	< 70	≤ 49	mg/dL
	≥140	> 270	mg/dL
Alanine aminotransferase	≥ 2x ULN	Pfizer std flag for PCC	IU/L
	> 3x ULN	*Pfizer std flag for PCC*	IU/L
	> 5x ULN	>8x ULN	IU/L

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 3. Clinical Laboratory Parameters CCI [redacted]

Parameter	Flag Level*	Alert Level^	Conventional Units
Aspartate aminotransferase	≥ 2x ULN	*Pfizer std flag for PCC*	IU/L
	> 3x ULN	*Pfizer std flag for PCC*	IU/L
	> 5x ULN	>8x ULN	IU/L
Alkaline Phosphatase	≥ 2x ULN	*Pfizer std flag for PCC*	IU/L
	> 3x ULN	*Pfizer std flag for PCC*	IU/L
	> 5x ULN	*Pfizer std flag for PCC*	IU/L
Gamma Glutamyl Transferase	> ULN	*Pfizer std flag for PCC*	IU/L
Total Bilirubin	> 1.5x ULN	>3x ULN	mg/dL
Direct (Conjugated) Bilirubin	> ULN	*Pfizer std flag for PCC*	mg/dL
Estimated glomerular filtration rate (eGFR)**	<60	*Pfizer std flag for PCC*	mL/min/1.73 m^2
	<45	*Pfizer std flag for PCC*	mL/min/1.73 m^2
	<30	*Pfizer std flag for PCC*	mL/min/1.73 m^2

*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
** eGFR should be based on cystatin-C – see Appendix 2.4 for derivations.

The safety events CCI [redacted] as defined in Table 3 will be summarized and the frequency of laboratory abnormalities (both number of occurrences **and** number of participants).

Across all the laboratory-related analytes (for safety and pharmacodynamics/biomarkers), inclusion of results when assessed at scheduled nominal visits must be (1) collected prior to AM dose of double-blinded study medication, **and** should be summarized as noted **either** for criteria (2) **or** criteria (3):

(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

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 4. Laboratory Assessments (non-fasted)

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

Or

(3) 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 recorded on the laboratory CRF page) as it is well known that fasting status impacts numerical results (or there is insufficient data regarding effect of meals on results).

Table 5. Laboratory Assessments (fasted)

Biomarkers	Lipid-related	Other Blood-related
% CDT Plasma Glucose Plasma Insulin Total Bile Acids CK18-M30 CK18-M65 ProC3 ProC6 Plasma PCSK9 Individual components of ELF Score (ie, HA, PIIINP, TIMP-1) Adiponectin Individual components of HOMA-IR (i.e. Fasting Plasma Glucose and Fasting Plasma Insulin)	Total cholesterol HDL-C Direct LDL-C Triglycerides Direct VLDL Apolipoprotein A1 Apolipoprotein B total Apolipoprotein B100 Apolipoprotein B48 Apolipoprotein C3 Apolipoprotein E	CCI [redacted]

CCI [redacted], using these samples will be summarized in a stand alone report (outside of study-level CSR).

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

3.5.3. Other Standard Safety other than AEs

Individual participants' vital signs meeting the criteria outlined in Table 6 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 6. 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	Beats per minute (bpm)
	> 120	bpm

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

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

Individual participants' ECGs meeting the criteria outlined in Table 7, which align with the Pfizer standard criteria, will be summarized by randomized arm, as part of standard safety-related outputs.

Table 7. ECG Categories

Parameter	Flag Level*
Maximum Post-dose Corrected QT (Fridericia method) (QTcF) (msec)	≤ 450
	450 - ≤480
	480 - ≤500
	> 500
Maximum Increase from Baseline in QTcF (msec)	≤30
	30 - ≤60
	> 60
Maximum Post-dose PR (msec)	≥300
Maximum Increase from Baseline in PR (msec)	Baseline >200 and max ≥25% increase
	Baseline ≤200 and max ≥50% increase
Maximum Post-dose QRS (msec)	≥140
Maximum Increase from Baseline in QRS (msec)	≥50% increase
Heart Rate (bpm)	<40 >120

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

⇨ QTcF –derived by programming as follows –

QTcF (in msec) = [QT (in msec)/[cube root of RR interval (in sec)] with RR (in sec) = RR (in msec) x 1000

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

In addition, changes in background medications (i.e. final Total Daily Dose (TDD) in the study, i.e. End of Treatment, compared to Baseline (pre-Day 1) TDD) for agents used for glycemic control, lipid control, or blood pressure, and *selected* other specific background medications will be summarized as follows:

- Dose increase and decrease in concomitant medication for glycemic control, i.e. medications for
    - Type 1 and 2 Diabetes
    - Metformin dose categories will also be summarized by visit and treatment group
    - The relevant data can be obtained from the CM001_1 page (Anti-diabetic agents category) from the eCRF.
- Dose increase and separately decrease in concomitant lipid-modifying medications, i.e. medications for
    - HDL-C decreased
    - LDL-C increased
    - Triglycerides increased
    - The relevant data can be obtained from the CM001_2, 3 and 5 pages (Hypercholesterolemia category) from the eCRF.
- Dose increase and separately decrease in concomitant blood pressure management medications, i.e. medications for
    - Hypertension
    - The relevant data can be obtained from the CM001_4 pages (Anti-hypertension agents category) from the eCRF.
- Dose increase and separately decrease in concomitant therapies relating to:
    - Hypothyroidism
    - The relevant data can be obtained from the CM001_6 pages (Hypothyroidism medications category) from the eCRF.

To assess for changes in alcohol consumption during the study, changes from baseline in the Alcohol use disorders identification test (AUDIT) questionnaire total score and the Absolute Value in Percent carbohydrate deficient transferrin relative to total transferrin (%CDT), will also be summarized descriptively over time.

Drug-Induced Liver Injury (DILI) and Hy's Law Criteria

A summary table will categorize the number and percentage of participants in each treatment group who meet the CCI criteria, defined differently for participants with Normal baseline

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

and Abnormal baseline (see Appendix 6 of protocol for definition of potential Hy's Law and Potential DILI)

Participants meeting the AST/ALT and TBili criteria will be categorized as potential Hy's Law cases, and participants who meet the AST/ALT criteria but not the TBili criteria will be categorized as potential DILI cases.

eDish plots will be produced (Peak ALT vs Peak TBili and Peak AST vs Peak TBili), with sectors delineated by the following lines to indicate elevations in ALT, AST and TBili:

- two vertical lines at the point at which Peak ALT (or AST) = 3 x ULN and at 8 x ULN
- a horizontal line at the point at which Peak Total Bilirubin = 1 x ULN and at 2 x ULN

Separate plots will be produced for each treatment group, with males and females identified using different symbols.

4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS)

For purposes of analysis, the following analysis sets are defined:

Note that participants randomized at sites terminated by the Sponsor for Good Clinical Practice (GC) non-compliance will not have their data included in any summary outputs or analyses (i.e. they will be excluded from the Evaluable and Safety Analysis Sets), but their data will be retained in listings as part of the All Randomized Participants set below.

Table 8. Analysis Set Definitions

***Participants** Analysis Set*	***Description***
Enrolled	*All participants who sign the* CCI *ICD* and are enrolled into the Main study
Randomly assigned to investigational product	*All randomized participants*
Evaluable (i.e. Full Analysis Set)	*Defined according to the Full Analysis Set:* *All randomized participants who take at least 1 dose of investigational product who have provided baseline data for the primary endpoint (i.e.* evaluable *baseline biopsy data). Participants will be analyzed according to the treatment group they are randomized to.* • *The only exception to this is for participants randomized before Protocol Amendment 1 is enacted who are randomized to either of the 2 QD dosing regimens of DGAT2i alone, will be analyzed under the corresponding DGAT2i BID treatment group.*
Safety	*All participants who take at least 1 dose of investigational product. Participants will be analyzed according to the treatment they actually received.* • *The only exception to this is for participants randomized before Protocol Amendment 1 is enacted who are randomized to either of*

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 8. Analysis Set Definitions

	the 2 QD dosing regimens of DGAT2i alone, will be analyzed under the corresponding DGAT2i BID treatment group.
Defined Analysis Data Set (at the data level) – endpoint specific	***Description***
All biopsy-derived endpoints	
Biopsy-based Composite Estimand	*If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have* evaluable *Week 48 Biopsy data, they will be considered to be a non-responder*. If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable biopsy data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.
Biopsy-based Composite Estimand for 'Worsening from Baseline' exploratory liver biopsy endpoints	If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, they will be considered to be a responder. If a participant improves or remains the same compared to baseline, then that participant will be termed as a non-responder. If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable biopsy data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.
Biopsy-based Hypothetical *Estimand*	*Any available biopsy data for all participants collected before or at the time of treatment withdrawal will be included in the analysis, even if they have withdrawn from treatment for lack of efficacy or toleration. Endpoint data collected after treatment withdrawal will be censored.* This analysis will be carried out under the assumption that missing data at Week 48 are missing at random conditional on dose and the strata formed by the covariates
All MRI-PDFF-derived endpoints (in Imaging Substudy)	
MRI-based Hypothetical *Estimand*	*Any available MRI data for all participants collected before or at the time of treatment withdrawal will be included in the analysis, even if they have withdrawn from treatment for lack of efficacy or toleration. Endpoint data collected after treatment withdrawal will be censored.* Missing data at Week 48 will not be imputed. MMRM adjusted for covariates will be used to obtain LS mean estimate at Week 48 for each dose, which will then be used in fitting the Bayesian Emax model

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

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

4.1. Treatment Misallocations

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

- Randomized but not treated (i.e. if the participant has a randomization number but no treatment or PK data), then the participant will be excluded from all the statistical analyses.
- Randomized but took incorrect treatment, then the participant will be reported under his/her randomized group for all efficacy analyses. The participant will be reported for safety and Pharmacokinetic (PK) analysis under his/her actual randomized arm.
- Participants randomized to the incorrect stratum at baseline will be allocated to the correct stratum for analyses using SCR2 biopsy eligibility results collected in the study.
- Treated but not randomized, then by definition the participant 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 and PK analyses.

Note: Participants randomized to either of the 2 QD dosing regimens of DGAT2i will be analyzed according to the corresponding BID dosing regimen of DGAT2i alone (maintaining the same total daily dose). If a participant has been re-screened then they will only be counted once in all summaries under the original unique SSID#. Information collected on the Demography page of the CRF can be used to determine any previous participant ID that a participant may have been screened under ***in this study***.

5. GENERAL METHODOLOGY AND CONVENTIONS

The study will be analyzed and reported once all randomized participants have completed the study (or discontinued), and following study-level data set release.

5.1. Hypotheses and Decision Rules

There are no formal hypothesis tests planned for this study, and a summary of the general analysis methodologies to be utilized is shown in Table 9 below.

Table 9. Summary of Analysis Methods

Treatment	Dose group	Comparator	Primary Analysis Methodology for each comparison
Placebo	Placebo	-	-
DGAT2i	25 mg BID	Placebo	Bayesian E_{max} DR modelling/ER modelling
	75 mg BID	Placebo	Bayesian E_{max} DR modelling/ER modelling
	150 mg BID	Placebo	Bayesian E_{max} DR modelling/ER modelling
	300 mg BID	Placebo	Bayesian E_{max} DR modelling/ER modelling

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Table 9. Summary of Analysis Methods

Treatment	Dose group	Comparator	Primary Analysis Methodology for each comparison
DGAT2i+ACCi	150 mg BID + 5 mg BID	150 mg BID	Pairwise comparison (Logistic Regression or ANCOVA)/Linear DR modelling (if required)
	300 mg BID + 10 mg BID	300 mg BID	Pairwise comparison (Logistic Regression or ANCOVA)/Linear DR modelling (if required)

All 90% credible or confidence intervals and 50% confidence intervals will be assessed as one-sided.

5.2. General Methods

For the assessment of dose-response, the following BID treatment groups (at a minimum) will be included in a Bayesian E_{max} dose-response model:

- Placebo
- DGAT2i 25 mg BID
- DGAT2i 75 mg BID – including those randomized to 150 mg QD
- DGAT2i 150 mg BID – including those randomized to 300 mg QD
- DGAT2i 300 mg BID

Other treatment groups will be analyzed to assess the comparisons as shown below:

- DGAT2i 150 mg BID + ACCi 5 mg BID vs Placebo and vs DGAT2i 150 mg BID
- DGAT2i 300 mg BID + ACCi 10 mg BID vs Placebo and vs DGAT2i 300 mg BID

5.2.1. Analyses for Binary Endpoints

Dose Response Analyses

For binary (responder) endpoints in the study, to be analyzed under a composite estimand strategy *all cases of withdrawal from treatment for lack of efficacy or toleration* or if a Week 48 biopsy is not performed will be considered to be non-responders. Such endpoints *will be analyzed by fitting a Bayesian E_{max} dose-response model. This model will be utilized to characterize the dose response across all BID treatment groups, to estimate the proportion of responders (and 90% credible interval) for each BID dose studied, and to estimate the placebo-adjusted proportion of responders for each dose (and 90% credible interval). If the E_{max} dose-response model cannot be fitted to the data, then other models that allow dose*

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

response to be estimated will be fitted, ie, linear, log-linear or exponential. No adjustment for multiple comparisons will be made.

The 3 parameter Emax model will be fitted to the response at Week 48 on placebo and all DGAT2i monotherapy BID doses – the actual dose to be modelled should be the Total Daily Dose (i.e. BID dose*2). If the model converges, model estimated parameters will be presented along with their 90% credible intervals. The model projected placebo adjusted treatment effect for each BID dose (i.e. Total Daily Dose) will also be presented along with their credible intervals.

The Bayesian estimation of the Emax model uses prior distributions on the placebo response (E0), dose that produces half maximal drug effect (ED_{50}) and Emax parameters. The specification of the Emax model and the prior distributions for some of its parameters are based on three meta-analyses of clinical dose response from more than 100 compounds (Thomas, 2014; Thomas & Roy, 2017; Wu, 2017).
The DGAT2i ED_{50} for the histological-based primary endpoint has been assumed to be slightly higher than the projected EC_{50} for reduction in liver fat which was estimated from a preliminary population pharmacokinetic model based on previous clinical studies with DGAT2i. Therefore, the DGAT2i ED_{50} was estimated to be approximately 30 mg BID (TDD = 60 mg). The substantial uncertainty in ED_{50} values was assessed in the meta-analyses of clinical dose response studies (Thomas, 2014). Therefore, based on the meta-analyses, a scaled t-distribution (with 5 degrees of freedom) is planned to be used that is focused on the initial projected ED_{50}=30 mg BID, with a scale parameter of 0.6. A normal prior distribution for the logit of the placebo response centered at logit(0.16) with a prior standard deviation of 2.0 (logistic scale) is planned to be used, which yields a weak diffuse prior distribution for the placebo response. A normal prior distribution for the Emax parameter will also be used, centred on the anticipated Emax (0.6) on the logistic scale. This prior distribution will also be diffuse on the logistic scale with a prior standard deviation of 2.0.

The proportion of primary endpoint responders in each treatment group will be analyzed by fitting a Bayesian Emax model to the data to characterize the dose response across all treatment groups and to estimate the proportion of responders (and 90% credible interval) for each dose studied, and to estimate the placebo-adjusted proportion of responders for each dose (and 90% credible interval). If an Emax model cannot be fitted to the data, then other models that allow dose response to be estimated will be fitted, i.e. linear, log-linear or exponential.

Note that a similar methodology will be used for the exploratory 'worsening from baseline' liver biopsy endpoints, but for this Estimand 3.1 will be utilized.

Pairwise Treatment Comparisons

The comparison of the proportion of responders for binary responder endpoints for DGAT2i BID doses vs placebo, and DGAT2i+ACCi BID doses vs placebo and corresponding DGAT2i BID monotherapy doses *will be analyzed using logistic regression models,* including baseline (SCR2 Eligibility) fibrosis stage (F2/F3) as a factor, *to estimate the*

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

proportion of responders in each treatment group and odds ratio and risk difference *(and corresponding 90% CIs) for each treatment comparison.* Risk difference and 2-sided 90% confidence interval for risk difference will be calculated by using the observed placebo rate and estimated odds ratio from the logistic regression model. *For comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% CI will also be provided. No adjustment for multiple comparisons will be made.*

5.2.2. Analyses for Continuous Endpoints

Dose Response Analyses

Continuous endpoints will be analyzed using a Bayesian Emax dose-response model, fitted to LS means obtained from the MMRM model fitted to the continuous endpoint to be analyzed. This Bayesian Emax model will be utilized to characterize the dose response across all BID treatment groups, to estimate the posterior mean relative change from baseline (and 90% credible interval) for each BID dose studied, and to estimate the placebo-adjusted posterior mean relative change from baseline for each dose (and 90% credible interval). If the E_{max} dose-response model cannot be fitted to the data, then other models that allow dose response to be estimated will be fitted, ie, linear, log-linear or exponential. No adjustment for multiple comparisons will be made.

A prior of CCI % for the placebo response has been estimated from the CCI study and a prior standard deviation of CCI times the residual SD from the MMRM model will be used, which yields a weak diffuse prior distribution for the placebo response. The mean of the prior distribution of the difference in response between the target dose and placebo is set to zero so that the prior of the mean difference over placebo is centred at no effect. The SD of the prior distribution of the difference in response between the target dose and placebo will be CCI times the residual SD from the MMRM model.

Pairwise Treatment Comparisons

To estimate pairwise treatment comparisons for DGAT2i BID doses vs placebo, and DGAT2i+ACCi BID doses vs placebo and corresponding DGAT2i BID monotherapy doses, continuous endpoints *will be analyzed using an ANCOVA performed on log-transformed* relative *change from baseline with treatment (where relative change from baseline = post baseline value/baseline value) and baseline fibrosis stage (F2/F3) as factors, and log-transformed baseline* value of the endpoint *as a covariate. Estimates of the mean relative changes for each treatment comparison and the corresponding 90% CI will be obtained from the model. The LS mean relative changes* (RC) *and their CIs will be exponentiated, and percent change will then be determined as follows: Percent change = 100* (RC – 1). For comparisons of DGAT2i + ACCi BID vs the corresponding DGAT2i BID monotherapy doses the corresponding 50% CI will also be provided. No adjustment for multiple comparisons will be made.*

Descriptive summaries of the observed values and percent change from baseline for each treatment group will also be produced.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Treatment effects over time

A Mixed-effects model with repeated measures *(MMRM) will be used for the analysis of continuous endpoints with more than one post-baseline collection time point. All observed data collected during the post-baseline treatment period will be utilized*, including follow-up data in order to assess the impact of stopping treatment on the endpoint [redacted]. *The MMRM analysis will be performed with treatment, study week (nominal timepoint) and treatment-by-study week interaction as fixed effects, and baseline value of the analysis endpoint* as a covariate, baseline (SCR2 Eligibility) fibrosis stage (F2/F3) as a factor. *If the endpoint being analyzed is log-transformed, the baseline value of the analysis endpoint will also be log-transformed. Repeated measures model with unstructured* covariance *matrix will be utilized. If this does not converge then compound symmetry structure will be considered. Additionally, the number of covariates may be reduced to improve model fit. Estimates of treatment effects will be assessed using* Least Square Means *(LSMs) and CIs at each time point. LSM difference for each comparison along with the corresponding 90% CI will be provided.* For log-transformed endpoints, the mean relative changes (RC) and their CIs will be exponentiated and percent change will then be determined as follows: Percent change = 100* (RC – 1). *If there are major deviations from the statistical assumptions underlying this model then alternative transformations (eg, log) or non-parametric analyses may be presented.*

5.2.3. Analyses for Categorical Endpoints

Not applicable

5.2.4. Analyses for Time-to-Event Endpoints

The occurrence of adjudicated time-to-event endpoints *over time up to Week 48 will be summarized descriptively by treatment group.*

5.3. Methods to Manage Missing Data

For composite estimands, if a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have any evaluable Week 48 Biopsy data, they will be considered to be a non-responder (i.e. response will be imputed as 'non-response'). If they withdraw from the study for other reasons (not related to lack of efficacy or toleration) and have evaluable biopsy data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.

For hypothetical estimands, for the biopsy responder endpoints, analysis will be carried out under the assumption that missing data at Week 48 are missing at random conditional on dose and the strata formed by the covariates., The unconditional estimate of treatment effect will be obtained through covariate adjustment (ie, averaged over strata formed by baseline fibrosis status (F2/F3) and metabolic status (with or without T2DM)). For the continuous endpoint (i.e. percentage change from baseline in liver fat), missing data at Week 48 will not be imputed. MMRM adjusted for covariates will be used to obtain LS mean estimate at Week 48 for each dose, which will then be used in fitting the Bayesian Emax model.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

For quantitative biopsy analyses, if the SCR2 Change from Baseline biopsy assessment is deemed to be 'Non-evaluable' the SCR2 Eligibility biopsy assessment will be used instead. If the Week 48 Change from Baseline biopsy assessment is deemed to be 'Non-evaluable', the Week 48 IPV biopsy assessment will be used instead (if this is available, as not all participants had a Week 48 IPV assessment).

Table 10. Handling non-evaluable histological change-from-baseline data

Scenario	Handling
CFB-SCR = nonevaluable (for NAS or fibrosis component(s) required for endpoint)	• Utilize SCR2 eligibility results for whichever component(s) is/are nonevaluable (ie, consensus results if available); if only parallel is available, select results from pathologists that had both the lowest NAS score and/or lowest fibrosis stage • Qualitative assessment will only include data from whichever component is evaluable, if any (either NAS components or fibrosis)
CFB-Wk48/DC = nonevaluable (for NAS or fibrosis component(s) required for endpoint)	• If case did not complete IPV → count as non-responder • If case completed IPV, utilize Week 48-IPV results for whichever component(s) is/are nonevaluable as Week 48 result in CFB (ie, IPV consensus results if available); if only parallel is available, select results from pathologist that had highest NAS score and/or fibrosis stage • Qualitative assessment will only include data from whichever component is evaluable, if any (either NAS components or fibrosis)

Observed cases (including any alternative cases utilized for analysis as described above) will be used when analyzing variables using an MMRM.

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

PK Concentrations below the limit of quantification

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

Deviations, missing concentrations, and anomalous values for PK

In summary tables, statistics will be calculated having set concentrations to missing if one of the following cases is true:

1. A concentration has been collected as ND (ie not done) or NS (ie no sample),

2. A deviation in visit window is of sufficient concern or a concentration has been flagged anomalous by the pharmacokineticist.

Note that summary statistics will not be presented at a particular visit if more than 50% of the PK data are missing.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Efficacy and Pharmacodynamic (PD) endpoints

- Missing data for all efficacy and PD endpoints 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) 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".
- Laboratory parameters that are not correctly collected (eg, collected in non-fasted when required to be collected in fasted state) should not be used and set to missing.

Date of Birth

- Only the year of birth is collected in the CRF, therefore the day and the month will be imputed as 01-Jan in order to derive a complete Date of Birth.

6. ANALYSES AND SUMMARIES

6.1. Primary Endpoint(s)

6.1.1. Proportion of participants achieving centrally adjudicated resolution of NASH without worsening of fibrosis or improvement in fibrosis by ≥1 stage without worsening of NASH, or both, at Week 48

6.1.1.1. Main Analysis

- Estimand strategy: Composite (Section 2.1.1).
- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided evaluable baseline data for the primary endpoint (i.e. baseline biopsy data).
- Analysis methodology: For assessing dose response, the proportion of responders at Week 48 will be analyzed using a Bayesian Emax dose-response model. For assessing pairwise comparisons, the proportion of responders at Week 48 will be analyzed using logistic regression models with treatment and baseline fibrosis stage (F2/F3) as factors (Section 5.2.1).
- Intercurrent events and missing data: If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, they will be considered to be a non-responder. If they

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

withdraw for other reasons (not related to lack of efficacy or toleration) and have evaluable biopsy data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.

- The number and proportion of responders for the primary endpoint at Week 48 will be presented for each treatment arm.

- The model-estimated proportion of responders (and 90% credible interval) for each treatment arm, and the placebo-adjusted proportion of responders for BID each dose (and 90% credible interval), and odds ratio and risk difference (and corresponding 90% confidence interval) for other treatment comparisons will be presented at Week 48. For comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence interval will also be provided.

6.1.1.2. Sensitivity/Supplementary Analyses

Sensitivity Analyses

In order to assess the sensitivity of the dose Bayesian Emax dose-response model to the location and informativeness of the prior distributions, sensitivity analyses may be performed that will utilize alternative prior distributions.

Supplementary Analyses

An analysis that assesses the primary endpoint using a hypothetical estimand strategy (Section 2.1.3) will be performed. Using a hypothetical estimand strategy in the study population of randomized and treated participants with biopsy-confirmed NASH and fibrosis, as assessed by the sponsor-identified central pathologist(s), the treatment effect of DGAT2i doses administered alone and coadministration of DGAT2i + ACCi relative to placebo, and coadministration of DGAT2i + ACCi, relative to DGAT2i alone will be estimated, in terms of estimating the proportion of 'clinical responders', defined as participants achieving centrally adjudicated: resolution of NASH without worsening of fibrosis or improvement in fibrosis by ≥1 stage without worsening of NASH, or both, at Week 48, assuming all patients had remained in the trial and received treatment as planned without withdrawal up to 48 weeks. All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models with a logit transformation, or logistic regression models.

Additional analyses may also be performed to assess participants whose SCR2 assessment when as part of the final paired biopsy assessment was different to the original SCR2 assessment for eligibility. Summaries of proportion of participants with No Change, Improvement or Worsening in each component of NAS and Fibrosis from the qualitative comparison of baseline to the end of treatment biopsy results will be provided. Shift tables of Baseline/SCR2 vs Week 48/EOT individual NAS component scores and fibrosis grades will be produced. If a participant doesn't have a pair of CFB results (for SCR2 and Week 48/EOT) (even after assessing if other possible results are able to be utilized as described in Table 10), then their SCR2 Eligibility result will be tabulated vs a 'Missing' Week 48 result.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

To assess concordance of intra-pathologist reviews, a random sample of biopsies will be selected (10% for each of the screening biopsies amongst those ultimately randomized and a separate 10% for the end of treatment biopsies obtained) to be reviewed again by the 2 central pathologists (i.e. IPV review) – with the review process identical to that used for SCR2 eligibility. Kappa statistics (simple and weighted) and 90% CI will be estimated based on the initial and second assessments of each central pathologist. For the screening biopsy comparisons, the SCR2 Eligibility biopsy will be the initial assessment that is compared the second assessment. For the Week 48/EOT biopsy comparisons, the Week 48/end of treatment biopsy from the Blinded side-by-side (Change from Baseline) assessment will be the initial assessment that is compared the second assessment. Concordance of inter-pathologist reviews (Pathologist 1 vs Pathologist 2) will also be assessed for Baseline/SCR2 and Week 48/EOT separately. In addition, summaries of the proportion of biopsies at each visit that had to undergo consensus review and a summary of screening biopsies that remaining eligible when reviewed as part of the CFB assessment will be produced (by treatment group).

6.2. Secondary Endpoint(s)

6.2.1. Proportion of participants achieving centrally adjudicated resolution of NASH without worsening of fibrosis, at Week 48

6.2.1.1. Main Analysis

- Estimand strategy: Composite (Section 2.1.2).
- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided evaluable baseline data for the endpoint (i.e. baseline biopsy data).
- Analysis methodology: For assessing dose response, the proportion of responders at Week 48 will be analyzed using a Bayesian Emax dose-response model. For assessing pairwise comparisons, the proportion of responders at Week 48 will be analyzed an using logistic regression models with treatment and baseline fibrosis stage (F2/F3) as factors (Section 5.2.1).
- Intercurrent events and missing data: If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, they will be considered to be a non-responder. If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable biopsy data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.
- The number and proportion of responders for this secondary endpoint at Week 48 will be presented for each treatment arm.
- The model-estimated proportion of responders (and 90% credible interval) for each treatment arm, and the placebo-adjusted proportion of responders for BID each dose (and 90% credible interval), and odds ratio and risk difference (and corresponding 90% confidence interval) for other treatment comparisons will be presented at Week 48. For

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence interval will also be provided.

6.2.1.2. Sensitivity/Supplementary Analysis

Sensitivity Analyses

In order to assess the sensitivity of the dose Bayesian Emax dose-response model to the location and informativeness of the prior distributions, sensitivity analyses may be performed that will utilize alternative prior distributions.

6.2.2. Proportion of participants achieving centrally adjudicated improvement in fibrosis by ≥1 stage without worsening of NASH, at Week 48

6.2.2.1. Main Analysis

- Estimand strategy: Composite (Section 2.1.2).
- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided evaluable baseline data for the endpoint (i.e. baseline biopsy data).
- Analysis methodology: For assessing dose response, the proportion of responders at Week 48 will be analyzed using a Bayesian Emax dose-response model. For assessing pairwise comparisons, the proportion of responders at Week 48 will be analyzed an using logistic regression models with treatment and baseline fibrosis stage (F2/F3) as factors (Section 5.2.1).
- Intercurrent events and missing data: If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, they will be considered to be a non-responder. If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable biopsy data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.
- The number and proportion of responders for this secondary endpoint at Week 48 will be presented for each treatment arm.
- The model-estimated proportion of responders (and 90% credible interval) for each treatment arm, and the placebo-adjusted proportion of responders for BID each dose (and 90% credible interval), and odds ratio and risk difference (and corresponding 90% confidence interval) for other treatment comparisons will be presented at Week 48. For comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence interval will also be provided.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

6.2.2.2. Sensitivity/Supplementary Analysis

Sensitivity Analyses

In order to assess the sensitivity of the dose Bayesian Emax dose-response model to the location and informativeness of the prior distributions, sensitivity analyses may be performed that will utilize alternative prior distributions.

6.2.3. Proportion of participants achieving centrally adjudicated improvement in fibrosis by ≥2 stages without worsening of NASH, at Week 48

6.2.3.1. Main Analysis

- Estimand strategy: Composite (Section 2.1.2).
- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided evaluable baseline data for the endpoint (i.e. baseline biopsy data).
- Analysis methodology: For assessing dose response, the proportion of responders at Week 48 will be analyzed using a Bayesian Emax dose-response model. For assessing pairwise comparisons, the proportion of responders at Week 48 will be analyzed an using logistic regression models with treatment and baseline fibrosis stage (F2/F3) as factors (Section 5.2.1).
- Intercurrent events and missing data: If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, they will be considered to be a non-responder. If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable biopsy data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.
- The number and proportion of responders for this secondary endpoint at Week 48 will be presented for each treatment arm.
- The model-estimated proportion of responders (and 90% credible interval) for each treatment arm, and the placebo-adjusted proportion of responders for BID each dose (and 90% credible interval), and odds ratio and risk difference (and corresponding 90% confidence interval) for other treatment comparisons will be presented at Week 48. For comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence interval will also be provided.

6.2.3.2. Sensitivity/Supplementary Analysis

Sensitivity Analyses

In order to assess the sensitivity of the dose Bayesian Emax dose-response model to the location and informativeness of the prior distributions, sensitivity analyses may be performed that will utilize alternative prior distributions.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

6.2.4. Proportion of participants achieving centrally adjudicated improvement of ≥2 points in Total NAFLD Activity Score (NAS) without worsening of fibrosis, at Week 48

6.2.4.1. Main Analysis

- Estimand strategy: Composite (Section 2.1.2).
- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided evaluable baseline data for the endpoint (i.e. baseline biopsy data).
- Analysis methodology: For assessing dose response, the proportion of responders at Week 48 will be analyzed using a Bayesian Emax dose-response model. For assessing pairwise comparisons, the proportion of responders at Week 48 will be analyzed an using logistic regression models with treatment and baseline fibrosis stage (F2/F3) as factors (Section 5.2.1).
- Intercurrent events and missing data: If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, they will be considered to be a non-responder. If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable biopsy data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.
- The number and proportion of responders for this secondary endpoint at Week 48 will be presented for each treatment arm.
- The model-estimated proportion of responders (and 90% credible interval) for each treatment arm, and the placebo-adjusted proportion of responders for BID each dose (and 90% credible interval), and odds ratio and risk difference (and corresponding 90% confidence interval) for other treatment comparisons will be presented at Week 48. For comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence interval will also be provided.

6.2.4.2. Sensitivity/Supplementary Analysis

Sensitivity Analyses

In order to assess the sensitivity of the dose Bayesian Emax dose-response model to the location and informativeness of the prior distributions, sensitivity analyses may be performed that will utilize alternative prior distributions.

6.2.5. Percentage change from baseline in liver fat (assessed via MRI-PDFF) at Week 48

6.2.5.1. Main Analysis

- Estimand strategy: Hypothetical (Section 2.1.2).

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

- Analysis set: Evaluable (Section 4). All randomized participants in the imaging sub-study, who take at least 1 dose of investigational product who have provided evaluable baseline data for the endpoint, i.e. baseline MRI-PDFF data.

- Analysis methodology: For assessing dose response, the percentage change from baseline in liver fat at Week 48 will be analyzed using a Bayesian Emax dose-response model, fitted to LS means obtained from the MMRM model. For assessing pairwise comparisons, the percentage change from baseline in liver fat at Week 48 will be analyzed using an ANCOVA performed on log-transformed relative change from baseline in liver fat with treatment and baseline fibrosis stage (F2/F3) as factors, and log-transformed baseline liver fat as a covariate. All treatment effect contrasts will be obtained through fitting the Bayesian dose-response and/or exposure-response models, or an ANCOVA model (Section 5.2.2).

- Intercurrent events and missing data: Any available MRI data for all participants collected before or at the time of treatment withdrawal will be included in the analysis, but endpoint data collected after treatment withdrawal will be censored. Missing data due to lack of efficacy or toleration prior to Week 48 or due to other reasons (not related to lack of efficacy or toleration), will be assumed to be missing at random and will not be imputed. Instead, MMRM will be used to obtain LS mean estimate at Week 48 for each dose, which will then be used in fitting the Bayesian Emax model.

- The sample size, mean, standard deviation, median, minimum, and maximum at the baseline and postbaseline visits for observed liver fat and relative and absolute change from baseline will be presented for each treatment arm.

- The model-estimated mean relative changes for each treatment comparison and the corresponding 90% credible interval will be obtained from the model. The mean relative changes (RC) and their CIs will be exponentiated and percent change will then be determined as follows: Percent change = 100* (RC - 1). For comparisons of DGAT2i + ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence intervals will also be provided.

6.2.5.2. Sensitivity/Supplementary Analysis

Sensitivity Analyses

In order to assess the sensitivity of the dose Bayesian Emax dose-response model to the location and informativeness of the prior distributions, sensitivity analyses may be performed that will utilize alternative prior distributions.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

6.3. Other Endpoint(s)

6.3.1. Proportion of participants achieving improvement in fibrosis by ≥1 stage independent of changes in NAS, based on assessment by sponsor-identified central pathologist(s), at Week 48 compared to baseline

6.3.1.1. Main Analysis

- Estimand strategy: Composite (Section 2.1.2).
- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided evaluable baseline data for the endpoint (i.e. baseline biopsy data).
- Analysis methodology: For assessing dose response, the proportion of responders at Week 48 will be analyzed using a Bayesian Emax dose-response model. For assessing pairwise comparisons, the proportion of responders at Week 48 will be analyzed an using logistic regression models with treatment and baseline fibrosis stage (F2/F3) as factors (Section 5.2.1).
- Intercurrent events and missing data: If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, they will be considered to be a non-responder. If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable biopsy data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.
- The number and proportion of responders for this tertiary endpoint at Week 48 will be presented for each treatment arm.
- The model-estimated proportion of responders (and 90% credible interval) for each treatment arm, and the placebo-adjusted proportion of responders for BID each dose (and 90% credible interval), and odds ratio and risk difference (and corresponding 90% confidence interval) for other treatment comparisons will be presented at Week 48. For comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence interval will also be provided.

6.3.2. Proportion of participants achieving resolution of NASH without worsening of fibrosis and improvement in fibrosis by ≥1 stage without worsening of NASH, based on assessment by sponsor-identified central pathologist(s), at Week 48 compared to baseline

6.3.2.1. Main Analysis

- Estimand strategy: Composite (Section 2.1.2).
- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided evaluable baseline data for the endpoint (i.e. baseline biopsy data).

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

- Analysis methodology: For assessing dose response, the proportion of responders at Week 48 will be analyzed using a Bayesian Emax dose-response model. For assessing pairwise comparisons, the proportion of responders at Week 48 will be analyzed an using logistic regression models with treatment and baseline fibrosis stage (F2/F3) as factors (Section 5.2.1).

- Intercurrent events and missing data: If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, they will be considered to be a non-responder. If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable biopsy data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.

- The number and proportion of responders for this tertiary endpoint at Week 48 will be presented for each treatment arm.

- The model-estimated proportion of responders (and 90% credible interval) for each treatment arm, and the placebo-adjusted proportion of responders for BID each dose (and 90% credible interval), and odds ratio and risk difference (and corresponding 90% confidence interval) for other treatment comparisons will be presented at Week 48. For comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence interval will also be provided.

6.3.3. Proportion of participants with progression of fibrosis by ≥1 stage (defined as an increase of at least 1 stage in the Brunt-Kleiner scale compared to baseline) independent of changes in NAS, at Week 48

6.3.3.1. Main Analysis

- Estimand strategy: Composite (Section 2.1.3).

- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided evaluable baseline data for the endpoint (i.e. baseline biopsy data).

- Analysis methodology: For assessing dose response, the proportion of responders at Week 48 (defined as the proportion of participants with worsening of fibrosis by ≥1 stage on liver biopsy, as assessed by sponsor identified central pathologist(s), at Week 48) will be analyzed using a Bayesian Emax dose-response model. For assessing pairwise comparisons, the proportion of responders at Week 48 will be analyzed using logistic regression models with treatment and baseline fibrosis stage (F2/F3) as factors (Section 5.2.1).

- Intercurrent events and missing data: If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 biopsy data, they will be considered to be a responder (as this assumes that these participants will also have worsened). If a participant has improved or

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

maintained the same level of fibrosis compared to baseline, then that participant will be termed as a non-responder.

- If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.

- The number and proportion of responders for this endpoint at Week 48 will be presented for each treatment arm.The model-estimated proportion of responders (and 90% credible interval) for each treatment arm, and the placebo-adjusted proportion of responders for BID each dose (and 90% credible interval), and odds ratio and risk difference (and corresponding 90% confidence interval) for other treatment comparisons will be presented at Week 48. For comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence interval will also be provided.

6.3.4. Proportion of participants with progression of fibrosis by ≥1 stage and worsening of ≥2 points in Total NAS at Week 48

6.3.4.1. Main Analysis

- Estimand strategy: Composite (Section 2.1.3).

- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided evaluable baseline data for the endpoint (i.e. baseline biopsy data).

- Analysis methodology: For assessing dose response, the proportion of responders at Week 48 (defined as the proportion of participants with worsening of fibrosis by ≥1 stage and worsening of ≥2 points in Total NAS on liver biopsy, as assessed by sponsor identified central pathologist(s), at Week 48) will be analyzed using a Bayesian Emax dose-response model. For assessing pairwise comparisons, the proportion of responders at Week 48 will be analyzed using logistic regression models with treatment and baseline fibrosis stage (F2/F3) as factors (Section 5.2.1).

- Intercurrent events and missing data: If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 biopsy data, they will be considered to be a responder (as this assumes that these participants will also have worsened). If a participant has improved or maintained the same level of fibrosis compared to baseline, then that participant will be termed as a non-responder.

- If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

- The number and proportion of responders for this endpoint at Week 48 will be presented for each treatment arm.

- The model-estimated proportion of responders (and 90% credible interval) for each treatment arm, and the placebo-adjusted proportion of responders for BID each dose (and 90% credible interval), and odds ratio and risk difference (and corresponding 90% confidence interval) for other treatment comparisons will be presented at Week 48. For comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence interval will also be provided.

6.3.5. Proportion of participants with Worsening of ≥2 points in Total NAS (defined as an increase of at least 2 points in Total NAS compared to baseline) independent of changes in fibrosis at Week 48

6.3.5.1. Main Analysis

- Estimand strategy: Composite (Section 2.1.3).

- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided evaluable baseline data for the endpoint (i.e. baseline biopsy data).

- Analysis methodology: For assessing dose response, the proportion of responders at Week 48 (defined as the proportion of participants with worsening of ≥2 points in Total NAS independent of changes in fibrosis, on liver biopsy, as assessed by sponsor identified central pathologist(s), at Week 48) will be analyzed using a Bayesian Emax dose-response model. For assessing pairwise comparisons, the proportion of responders at Week 48 will be analyzed using logistic regression models with treatment and baseline fibrosis stage (F2/F3) as factors (Section 5.2.1).

- Intercurrent events and missing data: If a participant discontinues treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 biopsy data, they will be considered to be a responder (as this assumes that these participants will also have worsened). If a participant's Total NAS has improved or stayed the same compared to baseline, then that participant will be termed as a non-responder. If they withdraw from treatment for other reasons (not related to lack of efficacy or toleration) and have evaluable data at withdrawal or Week 48, these data will be utilized to determine whether they are a responder or not.

- The number and proportion of responders for this endpoint at Week 48 will be presented for each treatment arm.

- The model-estimated proportion of responders (and 90% credible interval) for each treatment arm, and the placebo-adjusted proportion of responders for BID each dose (and 90% credible interval), and odds ratio and risk difference (and corresponding 90% confidence interval) for other treatment comparisons will be presented at Week 48. For

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

comparisons of DGAT2i+ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence interval will also be provided.

6.3.6. Other Continuous Tertiary/Exploratory Endpoints

The following continuous tertiary/exploratory endpoints will be summarized and analyzed as described below. The analyses of these endpoints are assessed for efficacy up to Week 48, but the analyses of blood based biomarkers will also include data up to and including Week 50 to allow an assessment of off-treatment effects:

- Percent change in liver fat (assessed via MRI-PDFF in Imaging substudy population), over time up to Week 48
- Percent change in liver volume (assessed via MRI in Imaging substudy population), over time up to Week 48
- Percent change in CCI (assessed via MRI in Imaging substudy population), over time up to Week 48
- Percent change in CCI (assessed via MRI in Imaging substudy population), over time up to Week 48
- Percent change from baseline in liver fat as assessed via CAP™ measure by FibroScan®, (entire study population), over time up to Week 48
- Percent change from baseline in liver stiffness (assessed via VCTE™ using FibroScan® (entire study population), over time up to Week 48
- Percent change from baseline in Liver Function Tests – ALT, AST, ALP, GGT, Total Bilirubin, Direct Bilirubin, and Total Bile Acids, over time up to Week 48
- Percent change from baseline, overtime up to Week 48 in Mechanism-Related and Metabolic Parameters – PCSK9, Adiponectin, Platelet count, eGFR-CKD-EPI-Cystatin-C, and Body Weight.
- Percent change from baseline, overtime up to Week 48 in CK18-M30 and CK18-M65
- Percent change from baseline, overtime up to Week 48 in ProC3 and ProC6
- Percent change from baseline, overtime up to Week 48 in ELF test (including the individual parameters assayed - Serum hyaluronic acid (HA), Serum amino-terminal propeptide of type III procollagen (PIIINP) and Serum tissue inhibitor of metalloproteinases 1 (TIMP-1))
- Percent change from baseline, overtime up to Week 48 in hs-CRP

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

- Percent change from baseline, overtime up to Week 48 Apolipoprotein A1, B (total), B100, B48, C3, E (Potentially mechanism-related parameters)
- Percent change from baseline, overtime up to Week 48 inFasting lipid panel (total cholesterol, direct LDL-C, HDL-C, triglycerides, direct VLDL), , for all participants.
- Change from baseline, over time up to Week 48, in HbA1C, FPG, FPI, HOMA-IR (for all participants (except HbA1c) and also participants with T2DM only (including HbA1c))
- Change from baseline, over time to Week 48 in abdominal pain and bloating questions from the NASH Symptom Diary
- Change from baseline, over time to Week 48 in fatigue, sleep disturbance, and daytime sleepiness questions from the NASH Symptom Diary
- Change from baseline, over time to Week 48 in fatigue as measured by PROMIS Fatigue questionnaire
- Change from baseline, over time to Week 48 in HRQoL sub-scale scores as measured by NASH-Check
- Change from baseline, over time to Week 48 in PGI-S
- Absolute score over time in PGI-S
- Absolute score over time in PGI-C

6.3.6.1. Main Analysis

- Estimand strategy: Not applicable
- Analysis set: Evaluable (Section 4). All randomized participants who take at least 1 dose of investigational product who have provided baseline data for the endpoint, and in addition have at least 1 evaluable data point post randomization.
- Analysis methodology: A MMRM will be used for the analysis of continuous endpoints with more than one post-baseline collection time point. The MMRM analysis will be performed with treatment, study week (nominal timepoint) and treatment-by-study week interaction as fixed effects, and baseline fibrosis stage (F2/F3) as a factor and baseline value of the analysis endpoint as a covariate. If the endpoint being analyzed is log-transformed (i.e. for all percentage changes from baseline in continuous endpoints which will be derived from the analysis of log-transformed relative changes from baseline), the baseline value of the analysis endpoint will also be log-transformed. Repeated measures model with unstructured covariance matrix will be utilized (Section 5.2.2).
- Intercurrent events and missing data: *All observed data collected during the post-baseline treatment period will be utilized.* Missing data will not be explicitly imputed.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

- The sample size, mean, standard deviation, median, minimum, and maximum at the baseline and postbaseline visits for observed values and relative and absolute change from baseline will be presented for each treatment arm.
    - 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.
- For percentage changes from baseline the model-estimated mean relative changes for each treatment comparison and the corresponding 90% CI will be obtained from the model. The mean relative changes (RC) and their CIs will be exponentiated and percent change will then be determined as follows: Percent change = 100* (RC - 1). For comparisons of DGAT2i + ACCi BID doses vs corresponding DGAT2i BID monotherapy doses corresponding 50% confidence intervals will also be provided.
- CCI [redacted] permit retrospective analysis of CCI [redacted])
    - These results may or may not be generated in the context of the present study.

6.3.7. Event-based Tertiary/Exploratory Endpoints

6.3.7.1. Main Analysis

The following endpoints will be summarized descriptively by treatment group (Section 5.2.4) for randomized participants, separately for pre-randomization events and treatment-emergent events (as SAEs will be collected prior to randomization so could also be sent for adjudication):

- Occurrence of any one of the following over time up to Week 48 based on final adjudication:

Deaths

- All cause mortality (deaths regardless of cause)
- Death due to Myocardial Infarction
- Death due to Cerebrovascular Accident
- Death due to Hepatic failure

For non-fatal events.

- Cardiovascular events
    - Myocardial infarction

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

- Cerebrovascular event (stroke) or transient ischaemic attack
- Hospitalizations due to unstable angina
- Hospitalization for heart failure

- Liver events
  - Hepatic encephalopathy
  - Upper Gastrointestinal bleed due to portal hypertension
  - Liver transplant evaluation
  - New diagnosis of hepatocellular carcinoma
  - Ascites
  - Acute kidney injury due to hepatorenal syndrome
  - Drug induced liver injury
  - Increase of Model of end-stage liver disease (MELD) score from ≤12 to ≥15 indicating listing for liver transplant (See Appendix 2.4.2)

6.3.8. Pharmacokinetic (PK) Endpoints

6.3.8.1. Main Analysis

CCI concentration data will be listed and summarized by treatment group, visit and nominal timepoint for DGAT2i and ACCi separately.

As a general rule, samples collected in those assigned placebo will not undergo analysis for DGAT2i and ACCi. However, when analysis is undertaken, for example: for cases of quality issues at selected site(s), PK results will be listed.

Any samples that have been recorded with a status of 'Label information and paperwork disagree' will be excluded from any PK summaries, as these samples are considered to be unreliable with the sample details at the analyzing laboratory not matching sites' source documents.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

6.4. Subset Analyses

To evaluate the effect of treatment on liver fat and liver volume as assessed by MRI-PDFF, these imaging endpoints will only be analyzed in the subset of participants included in the Imaging substudy population. The analysis of these endpoints is included in Sections 6.2.5 and 6.3.6.

The following subset analysis will also be performed:

- Analysis of glycemic endpoints in the Type 2 diabetes mellitus (T2DM) population

6.5. Baseline and Other Summaries and Analyses

Summaries that account for all participants who offered at least consent for pre-qualification and whether screen failed or randomized will be presented.

6.5.1. Baseline Summaries

Demographic and other baseline characteristics such as participant age, gender, weight, body mass index (BMI), waist circumference, region, stratification factors, ethnicity and race, will be summarized by randomized arm.

For each efficacy endpoint (ie primary, secondary or tertiary endpoint), baseline values will be listed and descriptively summarized by randomized arm.

6.5.2. Study Conduct and Participant Disposition

The number of participants randomized, treated, completing and discontinuing study medication and from the study, will be summarized by randomized arm. For participants who did not complete study medication or the study, the reasons for withdrawal will be presented.

6.5.3. Study Treatment Exposure

The number of days of study treatment exposure will be summarized by randomized treatment group and number of participants with ≤8 weeks, >8 to 24 weeks, >24 to 36 weeks, >36 to 48 weeks (+4 days – to allow for the maximum window for administration of study treatment), and >48 weeks (+5 days) exposure respectively. The formula to compute compliance is defined in Appendix 2.5. Overall compliance (Day 1 to Week 48 or premature termination during treatment phase) will be summarized using descriptive summary statistics and by frequency n (%) in the following two categories by randomized arm:

- <50%
- ≥50 and < 80%, inclusive
- ≥ 80%

6.5.4. Medical History

The number and proportion of participants with each of the up to 12 prespecified medical histories CCI will be summarized by randomized treatment group. In addition, the

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

number of participants with 0, 1, 2, 3, 4, 5, 6, 7, 8 or >8 of these prespecified medical histories will be summarized.

6.5.5. Concomitant Medications and Nondrug Treatments

The number and proportion of participants taking specific concomitant medications CCI [redacted] for medical histories CCI [redacted] on Day 1 and as concomitant medications at the end of treatment/Week 48 will be summarized by randomized treatment group, separately for each specific medical history. In addition, the number and proportion of participants who have increased or decreased their dose of these medications will also be summarized by randomized treatment group, and the number of participants with 0, 1, 2, 3, 4, 5 or >5 of these specific medications CCI [redacted] will be summarized. All other concomitant medications will be summarized by randomized treatment group.

6.6. Safety Summaries and Analyses

All safety analyses will be performed on the safety population. These will be presented in tabular and/or graphical format and summarized descriptively and will follow Pfizer standards as appropriate. Further details are provided below.

6.6.1. Adverse Events

MedDRA will be used to classify all AEs with respect to system organ class and preferred term. Summaries of AEs will include treatment-emergent AEs according to treatment group.

Furthermore, a 3-tier approach will be used to summarize AEs. Tier-1 consists of pre-specified adverse events CCI [redacted] *and will include AEs or collections of AEs related to identified adverse drug reactions in this program related to either DGAT2i or ACCi. Where available, standard MedDRA queries will be used to pool different AE terms that are related to the Tier-1 AEs. The precise AE terms that will contribute to the Tier-1 endpoints will be determined prior to unblinding. For these events, the percentage of participants with incident AE, the risk difference* (using an unconditional exact test to compare each active treatment arm vs placebo)*, its 95% confidence interval, and p-value will be provided. The confidence intervals and p-values are not adjusted for multiplicity and are provided for screening purposes only.*

Tier-2 AEs are those that are not Tier-1, but are common, occurring in at least 4 participants in any treatment arm. The cut-off of at least 4 events was chosen because the 95% confidence interval for the between-group difference in percent incidence will always include zero when treatment groups of equal size each have less than 4 events and so adds little to the interpretation of potentially meaningful differences. For these events, the percentage of participants with incident AE, the risk difference and its 95% confidence interval will be provided. The CIs are for estimation purposes only.

For Tier 1 and Tier 2 AEs, TEAEs, by preferred term, will be sorted in descending point estimate of risk difference. For table/graphic output, the following footnotes will be included to provide proper interpretation of p-values and/or confidence intervals, and to describe how comparison is conducted. E.g. P-values and confidence intervals are not adjusted for multiplicity

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

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 <Active treatment versus Placebo>.

Tier-3 AEs are all other AEs (neither Tier-1 nor Tier-2).

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

6.6.2. Laboratory Data

Laboratory parameters will be summarized descriptively over time, and as described in Section 3.5.2. Individual plots of absolute values in Platelet Count, aPTT, INR vs time for participants with Platelet Count<100,000/mm3 after Day 1 to Week 48 will be produced as well as individual plots of absolute values in TG, FPG, FPI, ApoC3 vs time for participants with Fasting TG $\geq$ 400 mg/dL after Day 1 to Week 48.

6.6.3. Vital Signs

Changes from baseline in systolic blood pressure, diastolic blood pressure and pulse rate will be summarized by treatment and time. The number (%) of participants with maximum increases from baseline will be tabulated by treatment as defined in Section 3.5.3. *Numbers and percentages of participants meeting the categorical criteria will be provided and individual values listed in the study report. No formal inferential statistics will be applied to the vital signs data.*

6.6.4. Electrocardiograms

Changes from baseline for the ECG parameters QT interval, heart rate, QTcF interval, PR interval, and QRS interval will be summarized by treatment and time. For purposes of reporting study-level results, QTcF interval will be derived using Fridericia's heart rate correction formula applied to databased QT interval, and RR interval. The number (%) of participants with maximum increases from baseline will be tabulated by treatment as defined in Section 3.5.3. *Numbers and percentages of participants meeting the categorical criteria will be provided and individual values listed in the study report. No formal inferential statistics will be applied to the ECG data.*

7. INTERIM ANALYSES

7.1. Introduction

Interim analyses (IAs) *will be performed to assess safety after approximately 33%, 67%, and 100% of the planned total sample size has been randomized in the study* (and these would primarily be for External Data Monitoring Committee (E-DMC) safety reviews while the study is ongoing). If required, *interim analysis results* from a single IA *may be used for*

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

other purposes such as: *internal business decisions regarding future study planning, conducting a sample size re-estimation, adapting the safety-related endpoints in the study after the interim analysis, or early unblinding to facilitate Population PK/PD model development. Before any interim analysis is instigated, the details of the objectives, decision criteria, dissemination plan, and method of maintaining the study blind as per Pfizer's SOPs will be documented and approved in an E-DMC charter* or a separate interim analysis plan as needed.

7.2. Interim Analyses and Summaries

Any analyses and summaries to be included in the E-DMC reviews will be documented in the E-DMC Charter or a separate interim analysis plan.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

8. REFERENCES

Thomas N, Sweeney K, Somayaji V. Meta-analysis of clinical dose response in a large drug development portfolio. Statistics in Biopharmaceutical Research 2014;6(4):302-17.

Thomas N, Roy D. Analysis of clinical dose-response in small-molecule drug development: 2009-2014. Statistics in Biopharmaceutical Research 2017;9(2):137-46.

Wu J, Banerjee A, Jin B et al. Clinical dose–response for a broad set of biological products: A model-based meta-analysis. Stat Methods in Med Res 2018;27(9):2694-2721

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

9. APPENDICES

Appendix 1. Summary of Efficacy Analyses

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Primary: Proportion of participants achieving centrally adjudicated resolution of NASH without worsening of fibrosis or improvement in fibrosis by ≥1 stage without worsening of NASH, or both, at Week 48	Summary	Evaluable	Observed data.	N/A
	Main analysis (Estimand 1.1)	Evaluable	Participants who discontinue treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, will be considered to be non-responders. All data collected at or after withdrawal from treatment for other reasons (not related to lack of efficacy or toleration) will be included.	Bayesian Emax for DR, Logistic Regression for pairwise comparisons
	Sensitivity Analyses	Evaluable	Same as main analysis, but may utilize alternative prior distributions for Bayesian Emax modelling.	Bayesian Emax for DR
	Supplementary analysis (Estimand 1.2)	Evaluable	All data collected before or at the time of treatment withdrawal will be included. Endpoint data collected after treatment withdrawal will be censored. Analysis will be	Bayesian Emax for DR, Logistic Regression for pairwise comparisons

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
			carried out under the assumption that missing biopsies at Week 48 are missing at random conditional on dose and the strata formed by the covariates (fibrosis status and T2DM status).	
Secondary: Proportion of participants achieving centrally adjudicated resolution of NASH without worsening of fibrosis at Week 48	Summary	Evaluable	Observed data.	N/A
	Main analysis (Estimand 2.1)	Evaluable	Participants who discontinue treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, will be considered to be non-responders. All data collected at or after withdrawal from treatment for other reasons (not related to lack of efficacy or toleration) will be included.	Bayesian Emax for DR, Logistic Regression for pairwise comparisons
	Sensitivity Analyses	Evaluable	Same as main analysis, but may utilize alternative prior distributions for Bayesian Emax modelling.	Bayesian Emax for DR

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Secondary: Proportion of participants achieving centrally adjudicated improvement in fibrosis by ≥1 stage without worsening of NASH at Week 48	Summary	Evaluable	Observed data.	N/A
	Main analysis (Estimand 2.2)	Evaluable	Participants who discontinue treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, will be considered to be non-responders. All data collected at or after withdrawal from treatment for other reasons (not related to lack of efficacy or toleration) will be included.	Bayesian Emax for DR, Logistic Regression for pairwise comparisons
	Sensitivity Analyses	Evaluable	Same as main analysis, but may utilize alternative prior distributions for Bayesian Emax modelling.	Bayesian Emax for DR
Secondary: Proportion of participants achieving centrally adjudicated improvement in fibrosis by ≥2 stages without worsening of NASH at Week 48	Summary	Evaluable	Observed data.	N/A
	Main analysis (Estimand 2.3)	Evaluable	Participants who discontinue treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, will be considered to be non-	Bayesian Emax for DR, Logistic Regression for pairwise comparisons

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
			responders. All data collected at or after withdrawal from treatment for other reasons (not related to lack of efficacy or toleration) will be included.	
	Sensitivity Analyses	Evaluable	Same as main analysis, but utilize alternative prior distributions for Bayesian Emax modelling.	Bayesian Emax for DR
Secondary: Proportion of participants achieving centrally adjudicated improvement of ≥2 points in Total NAFLD Activity Score (NAS) without worsening of fibrosis at Week 48	Summary	Evaluable	Observed data.	N/A
	Main analysis (Estimand 2.4)	Evaluable	Participants who discontinue treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, will be considered to be non-responders. All data collected at or after withdrawal from treatment for other reasons (not related to lack of efficacy or toleration) will be included.	Bayesian Emax for DR, Logistic Regression for pairwise comparisons

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
	Sensitivity Analyses	Evaluable	Same as main analysis, but utilize alternative prior distributions for Bayesian Emax modelling.	Bayesian Emax for DR
Secondary: Percentage change in liver fat (assessed via MRI-PDFF) at Week 48	Summary	Evaluable	Observed data.	N/A
	Main analysis (Estimand 2.5)	Evaluable	All data collected before or at the time of withdrawal will be included. Endpoint data collected after treatment withdrawal will be censored. Missing data at Week 48 will not be imputed. MMRM will be used to obtain LS mean estimate at Week 48 for each dose, which will then be used in fitting the Bayesian Emax model.	Bayesian Emax for DR, ANCOVA for pairwise comparisons
	Sensitivity Analyses	Evaluable	Same as main analysis, but utilize alternative prior distributions for Bayesian Emax modelling.	Bayesian Emax for DR
	Summary	Evaluable	Observed data.	N/A

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Exploratory: Proportion of participants achieving improvement in fibrosis by ≥1 stage independent of changes in NAS, at Week 48	Main analysis (Estmand 3.2)	Evaluable	Participants who discontinue treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, will be considered to be non-responders. All data collected at or after withdrawal from treatment for other reasons (not related to lack of efficacy or toleration) will be included.	Bayesian Emax for DR, Logistic Regression for pairwise comparisons
Exploratory: Proportion of participants achieving resolution of NASH without worsening of fibrosis and improvement in fibrosis by ≥1 stage without worsening of NASH, at Week 48	Summary	Evaluable	Observed data.	N/A
	Main analysis (Estmand 3.3)	Evaluable	Participants who discontinue treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, will be considered to be non-responders. All data collected at or after withdrawal from treatment for other reasons (not related to lack of efficacy or toleration) will be included.	Bayesian Emax for DR, Logistic Regression for pairwise comparisons

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Exploratory: Proportion of participants with Worsening of ≥2 points in Total NAS (defined as an increase of at least 2 points in Total NAS compared to baseline) independent of changes in fibrosis at Week 48	Summary	Evaluable	Observed data.	N/A
	Main analysis (Estimand 3.1)	Evaluable	Participants who discontinue treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, will be considered to be responders (i.e. they will be assumed to have worsened). All data collected at or after withdrawal from treatment for other reasons (not related to lack of efficacy or toleration) will be included.	Bayesian Emax for DR, Logistic Regression for pairwise comparisons
Exploratory: Proportion of participants with progression of fibrosis by ≥1 stage (defined as an increase of at least 1 stage in the Brunt-Kleiner scale compared to baseline) independent of changes in NAS at Week 48	Summary	Evaluable	Observed data.	N/A
	Main analysis (Estmand 3.1)	Evaluable	Participants who discontinue treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, will be considered to be responders (i.e. they will be assumed to have worsened). All data collected at or after	Bayesian Emax for DR, Logistic Regression for pairwise comparisons

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
			withdrawal from treatment for other reasons (not related to lack of efficacy or toleration) will be included.	
Exploratory: Proportion of participants with progression of fibrosis by ≥1 stage and worsening of ≥2 points in Total NAS at Week 48	Summary	Evaluable	Observed data.	N/A
	Main analysis (Estimand 3.1)	Evaluable	Participants who discontinue treatment or from the study due to lack of efficacy or toleration prior to Week 48 or if they do not have evaluable Week 48 Biopsy data, will be considered to be responders (i.e. they will be assumed to have worsened). All data collected at or after withdrawal from treatment for other reasons (not related to lack of efficacy or toleration) will be included.	Bayesian Emax for DR, Logistic Regression for pairwise comparisons

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Appendix 2. Data Derivation Details

Appendix 2.1. Definition and Use of Visit Windows in Reporting

Each collected assessment will be mapped to its actual visit window according the following algorithm described in the table below:

Visit label	Protocol defined Window (days)	Visit	Visit window as part of final reporting*
			For efficacy, PD, PK and safety endpoints
Pre-Qualification	N/A	1	N/A
Screen 1	N/A	2	N/A
Screen 2	N/A	3	N/A
Run-In	N/A	3	N/A
Week -2	-14 ± 4	4	• Between Week 0 and Week -2, data are used to define baseline • Baseline will be result closest ***prior to*** dosing on randomized regimen on Day 1 [Visit 5] • Window for baseline measurements is from Day -22 to Day 1 [Visit 5] pre-dose (The only exception is for Height, for which any measurement from PreQ to Day 1 pre-dose can be used, and for Fibroscan® for which the last evaluable measurement from SCR1 to Day 1 pre-dose can be used) • Label as 'BASELINE'
Week 0	1	5	
Week 2	14 ± 4	6	'Date of assessment' – 'Baseline date' + 1 = [8, 21]; ie, -6 days to +7 days
Week 4	28 ± 4	7	'Date of assessment' – 'Baseline date' + 1 = [22, 35]; ie, -6 days to +7 days
Week 6	42 ± 4	8	'Date of assessment' – 'Baseline date' + 1 = [36, 49]; ie, -6 days to +7 days
Week 8	56 ± 4	9	'Date of assessment' – 'Baseline date' + 1 = [50, 70]; ie, -6 days to +14 days
Week 12	84 ± 4	10	'Date of assessment' – 'Baseline date' + 1 = [71, 98]; ie, -13 days to +14 days
Week 16	112 ± 4	11	'Date of assessment' – 'Baseline date' + 1 = [99, 140]; ie, -13 days to +28 days
Week 24	168 ± 4	13	'Date of assessment' – 'Baseline date' + 1 = [141, 196]; ie, -27 days to +28 days
Week 32	224 ± 4	15	'Date of assessment' – 'Baseline date' + 1 = [197, 252]; ie, -27 days to +28 days
Week 36**	252 ± 4	16	'Date of assessment' – 'Baseline date' + 1 = [197, 308]; ie, -55 days to +56 days
Week 40	280 ± 4	17	'Date of assessment' – 'Baseline date' + 1 = [253, 308]; ie, -27 days to +28 days
Week 48	336 ± 4	19	'Date of assessment' – 'Baseline date' + 1 = [309, 343]; ie, -27 days to +7 days
Week 50 (1st Follow-up visit)		20	'Date of assessment' – 'Date of last dose' + 1 = [7, 27]; ie, any time after 7 days and up to 28 days after last dose.
Week 52 (2nd Follow-up visit)	≥28- and ≤35- days post last dose at		'Date of assessment' – 'Date of last dose + 1 ≥ 28; ie, any time after 28 days after last dose

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Visit label	Protocol defined Window (days)	Visit	Visit window as part of final reporting*
			For efficacy, PD, PK and safety endpoints
	Week 48 visit in the protocol		

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

** This visit window will only be used for Patient Reported Outcomes reported at Week 36. The other post-baseline visits at which these data are collected are Weeks 12, 24 and 48, therefore the Week 32 and Week 40 analysis windows have not been adjusted to allow for the Week 36 analysis window.

Special considerations:

- For participants withdrawn early, the 1st follow-up visit window (F/U) will include data up to 4 days after the last dose of study medication, and the visit label is defined in accordance with the nominal label as recorded on the "E_TERM/FOLLOW_UP" Administration CRF.
- Data from planned and unplanned visits will be windowed, as well as data on study drug and for participants who may have discontinued study medication but are continuing in the study.
- If 2 or more observations fall within the visit window for Week N (regardless of whether they are planned or unplanned), 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.
- 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)'
- For Liver Biopsy, MRI and Fibroscan® assessments, any Week 48 assessments would be considered to be evaluable, regardless of when they were performed relative to the last dose.
- In addition, the Fibroscan® endpoint of CAP™ and VCTE™ will require conduct of FibroScan® following a fast of at least 4 hours and based on ≥ 10 valid measurements, to be evaluable for analysis, and an additional criterion should be applied for liver stiffness (VCTE™) only:
    - For median stiffness (kPa) ≥7.1 kPa:
        - Interquartile stiffness Range (kPa) divided by median stiffness (kPa) x 100 should be ≤30%
        - This needs to be derived from the Interquartile stiffness Range (kPa) and Median stiffness (kPa) as recorded in the database
    - For median stiffness (kPa) <7.1 kPa
        - any Interquartile stiffness Range (kPa) result is acceptable, so this additional criterion does not need to be applied to these values

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

- Assesment of evaluable CAP™ (dB/m) result will not use Interquartile attentuation Range (dB/m).
- For Fibroscan® Baseline derivation, the last evaluable measurement from SCR1 to Day 1 pre-dose can be used (i.e. closest results obtained prior to dosing on Day 1 (Visit 5)).

- Patient Reported Outcomes: All PRO endpoints will be analyzed according to the analysis windows above (and the Week 36 analysis window will only be used for the PROs completed at Week 36). Note that only the daily assessments recorded in the e-Diary prior to dosing on Day 1 [Visit 5] will be included in the derivation of Baseline NASH Symptom Diary scores. Any daily assessments captured in the e-Diary on Day 1 or later will be excluded from the Baseline scores. For PGI-S, as this is recorded twice at each visit, the records will be selected based on the below method:

  Closest to Target (first) – Selects the record closest to the target. If two records are equi-distant from the target (before and after), select the 1st record.

- For the following assessments, the above visit windows will not be used:
  - Liver Biopsy:

    For liver biopsy assessments, any pre-randomization biopsy data will be considered to be 'Baseline', and any post-randomization biopsy data collected on or after the Week 24 visit will be considered to be 'Week 48/End of Study'. This is so that any liver biopsies performed at the 'early termination' visit, which will be an unplanned visit (after Week 24), can be windowed appropriately so that data will be summarized and included in the analysis as required.
  - MRI: All MRI endpoints will be analyzed according to the nominal visit to which they relate.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Appendix 2.2. Endpoint Derivations

Appendix 2.2.1. % Liver Fat using MRI-PDFF and [redacted]

For each participant in the Imaging substudy 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:

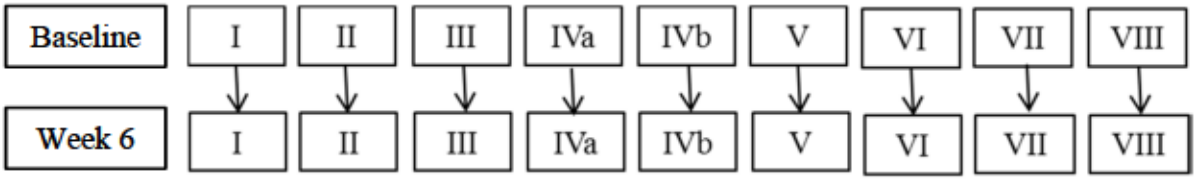
$WLPDFF = (SPDFF_{Segment\ II} + SPDFF_{Segment\ III} + SPDFF_{Segment\ IVa} + SPDFF_{Segment\ IVb} + SPDFF_{Segment\ V} + SPDFF_{Segment\ VI} + SPDFF_{Segment\ VII} + SPDFF_{Segment\ VIII} + (0.36 * SPDFF_{Segment\ I})) / 8.36$ (if all segments are assessed and not missing/mapping at Baseline, and on Week 6, Week 24, and Week 48).

Note: If data for all segments is not included, then adjust the denominator accordingly. This should be the number of all segments contributing to the numerator, except if this includes Segment 1, then add 0.36 instead of 1). For example, if the only segment not included is Segment 1 then the denominator=8, if segment 2 is not included then the denominator=7.36 ((7 x 1)+0.36) etc.

Calculation Rule for % Liver Fat when Assessed with MRI-PDFF:

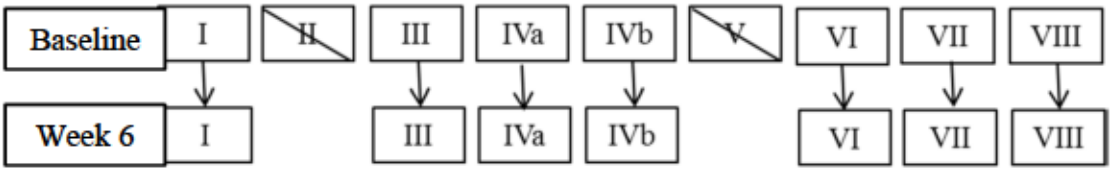
WLPDFF will be calculated on mapping non-missing segments at Baseline, and at Week 6, Week 24, and Week 48.

For instance, if results are reported for all segments at Baseline, and at Week 6, Week 24, and Week 48 (see below), the number of segments assessed is equal to 9.



If at Baseline, all segment data are available but, at Week 6, or Week 24, or Week 48, only 7 segments have results reported, WLPDFF will be calculated at Baseline, and at Week 6, Week 24, and Week 48 and using the matching individual segments providing values on all weeks. For instance, based on the schema below, segment II and Segment V will not be used to calculate WLPDFF on any of the time points reported. In this case, WLPDFF will be calculated using data in the 7 segments reported at all non-missing visits. A minimum of 5 segments need to be available at Baseline and at least one post-baseline visit in order to calculate WLPDFF and the participant should have a fasted status of 'Yes' (i.e. fasted for a duration of at least 4 hours). If a visit has fewer than 5 segments or the participant was not fasted, then the whole visit would be considered 'missing'.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)



WLPDFF and individual % liver fat measured in each individual segment will be included in source data. WLPDFF will be summarized by randomized arm and Week.
Note: If baseline at the individual segment level is zero by itself or after imputation, % change from baseline and relative change 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.

A similar derivation will be utilized to determine CCI (values of which will be provided for each segment), which reflects the presence of CCI and thus is reflective of CCI (combination of liver CCI and CCI).

Appendix 2.2.2. MRI-derived Liver Volume

No additional derivation is required for MRI-derived Liver Volume – a single value (per participant per visit) for liver volume (in mL) will be recorded on the database.

Appendix 2.2.3. MRI-derived CCI

No additional derivation is required for MRI-derived CCI – a single value (per participant per visit) for CCI (in mL) will be recorded on the database.

Appendix 2.2.4. Fasting Plasma Glucose and Fasting Plasma Insulin (T2DM Participants Only)

FPG and FPI values will be assessed at timepoints described in the Schedule of Activity 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)

Appendix 2.2.5. Homeostatic Model Assessment for Insulin Resistance (T2DM Participants Only)

The 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 follows:
$HOMA\text{-}IR = (Glucose_{Fasting\ Concentration} \times Insulin_{Fasting\ Concentration}) / 405$

where plasma glucose concentrations is reported in mg/dL and plasma insulin is reported in mU/L. As such, HOMA-IR is unit-less.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Appendix 2.3. Data Derivation for Patient Reported Outcomes

Appendix 2.3.1. NASH-CHECK

The NASH-Check is a NASH-specific health-related quality of life measure that was developed based on qualitative participant input. It consists of 31 items that measure 3 domains: symptoms (11-point NRS), day-to-day activities (5-point VRS), and emotions and lifestyle (4-point VRS), over the past 7 days.

NASH-CHECK Symptom Scale Scores

Single Symptom Item Scales

Five of the symptom scale scores are single items (**abdominal pain**, item 1; **abdominal bloating**, item 2; **fatigue**, item 3; **sleep**, item 4; **itchy skin**, item 9). Each item has an 11-point numerical rating scale with response anchors ranging from no symptom (scored 0) to worst possible or extreme symptom (scored 10), with a higher score indicating more severe symptoms. The item score is used as the corresponding symptom scale score and the item must be completed for that symptom scale score to be derived.

Symptom scale scores range 0 to 10, with higher scores indicating more severe symptoms.

Cognitive Symptom Scale

NASH-CHECK has 4 items assessing cognitive symptoms (items 5–8). Each item has an 11-point numerical rating scale with response anchors ranging from no symptom (scored 0) to worst possible or extreme symptom (scored 10), with a higher score indicating more severe symptoms.

The cognitive symptoms raw score is calculated as the mean of the non-missing item scores for items 5 to 8 (i.e., the sum of items 5–8 divided by the number of non-missing items). A minimum of 3 cognitive symptoms items need to be answered for a raw score to be derived; specifically, if a respondent misses 2 or more of these items, that person's cognitive symptoms score is set to missing. The raw score is used as the final scale score. The formulas for the raw and scale scores are as follows:

Cognitive symptoms raw score = sum of non-missing scores for items 5 to 8 / number of non-missing items

Cognitive symptoms scale score = cognitive symptoms raw score

Cognitive symptoms scale scores range 0 to 10, with higher scores indicating more severe symptoms.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

NASH-CHECK HRQOL Scale Scores

Activity Limitations Scale

NASH-CHECK has 8 items assessing activity limitations due to NASH (items 10–17). Each item has a 5-point verbal rating scale ranging from "no difficulty" (scored 0) to "unable to do" (scored 4), with a higher score indicating greater problems.

The activity limitations raw score is calculated as the mean of the non-missing item scores for items 10 to 17 (i.e., the sum of items 10–17 divided by the number of non-missing items). A minimum of 5 activity limitations items need to be answered for a raw score to be derived; specifically, if a respondent misses 4 or more of these items, that person's activity limitations score is set to missing. Finally, the raw score is transformed to range 0 to 10 to form the final scale score. The formulas for the raw and scale scores are as follows:

Activity limitations raw score = sum of non-missing scores for items 10 to 17 / number of non-missing items

Activity limitations scale score = (activity limitations raw score / 4) * 10

Activity limitations scale scores range 0 to 10, with higher scores indicating greater activity limitations.

Emotional Impact Scale

NASH-CHECK has 4 items assessing emotional impact due to NASH (items 18–21). Each item has a 4-point verbal rating scale ranging from "not at all" (scored 0) to "very much" (scored 3), with a higher score indicating greater problems.

The emotional impact raw score is calculated as the mean of the non-missing item scores for items 18 to 21 (i.e., the sum of items 18–21 divided by the number of non-missing items). All 4 emotional impact items need to be answered for a raw score to be derived; specifically, if a respondent misses 1 or more of these items, that person's emotional impact score is set to missing. Finally, the raw score is transformed to range 0 to 10 to form the final scale score. The formulas for the raw and scale scores are as follows:

Emotional impact raw score = sum of non-missing scores for items 18 to 21 / number of non-missing items

Emotional impact scale score = (emotional impact raw score / 3) * 10

Emotional impact scale scores range 0 to 10, with higher scores indicating greater emotional impact.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Social Impact Scale

NASH-CHECK has 7 items assessing social impact due to NASH (items 22–28). Each item has a 4-point verbal rating scale ranging from "not at all" (scored 0) to "very much" (scored 3), with a higher score indicating greater problems.

The social impact raw score is calculated as the mean of the non-missing item scores for items 22 to 28 (i.e., the sum of items 22–28 divided by the number of non-missing items). A minimum of 5 social impact items need to be answered for a raw score to be derived; specifically, if a respondent misses 3 or more of these items, that person's social impact score is set to missing. Finally, the raw score is transformed to range 0 to 10 to form the final scale score. The formulas for the raw and scale scores are as follows:

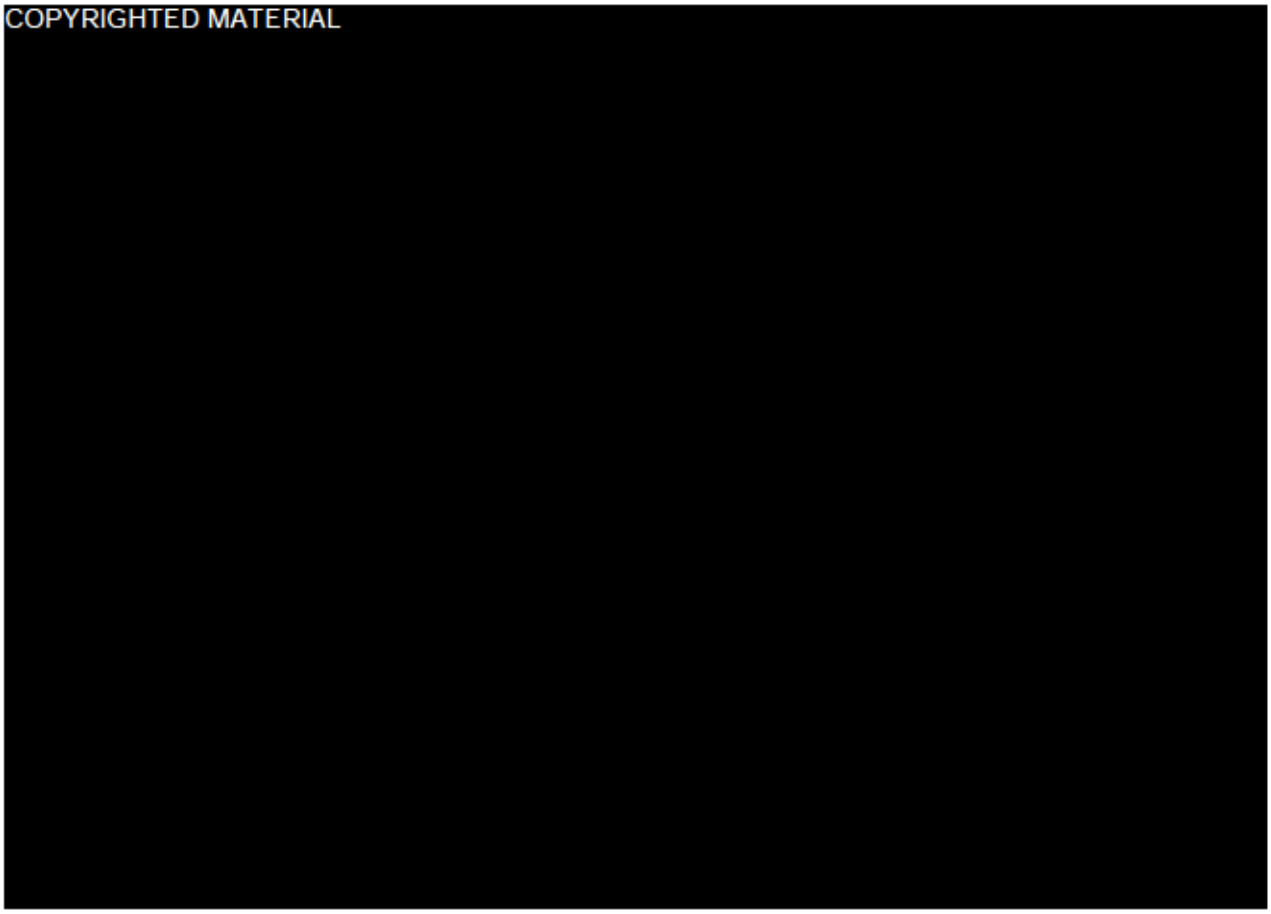
Social impact raw score = sum of non-missing scores for items 22 to 28 / number of non-missing items

Social impact scale score = (social impact raw score / 3) * 10

Social impact scale scores range 0 to 10, with higher scores indicating greater social impact.

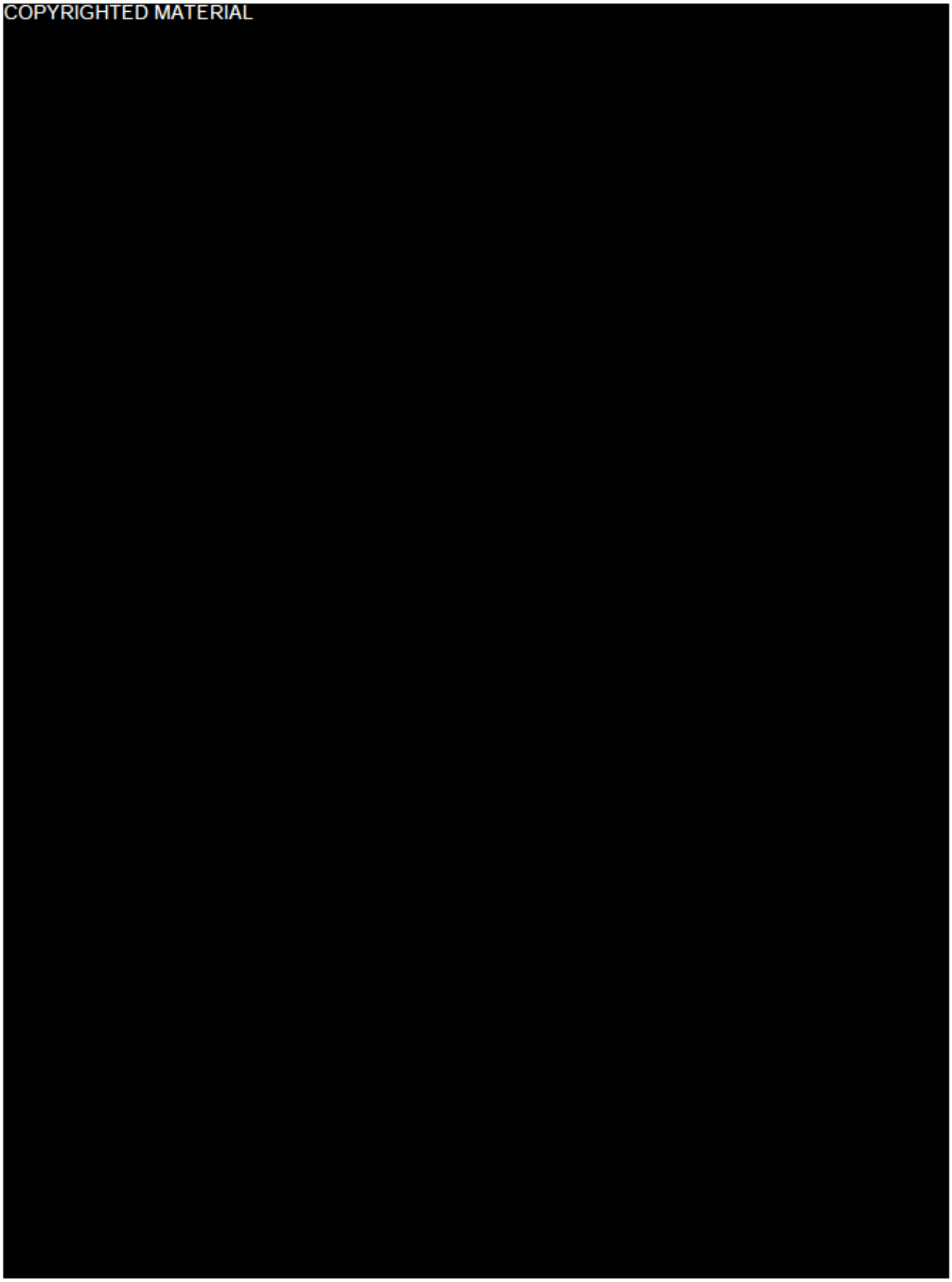
- **NASH-Check 31-Item Version and 28-Item Version item mapping:**

COPYRIGHTED MATERIAL



090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

COPYRIGHTED MATERIAL



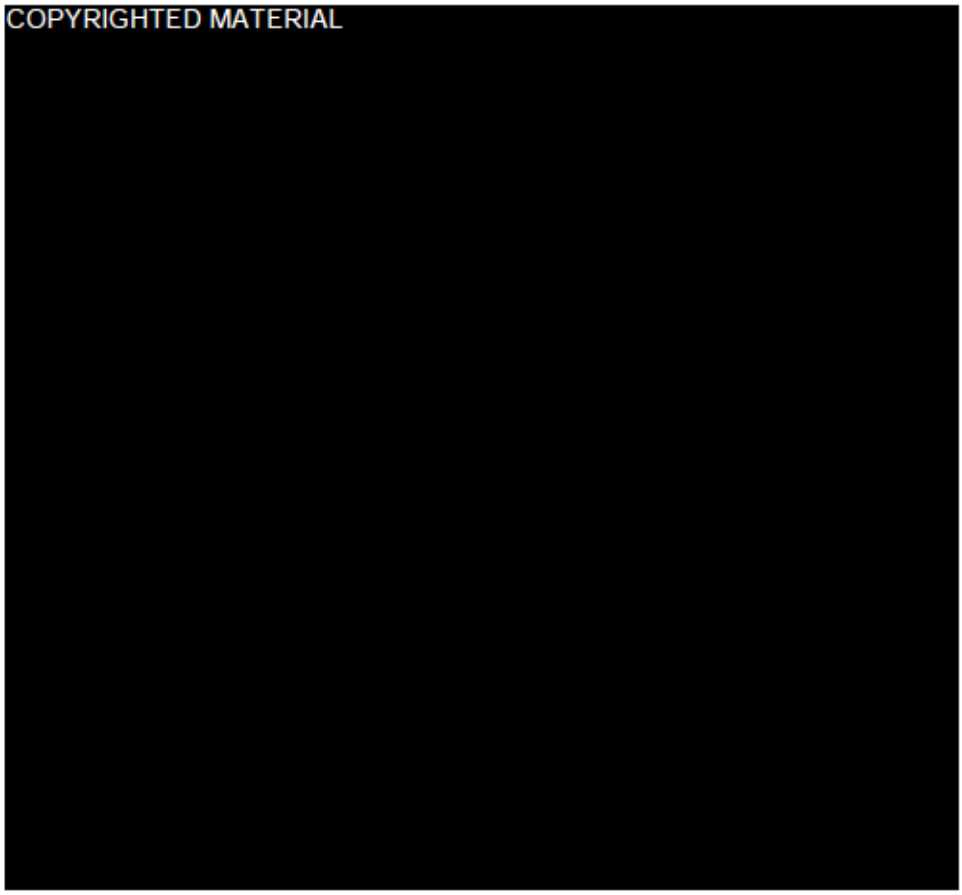
090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Appendix 2.3.2. PROMIS Fatigue Custom 9-Item Version

The PROMIS Fatigue is a self-reported measure that assesses a range of symptoms in the past 7 days from mild subjective feelings of tiredness to an overwhelming, debilitating, and sustained sense of exhaustion that likely decreases one's ability to execute daily activities and function normally in family or social roles.

The short form consists of 9 items that participants will be asked to rate from 1: "Never" to 5: "Always". A global raw score ranging from 9 to 45 is calculated and can be translated into a T-score (Mean= 50, SD = 10) using either the applicable score conversion table provided by the developer (see below) or using online scoring service (http://www.healthmeasures.net/score-and-interpret/calculate-scores/scoring-instructions).

Both the global raw score and the T-score will be summarized, and only if all items have been answered (no missing responses are allowed).



Appendix 2.3.3. NASH Symptom Diary

The NASH Symptom Diary is a daily, self-administered questionnaire that measures symptoms of NASH. This measure was developed by Pfizer based on qualitative participant input as well as review of other data sources (ie, literature and other existing measures). The measure consists of a total of 5 items that ask participants to rate the severity of their

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

abdominal symptoms (pain and bloating [2 items]), fatigue (1 item), sleep disturbance (1 item), and daytime sleepiness (1 item), in the past 24-hours on an 11-point NRS that ranges from 0 to 10. The NASH daily symptom diary entries will be identified from each date/time-stamped (Refer to Date/Time of Collection in Section 3.11 of the DTS).

NASH Symptom Diary Scoring

NASH Symptom Diary is a newly developed measure. The current scoring algorithm is preliminary. The final scoring algorithm will be developed based on psychometric evaluation of the measure once C2541013 data is available.

- Each item is scored individually
- The score for each item is defined as the average of daily data over the assessment period (14 consecutive days)
    - Assessment period is defined in the study protocol (Section 8.1.3, Table 4)
    - Min of 50% of daily data (min of 7 days, and these do not need to be consecutive but must be within the same 14 day period) during the assessment period is required to calculate the score
    - If <7 days of scores during assessment period, then average daily score should be set to missing
    - Analyse Change from Baseline in (1) Pain. (2) Bloating. (3) Fatigue. (4) Sleep disturbance, and (5) Daytime sleepiness average daily score.

Appendix 2.4. Data Derivation for Non Standard Safety Endpoints

Appendix 2.4.1. ELF Test

ELF score (highlighted below) will be calculated using the LabCorp reported sensitive results for following 3 parameters assayed using the Centaur (ROW) or Centaur XP (China) chemiluminescence assay from Siemens –

1. Serum hyaluronic acid (HA) – reported in *ng/mL*
    - assay range = 1.6 to 1000 ng/mL
2. Serum amino-terminal propeptide of type III procollagen (PIIINP) – reported in *ng/mL*
    - assay range = 0.5 to 150 ng/mL
3. Serum tissue inhibitor of metalloproteinases 1 (TIMP-1) – reported in *ng/mL*
    - assay range = 3.5 to 1000 ng/mL

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

ADVIA Centaur and Centaur XP:

ELF score = $2.278 + 0.851 \ln(C_{HA}) + 0.751 \ln(C_{P3NP}) + 0.394 \ln(C_{TIMP1})$

ADVIA Centaur CP:

ELF score = $2.494 + 0.846 \ln(C_{HA}) + 0.735 \ln(C_{P3NP}) + 0.391 \ln(C_{TIMP1})$

Concentrations (C) of each assay are in ng/mL

The above highlighted formula should be added as a programming footer for the respective outputs, as follows:
"ELF Score using Centaur and Centaur XP assay derived as $2.278 + 0.851 \ln(C_{HA}) + 0.751 \ln(C_{P3NP}) + 0.394 \ln(C_{TIMP1})$, where concentrations (C) of each assay are in ng/ml."

Appendix 2.4.2. MELD score

The MELD score for all the participants to be derived using the below mentioned formula:-

Initial $MELD_{(i)}$ score = $0.957 \times Log_e$(creatinine mg/dL) + $0.378 \times Log_e$(bilirubin mg/dL) + $1.120 \times Log_e$(INR) + 0.643

Note: Laboratory values < 1.0 will be set to 1.0 when calculating the initial $MELD_{(i)}$ score. The maximum MELD score is 40. The MELD score derived from this calculation will be rounded to the tenth decimal place and then multiplied by 10.

If the initial $MELD_{(i)}$ score is greater than 11, then the score should be re-calculated using the below formula:-

$MELD = MELD_{(i)} + 1.32*(137-Na) - [0.033*MELD_{(i)}*(137-Na)]$

Notes:-

1. If the creatinine value is >4.0 mg/dL then set this to 4.0.

2. If Sodium (Na in the second equation) is <125 mmol/L then set this to 125.

3. If Sodium (Na in the second equation) is >137 mmol/L then set this to 137.

4. All units should be in Conventional Units (conversions shown in table below)

To be included in Adjudicated Events listing, (MELD) score must increase from ≤12 at baseline to ≥15 on at least 2 consecutive occasions at least 4 weeks apart and without an intermediate score <15.

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Conversion of SI units to conventional units:

Laboratory parameter (unit in MELD-Na)	SI result → conventional unit result
Serum creatinine (mg/dL)	µmol/L x **0.0113** → result in mg/dL
Serum sodium (mEq/L)	mmol/L result x **1** = result in mEq/L
Total bilirubin (mg/dL)	µmol/L x **0.0585** → result in mg/dL
INR (Unitless parameter)	Result in SI unit x **1** = result in conventional unit
Source for conversions: https://www.labcorp.com/resource/si-unit-conversion-table	

Appendix 2.4.3. Estimated Glomerular Filtration Rate

The Estimated Glomerular Filtration Rate (eGFR) will be calculated using cystatin-C, as follows:

If cystatin-C>0.8 mg/L:

eGFR (mL/min/1.73 m^2) = 133 x (cystatin-C/0.8)$^{-1.328}$ x 0.996^{Age}

(multiply whole equation by 0.932 if female)

If cystatin-C≤0.8 mg/L:

eGFR (mL/min/1.73 m^2) = 133 x (cystatin-C/0.8)$^{-0.499}$ x 0.996^{Age}

(multiply whole equation by 0.932 if female)

Where Age is the participant's age at screening (collected at SCR1 visit) in years.

For the entire study population, eGFR will be summarized by randomized arm.

Ref: Inker LA, Schmid CH, Tighiouart H, et al. Estimating Glomerular Filtration Rate from Serum Creatinine and Cystatin C. N Engl J Med 2012; 367:20-9.

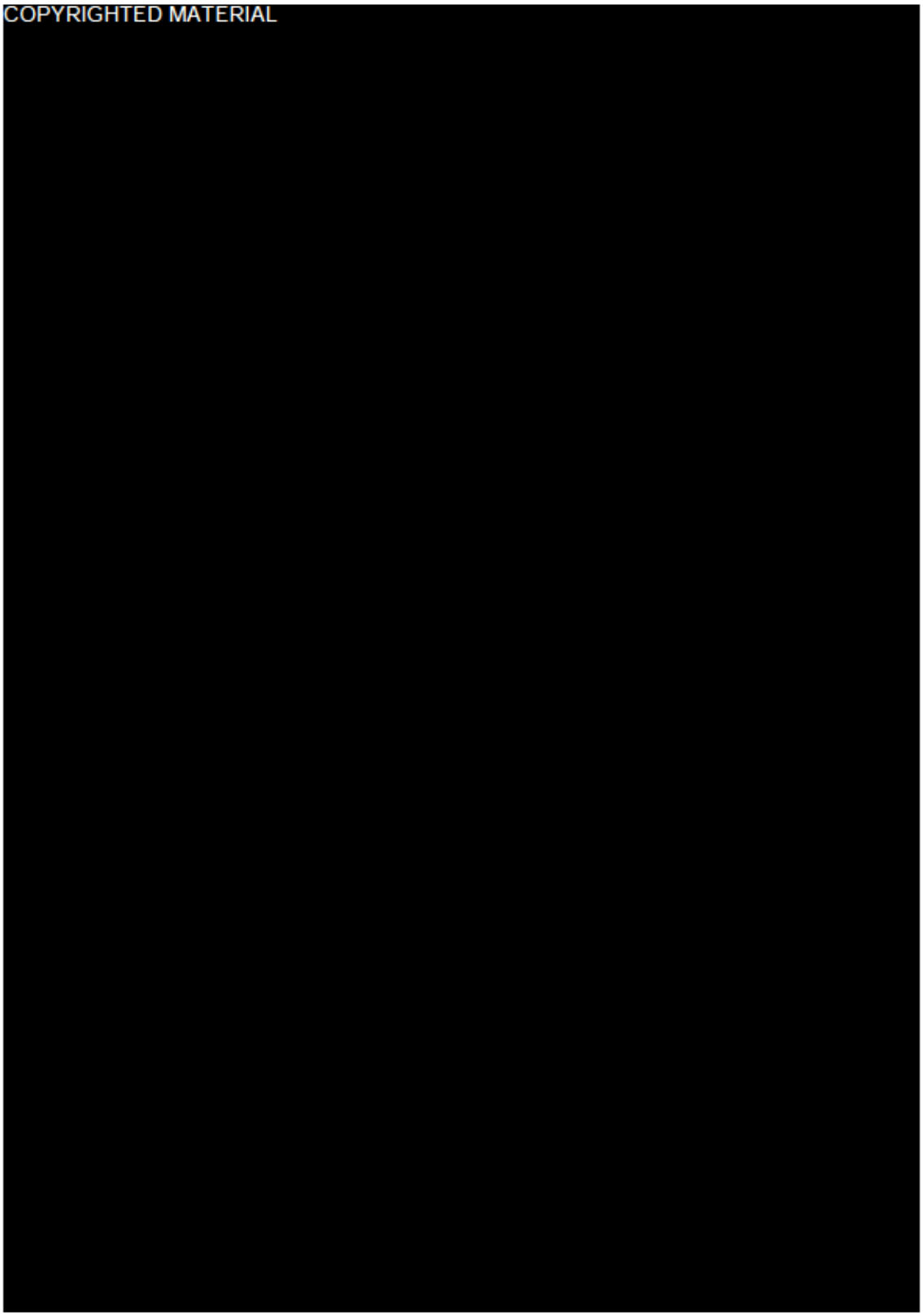
Appendix 2.4.4. AUDIT Total Score

The AUDIT total score is calculated as the sum of scores provided for each of the 10 questions of the AUDIT questionnaire. The range of possible AUDIT total scores is from 0 to 40.

Not all questions in the AUDIT questionnaire have to be answered (see below), however, it is not possible to skip questions in the eDiary (if they should be answered). Therefore there should not be any missing responses (other than those that are allowed to be missing based on the responses to Qs 1, 2 and 3).

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

COPYRIGHTED MATERIAL



090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Ref: Babor TF, Higgins-Biddle JC, Sanuders JB, Monteiro MG. AUDIT – the alcohol use disorders identification test – guidelines for use in primary care – 2nd edition. World

Health Organization; Department of Mental Health and Substance Dependence – 2001.

Appendix 2.5. Definition of Protocol Deviations That Relate to Statistical Analyses/Populations

Appendix 2.5.1. Compliance

Compliance with double-blinded study medication will be calculated using data captured on the Double Blinded treatment CRFs.

Compliance between any two visits in the study will be calculated as:

$$\% \text{ Compliance} = \frac{\#\text{ of pills actually taken since the previous visit} \times 100}{\#\text{ of pills expected to be taken since the previous visit}}$$

The # **of pills actually taken since the previous visit** will be taken directly from the "Actual Dose" field in the Double Blinded treatment CRFs (sites will be instructed to enter the total number of tablets that the participant has consumed since the last on-site visit).

The # **of pills expected to be taken since the previous visit** = (Date of current visit – Date of previous visit)*6

Compliance across the overall study treatment period will be calculated as:

$$\% \text{ Compliance} = \frac{\#\text{ of pills actually taken during the double-blind treatment phase} \times 100}{\#\text{ of pills expected to be taken during the double-blind treatment phase}}$$

The # **of pills actually taken during the double-blind treatment phase** = sum of all "Actual dose" fields for all visits during the double-blind treatment phase (i.e. the sum of the total number of tablets that the participant has consumed during the double-blind treatment phase)

The # **of pills expected to be taken during the double-blind treatment phase** = (Week 48 (V19) or Early Term visit date – Baseline visit date (V5)+1)*6

Appendix 3. Data Set Descriptions

Not applicable

Appendix 4. Statistical Methodology Details

Not applicable

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Appendix 5. List of Abbreviations

Abbreviation	Term
%CDT	Percent carbohydrate deficient transferrin relative to total transferrin
ACCi	Acetyl-CoA carboxylase inhibitor (PF-05221304)
AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
aPTT	Activated partial thromboplastin time
AST	Aspartate aminotransferase
AUDIT	Alcohol use disorders identification test
BID	Bis in die (twice-a-day)
BLQ	Below the limit of quantitation
BMI	Body mass index
BPM	Beats per minute
CAP™	Controlled attenuation parameter
CaPS	CDISC and Pfizer standards
CBL	Change from baseline
CDISC	Clinical Data Interchange Standards Consortium
CDT	Carbohydrate deficient transferrin
CI	Confidence interval
CK18-M30	Cytokeratin-18-M30 fragment
CK18-M65	Cytokeratin-18-M65 fragment
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration equation
CRF	Case report form
DILI	Drug-induced liver injury
DGAT2i	Diacylglycerol acyltransferase 2 inhibitor (PF-06865571)
DR	Dose response
DTS	Data Transfer Specification
EC	Ethics Committee
ECG	Electrocardiogram
E-DMC	External data monitoring committee
eGFR	Estimated glomerular filtration rate
ELF	Enhanced Liver Fibrosis
ER	Exposure response
FAS	Full analysis set
FPG	Fasting plasma glucose
FPI	Fasting plasma insulin
FSH	Follicle-stimulating hormone
GGT	Gamma-glutamyl transferase
HA	Health Authority
HbA1C	Glycated hemoglobin
HDL-C	High density lipoprotein cholesterol

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Abbreviation	Term
HIV	Human immunodeficiency virus
HOMA-IR	Homeostatic Model Assessment for Insulin Resistance
HRQoL	Health-related quality of life
hs-CRP	High-sensitive C-reactive protein
IA	Interim Analysis
ICD	Informed consent document
IME	Important Medical Event
INR	International normalized ratio
IPV	Intra-pathologist variability
IRB	Institutional Review Board
LDL-C	Low density lipoprotein-cholesterol
LFT	Liver function test
LLQ	Lower limit of quantification
LSM	Least-squares mean
MACE	Major adverse cardiovascular event
MCH	Mean corpuscular hemoglobin
MCHC	Mean corpuscular hemoglobin concentration
MCV	Mean corpuscular volume
MedDRA	Medical Dictionary for Regulatory Activities
MELD	Model of end-stage liver disease
MMRM	Mixed-effects model with repeated measures
MRI	Magnetic resonance imaging
MRI-PDFF	Magnetic resonance imaging using proton density fat fraction acquisition
N/A	Not applicable
NAFLD	Nonalcoholic fatty liver disease
NAS	NAFLD Activity Score
NASH	Nonalcoholic steatohepatitis
ND	Not done
NRS	Numeric rating scale
NS	No sample
PCC	Potential Clinical Concern
PCSK9	Proprotein convertase subtilisin/kexin type 9
PD	Pharmacodynamic(s)
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PK	Pharmacokinetic(s)
PRO	Patient-reported outcome
PROMIS	Patient-Reported Measurement Outcome Information System
ProC3	N-terminal propeptide of type III procollagen
ProC6	C-terminal fragment of α3 chain of procollagen type VI
PBL	Percent change from baseline
PIIINP	Amino-terminal propeptide of type III procollagen

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)

Abbreviation	Term
PR	Period of time from the onset of the P wave to the beginning of the QRS complex on an electrocardiogram
PT	Preferred term
QD	Quaque die (once-a-day)
QRS	Combination of Q-, R- and S- wave on an electrocardiogram representing ventricular depolarization
QTc	Corrected QT
QTcF	Corrected QT (Fridericia method)
RBC	Red blood cell
RC	Relatice change
RCBL	Relative change from baseline
RNA	Ribonucleic acid
ROW	Rest of World
RR	Respiratory rate
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SE	Standard error
SI	Système international
SOP	Standard operating procedure
SPDFF	Segmental PDFF
T2DM	Type 2 diabetes mellitus
TEAE	Treatment Emergent Adverse Event
TDD	Total Daily Dose
TIMP-1	Tissue inhibitor of metalloproteinases 1
UA	Urinalysis
ULN	Upper limit of normal
ULQ	Upper limit of quantification
VCTE™	Vibration-Controlled Transient Elastography
CCI	[redacted]
VLDL	Very low-density lipoprotein
VRS	Verbal rating scale
WBC	White blood cell
WLPDFF	Whole Liver PDFF

090177e1a09ab374\Approved\Approved On: 06-May-2024 15:03 (GMT)