STATISTICAL ANALYSIS PLAN	
Study Code	D7830C00004
Edition Number	4.0
Date	25-Jun-2025
EudraCT / EU CT Number	2022-001629-65 / 2023-509704-14

A Randomised, Double-blind, Placebo-controlled, Multi-centre Phase 2b Study to Evaluate the Efficacy, Safety and Tolerability of AZD2693 in Participants with Non-cirrhotic Non-alcoholic Steatohepatitis (NASH) with Fibrosis Who Are Carriers of the *PNPLA3* rs738409 148M Risk Allele

TABLE OF CONTENTS

TITLE PAGE 1
TABLE OF CONTENTS 2
LIST OF ABBREVIATIONS 7
AMENDMENT HISTORY 10
1 INTRODUCTION 18
2 CHANGES TO PROTOCOL PLANNED ANALYSES 18
3 DATA ANALYSIS CONSIDERATIONS 20
3.1 Timing of Analyses 20
3.2 Analysis Populations 20
3.3 General Considerations 22
3.3.1 General Study Level Definitions 23
3.3.2 Visit Window 26
3.3.3 Handling of Unscheduled Visits 29
3.3.4 Multiplicity/Multiple Comparisons 29
3.3.5 Handling of Protocol Deviations in Study Analysis 30
3.3.6 Analysis of data from China, Japan and East Asia cohorts 31
3.3.6.1 Definition of China cohort, Japan cohort and East Asia cohort 31
3.3.6.2 Analysis sets 31
3.3.6.3 Analysis considerations 31
3.4 Alternative approach for main and sensitivity analyses of all binary endpoints 32
4 STATISTICAL ANALYSIS 33
4.1 Study Population 33
4.1.1 Subject Disposition and Completion Status 33
4.1.1.1 Definitions and Derivations 33
4.1.1.2 Presentation 34
4.1.2 Analysis Sets 34
4.1.2.1 Definitions and Derivations 34
4.1.2.2 Presentation 35
4.1.3 Protocol Deviations 35
4.1.3.1 Definitions and Derivations 35
4.1.3.2 Presentation 35
4.1.4 Demographics 35
4.1.4.1 Definitions and Derivations 35
4.1.4.2 Presentation 35
4.1.5 Baseline Characteristics 36
4.1.5.1 Definitions and Derivations 36
4.1.5.2 Presentation 36
4.1.6 Disease Characteristics at Baseline 36
4.1.6.1 Definitions and Derivations 36

4.1.6.2 Presentation 36
4.1.7 Medical History and Concomitant Disease 36
4.1.7.1 Definitions and Derivations 36
4.1.7.2 Presentation 36
4.1.8 Prior and Concomitant Medications 37
4.1.8.1 Definitions and Derivations 37
4.1.8.2 Presentation 37
4.1.9 Study Drug Compliance 37
4.1.9.1 Definitions and Derivations 37
4.1.9.2 Presentation 38
4.2 Endpoint Analyses 38
4.2.1 Primary Endpoint: NASH resolution without worsening of fibrosis based on histology after Week 52 42
4.2.1.1 Definition 42
4.2.1.2 Derivations 43
4.2.1.3 Handling of Dropouts and Missing Data 44
4.2.1.4 Primary Analysis of Primary Endpoint 45
4.2.1.5 Sensitivity Analyses of the Primary Endpoint 45
4.2.1.6 Supplementary Analyses of the Primary Endpoint 48
4.2.1.7 Subgroup Analyses 48
4.2.2 Key Secondary Endpoint: At least one stage of liver fibrosis improvement with no worsening of NASH based on biopsy after Week 52 49
4.2.2.1 Definition 49
4.2.2.2 Derivations 49
4.2.2.3 Handling of Dropouts and Missing Data 49
4.2.2.4 Primary Analysis of Secondary Endpoint 49
4.2.2.5 Sensitivity Analyses of the Secondary Endpoint 49
4.2.2.6 Supplementary Analyses of the Secondary Endpoint 49
4.2.2.7 Subgroup Analyses 49
4.2.3 Secondary Endpoint: ≥ 2-point improvement from baseline in NAS based on biopsy after week 52 49
4.2.3.1 Definition 49
4.2.3.2 Derivations 49
4.2.3.3 Handling of Dropouts and Missing Data 50
4.2.3.4 Primary Analysis of the Secondary Endpoint 50
4.2.3.5 Sensitivity Analyses of the Secondary Endpoint 50
4.2.3.6 Supplementary Analyses of the Secondary Endpoint 50
4.2.3.7 Subgroup Analyses 50
4.2.4 Secondary Endpoint: improvement in fibrosis by at least one stage based on biopsy after Week 52 50
4.2.4.1 Definition 50
4.2.4.2 Derivations 50
4.2.4.3 Handling of Dropouts and Missing data 50
4.2.4.4 Primary Analysis of the Secondary Endpoint 50
4.2.4.5 Sensitivity Analyses of the Secondary Endpoint 50
4.2.4.6 Supplementary Analyses of the Secondary endpoint 50

4.2.4.7 Subgroup Analyses 50
4.2.5 Safety endpoint 50
4.2.6 Exploratory endpoint: resolution of NASH and improvement in fibrosis by at least one stage based on biopsy after Week 52 51
4.2.6.1 Definition 51
4.2.6.2 Derivations 51
4.2.6.3 Handling of Dropouts and Missing data 51
4.2.6.4 Primary Analyses of the Exploratory Endpoint 51
4.2.6.5 Additional Analyses of the Exploratory Endpoint 51
4.2.6.6 Subgroup Analyses 51
4.2.7 Exploratory endpoint: change from baseline to post-Week 52 in NASH-related histological parameters 51
4.2.7.1 Definition 51
4.2.7.2 Derivations 51
4.2.7.3 Handling of Dropouts and Missing data 51
4.2.7.4 Primary Analyses of the Exploratory Endpoint 52
4.2.7.5 Additional Analyses of the Exploratory Endpoint 52
4.2.7.6 Subgroup Analyses 52
4.2.8 Exploratory endpoint: change from baseline to post-Week 52 in circulating markers of hepatic injury and apoptosis or necrosis 52
4.2.8.1 Definition 52
4.2.8.2 Derivations 52
4.2.8.3 Handling of Dropouts and Missing data 52
4.2.8.4 Primary analyses of the Exploratory Endpoint 52
4.2.8.5 Additional Analyses of the Exploratory Endpoint 54
4.2.8.6 Subgroup Analyses 54
4.2.9 Exploratory endpoint: change from baseline to post-Week 52 in circulating biomarkers of fibrosis and fibrogenesis 55
4.2.9.1 Definition 55
4.2.9.2 Derivations 55
4.2.9.3 Handling of Dropouts and Missing data 55
4.2.9.4 Primary Analyses of the Exploratory Endpoint 55
4.2.9.5 Additional Analyses of the Exploratory Endpoint 55
4.2.9.6 Subgroup Analyses 55
4.2.10 Exploratory endpoint: change from baseline to post-Week 52 in lipids parameters 55
4.2.10.1 Definition 55
4.2.10.2 Derivations 55
4.2.10.3 Handling of Dropouts and Missing data 55
4.2.10.4 Primary Analyses of the Exploratory Endpoint 56
4.2.10.5 Additional Analyses of the Exploratory Endpoint 56
4.2.10.6 Subgroup Analyses 56
4.2.11 Exploratory endpoint: change from baseline to post-Week 52 in different non-invasive algorithms developed as marker of NAFLD or NASH disease activity 56
4.2.11.1 Definition 56

4.2.11.2 Derivations ... 56
4.2.11.3 Handling of Dropouts and Missing data ... 56
4.2.11.4 Primary Analyses of the Exploratory Endpoint ... 57
4.2.11.5 Additional Analyses of the Exploratory Endpoint ... 57
4.2.11.6 Subgroup Analyses ... 57
4.2.12 Exploratory endpoint: change from baseline to post-Week 52 in biomarkers for NASH comorbidities ... 57
4.2.12.1 Definition ... 57
4.2.12.2 Derivations ... 57
4.2.12.3 Handling of Dropouts and Missing data ... 57
4.2.12.4 Primary Analyses of the Exploratory Endpoint ... 57
4.2.12.5 Additional Analyses of the Exploratory Endpoint ... 57
4.2.12.6 Subgroup Analyses ... 57
4.2.13 Exploratory endpoint: pharmacokinetics of AZD2693 ... 58
4.2.14 Exploratory endpoint: immunogenicity of AZD2693 ... 58
4.2.15 Exploratory endpoint: change from baseline to post-Week 52 in patient-reported outcomes ... 58
4.2.15.1 Definition ... 58
4.2.15.2 Derivations ... 59
4.2.15.3 Handling of Dropouts and Missing data ... 62
4.2.15.4 Primary Analyses of the Exploratory Endpoint ... 62
4.2.15.5 Additional Analyses of the Exploratory Endpoint ... 62
4.2.15.6 Subgroup Analyses ... 63
4.2.16 Exploratory endpoint: liver stiffness and liver steatosis using imaging analysis ... 63
4.2.16.1 Definition ... 63
4.2.16.2 Derivations ... 63
4.2.16.3 Handling of dropouts and Missing data ... 63
4.2.16.4 Primary Analyses of the Exploratory endpoint ... 63
4.2.16.5 Additional Analyses of the Exploratory Endpoint ... 63
4.2.16.6 Subgroup Analyses ... 63
4.3 Pharmacodynamic Endpoint(s) (if not already covered as endpoint variables) ... 63
4.3.1 Analysis ... 63
4.3.2 Definitions and Derivations ... 63
4.3.3 Presentation ... 63
4.4 Pharmacokinetics (if not already covered as endpoint variables) ... 64
4.4.1 Pharmacokinetic concentration data ... 64
4.4.2 Pharmacokinetic parameters data ... 65
4.5 Immunogenicity (if applicable and if not already covered as endpoint variables) ... 65
4.6 Safety Analyses (if not already covered as endpoint variables) ... 67
4.6.1 Exposure ... 67
4.6.1.1 Definitions and Derivations ... 67
4.6.1.2 Presentation ... 68
4.6.2 Adverse Events ... 69

4.6.2.1 Definitions and Derivations 69
4.6.2.2 Presentation 69
4.6.3 Clinical Laboratory, Blood and Urine Samples 71
4.6.3.1 Definitions and Derivations 71
4.6.3.2 Presentations 76
4.6.4 Vital Signs 76
4.6.4.1 Definitions and Derivations 76
4.6.4.2 Presentations 77
4.6.5 Electrocardiogram 78
4.6.5.1 Definitions and Derivations 78
4.6.5.2 Presentations 78
4.6.6 Other Safety Assessments 79
4.6.6.1 Definitions and Derivations 79
4.6.6.2 Presentations 79
5 INTERIM ANALYSIS 80
6 REFERENCES 80
7 APPENDIX 80

LIST OF TABLES

Table 1 Analysis populations 20
Table 2 Sample size per dose level 22
Table 3 Study periods 26
Table 4 Analysis visit windows for non-biopsy/non-imaging parameters 27
Table 5 Analysis visit window for biopsy parameters 28
Table 6 Analysis visit window for liver imaging parameters 29
Table 7 List of all study objectives 39
Table 8 Description of non-alcoholic fatty liver disease activity score 43
Table 9 Description of fibrosis stage 44
Table 10 Formula for MMRM for log-normally parameters 54
Table 11 Normal reference ranges for INR and GLDH 75
Table 12 Vital signs normal reference ranges 77

LIST OF ABBREVIATIONS

List abbreviations and definitions of specialized or unusual terms, measurements, or units. Examples are provided below. These can be modified at study level.

Abbreviation or Specialized Term	Definition
ADA	Anti-drug antibody
AE	Adverse event
AESI	Adverse event of special interest
ApoB	Apolipoprotein B
APRI	AST to platelet ratio index
ASO	Antisense Oligonucleotide
AST	Aspartate aminotransferase/transaminase
ANCOVA	Analysis of covariance
BMI	Body mass index
CK-18	Cytokeratin 18
CSP	Clinical Study Protocol
CSR	Clinical Study Report
CMH	Cochran-Mantel-Haenszel
DILI	Drug Induced Liver Injury
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
FAS	Full analysis set
FIB-4	Fibrosis-4
GGT	Gamma Glutamyl Transferase
GLDH	Glutamate dehydrogenase
HDLc	High Density Lipoprotein cholesterol
IM	Immunogenicity
IP	Investigational Product
IPD	Important Protocol Deviation
IRT	Interactive Response Technology
iSMC	Independent Safety Monitoring Committee
LATAM	Latin American
LDLc	Low Density Lipoprotein cholesterol
LLN	Lower Limit of Normal
LLOQ	Lower Limit of Quantification
LSMD	Least square mean difference
LSmean	Least Squares Mean
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed model for repeated measures

Abbreviation or Specialized Term	Definition
NAS	NAFLD activity score
MRE	Magnetic Resonance Elastography
MRI-PDFF	Magnetic Resonance Imaging-based Protocol Density Fat Fraction
NASH	Non-alcoholic Steatohepatitis
NFS	NAFLD fibrosis score
PD	Pharmacodynamic
PGIS	Patient Reported Outcome Measure for Patient Global Impression of Severity
PNPLA3	Patatin-like Phospholipase Domain-containing Protein 3
PK	Pharmacokinetics
PRO	Patient Reported Outcome
PT	Preferred Term
RTSM	Randomisation and Trial Supply Management
SAE	Serious adverse event
SAF	Steatosis, Activity, Fibrosis
SAP	Statistical Analysis Plan
SC	Subcutaneous
SoA	Schedule of Activities
SOC	System Organ Class
T2DM	Type 2 diabetes mellitus
TBL	Total bilirubin
UACR	Urine albumin-creatinine ratio
ULN	Upper Limit of Normal
ULOQ	Upper Limit of Quantification

AMENDMENT HISTORY

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
N/A	5/9/2023	Initial approved SAP	Yes	N/A
Secondary endpoint(s)	25Oct2023	Heterozygous population removed from all objectives.	Yes	To align with the CSP amendment.
Data presentation	25Oct2023	Definitions of analyses sets adjusted. All analyses sets apart from enrolled set will include an extra condition: participant has to be homozygous for the risk allele.	No	To align with the objectives of the CSP amendment.
Data presentation	25Oct2023	Analysis window visits adjusted.	Yes	To better align with the CSP visit windows.
Data presentation	25Oct2023	Handling of unscheduled visits added. After assignment of analysis visit windows only scheduled time points are presented in Tables.	Yes	Clarification.
Data presentation	25Oct2023	Definition of Concomitant Medication adjusted.	Yes	Clarification.
Data presentation	25Oct2023	Selection of baseline observation presented in more detail.	Yes	Clarification.
Data presentation	25Oct2023	Definition of exposure adjusted.	Yes	Simplification.

Secondary and exploratory endpoints	15Feb2024	Subgroup analyses removed.	Yes	To reduce number of subgroup analyses.
Statistical analysis method for the exploratory endpoints.	15Feb2024	Distribution of continuous biomarkers will be explored.	Yes	To account for log-normal distributions.
Data presentation	15Feb2024	ADA definitions adjusted.	Yes	To align with definitions used in TFLs.
Data presentation	15Feb2024	Presentation of AESI clarified	Yes	Clarification
Data presentation	15Feb2024	Definition of baseline ALT for safety purpose simplified.	No	Simplification and clarification.
Data presentation	15Feb2024	Presentation and imputation rules for log-normally distributed parameters added.	Yes	Additional information needed to apply normality checks.
Data presentation	15Feb2024	Definition of actual arm added.	Yes	Clarification.
Data presentation	15Feb2024	Multiple textual corrections.	Yes	Corrections and clarifications.
Statistical analysis of the primary and secondary endpoint	6Mar2024	Supplementary analyses added.	No	Request from FDA.
Data presentation	6Mar2024	FAS and SAF for randomised heterozygous subjects defined.	No	To account for heterozygous subjects randomised before CSP amendment.
Data presentation	6Mar2024	Prioritize central labs in case if both local and central measurements are available in the same visit window.	Yes	Clarification.
Data presentation	6Mar2024	Allow only central labs for baseline ALT calculation	Yes	Clarification.

Visit Windows	14Jan2025	Extended analysis windows for liver imaging due to change in schedule of assessments	Yes	To account for extending visit week 26 assessments
Data presentation	15Jan2025	Study periods updated. AE and safety lab sections updated to mention that some summaries will be created for on-treatment an on-study periods.	Yes	Request from safety team
Data presentation	15Jan2025	Clarified definition of baseline for biopsy data (resubmission of screening will be prioritized)	Yes	Agreement with clinical team
Data presentation	16Jan2025	Important protocol deviations analysis population FAS will be used for the summary	Yes	To align with the listing population
Data presentation	20Jan2025	Participant Disposition summary updated	Yes	Clarification
Data presentation	20Jan2025	Removed text prior to Week 54 when looking at liver outcomes we consider all	Yes	Agreement with study team
Data presentation	29Jan2025	Updated significance level from 0.1 to 0.05 and confidence interval from 90% to 95%.	Yes	In response to Health Authority advice

Data presentation	29Jan2025	Updated all names of endpoints to align with the protocol text and added derivation for HOMA-IR in 4.2.12.2	Yes	In response to Health Authority advice and derivation added after discussion with clinical team
Data presentation	17Feb2025	Updated throughout the document the exploratory analysis based on choice to analyse change or percent change from baseline from study team	Yes	Agreement with the study team
Data presentation	24Feb2025	Added text on minimum presentation of results in case of few numbers of responders in either treatment group	Yes	Agreement with the statisticians' team
Data presentation	6Mar2025	Removed Breslow Day from subgroup analysis	Yes	Correction
Primary Endpoint	13Mar2025	Clarified worsening of fibrosis and concluding resolution of NASH	Yes	Clarification
Exploratory Endpoint	13Mar2025	Clarified ANCOVA model in 4.2.7.4 and added derivation in 4.2.11.2 for APRI and NFS	Yes	Clarification and derivation based on discussion with clinical team

Binary Endpoints	21Mar2025	Updated SAP text for sensitivity analyses 1 and 2 to further explain and clarify the approach	Yes	Clarification and explanation
Data Presentation	07Apr2025	Modified Table 10 based on the updated TLF shell	Yes	Agreement with statisticians' team
Exploratory Endpoint	08Apr2025	Updated SAP text for exploratory analyses to further explain and clarify the approach	Yes	Explanation, clarification
Safety Endpoint	08Apr2025	Added shift laboratory tables and other tables in safety part and vital signs figure	Yes	Request from safety team
Binary Endpoints	11Apr2025	Added response rates and random seed to sections of sensitivity analyses 1 and 2	Yes	In response to Health Authority advice
General revision throughout the document	18Apr2025	Align with changes in protocol amendments (considering v4.0 and v5.0)	Yes	Align with changes in versions 4 and 5 of protocol
Data presentation	01May2025	Added data handling for ADA data following AZ standards in sections 3.3.1 and	Yes	Intention to treat to include all subjects from ADA analysis set in the summary

Data presentation	15May2025	Added text in specific sections of Section 4 for the analyses expected for Asia Cohorts. Redefined the Asia Cohorts in Section 3.3.6.1	Yes	As per agreement with the statisticians' team.
Primary and key secondary endpoint	18May2025	Section 3.3.4 modified for clarification that multiple testing will be applied to primary and key secondary endpoint	Yes	Explanation, clarification
Biopsy/imaging window for post-Week 52	05Jun2025	Extended biopsy/imaging window for post-Week 52 in section 3.3.2	No	To include all cases/participants in the analysis
Shift table for safety parameters (low, normal, high categories)	05Jun2025	Replaced worst with minimum/maximum based on the safety parameter	No	Explanation, clarification
Addition of imputation rules for CM partial start date	05Jun2025	Added same imputation rule for CM partial start date as for AE. Clarified that imputations will take place for treated participants and not screen failures	No	To classify/summarize all CM to prior/concomitant medications
Intercurrent events	05Jun2025	Added listing of intercurrent events in section 4.2	No	To summarize/clarify the intercurrent events listing

ECG presentation	06Jun2025	Clarified that all ECG are listed and only QTcF and heart rate will be summarised	No	To summarize/clarify the intercurrent events listing
Data Presentation PK	06Jun2025	Explained how to handle the Non-Quantifiable Concentrations	No	Explanation, clarification
Data Presentation ADA	06Jun2025	Added ADA positive/negative status at the sample level	No	Explanation, clarification
Data Presentation - Safety	10Jun2025	Added to "Abnormalities in laboratory" the categories (>ULN, <LLN) for all safety parameters and to "Other tables" the categories >5 x baseline, >5xbaseline for >=12 weeks for the summary of participants with elevated ALT, section 4.6.3	No	Request by the safety team
New section about alternative approach in case primary analyses is not feasible for all binary endpoints	25Jun2025	Added section 3.4 to follow alternative for the primary analyses and all binary endpoints	No	To still assess the binary endpoints in case the primary stratified analyses will not be feasible.

1 INTRODUCTION

The purpose of this document is to provide details for the statistical analysis of study D7830C00004 supporting the clinical study report. The reader is referred to the CSP and the CRF for details of study conduct and data collection.

2 CHANGES TO PROTOCOL PLANNED ANALYSES

Change refers to	Description of change
Definition of the following analysis sets: Randomly assigned to study intervention, Safety, Full analysis set, PK analysis set, ADA set. Extra analysis sets defined: Full analysis set, Heterozygous subjects, Safety analysis set, Heterozygous subjects.	Addition of the following condition: participants must be homozygous for the risk allele. Separate analysis sets defined for randomised heterozygous subjects.
Definition of ADA set	Participants will be analysed according to the treatment they actually received instead of treatment they were randomised to.
Changes in Schedule of Activities for extra pre-dose platelet assessments after Visit 11 and ECG assessments	Removal of requirement for extra platelet count sample after Visit 11. Removal of ECG assessments from study Visit 4, Visit 7, and Visit 20. Removal of ECG recordings at 60 to 120 minutes pre-dose and 4 hours post-dose at Visit 3 and Visit 13.
Removal of cholesteryl ester biomarker.	Removal of exploratory endpoint ‘change from baseline in plasma cholesteryl ester 16:1/16:0 ratio’ and reference to cholesteryl ester biomarker.
Aligned definition of worsening of NASH to health authority guidelines	Definition of worsening of NASH updated to include ‘steatosis’ as a criterion.
Change to significance level and confidence interval.	The significance level has been changed from 0.1 to 0.05, and the confidence

	interval has been changed from 90% to 95%.
MRI-PDFF endpoint updated.	‘Liver volume’ removed and ‘volume corrected HFF’ added as an MRI-PDFF endpoint.

3 DATA ANALYSIS CONSIDERATIONS

3.1 Timing of Analyses

The following analyses are planned for the trial: multiple safety analyses and one final analysis. Safety analyses are planned every three months, with a possibility of changes to this schedule depending on the emerging data. The first safety analysis is planned 3 months after first dose of the first randomised participant. Safety analyses will be evaluated by an internal Safety Monitoring Committee (iSMC) within AstraZeneca. Details on safety analyses are described in the iSMC charter.

A primary analysis will be performed when all participants have completed the treatment period and last biopsy. Key efficacy and safety/tolerability data will be assessed. The design or the conduct of this study will not be impacted by the primary analysis results. The primary analysis will be performed by an unblinded team. The study team will remain blinded until the end of the follow-up. Further details regarding maintenance of the blind for the study team during execution of primary analysis, team responsibilities and other processes are described in the primary analysis charter.

Final analysis will be performed when all randomised participants have completed the study. No interim analyses are planned.

3.2 Analysis Populations

Table 1 Analysis populations

Population	Description
Enrolled	All participants who sign the ICF*
Randomly assigned to study intervention	All participants homozygous for the risk allele randomised into the study
Safety analysis set	All participants homozygous for the risk allele randomised into the study who have received at least one dose of study intervention. Participants will be analysed according to the treatment they actually received.
Safety analysis set, Heterozygous subjects	All participants heterozygous for the risk allele randomised into the study who have received at least one dose of study

	intervention. Participants will be analysed according to the treatment they actually received.
Full analysis set	All participants homozygous for the risk allele randomised to study intervention who received at least one dose of study intervention. Participants will be analysed according to the treatment they were randomised to.
Full analysis set, Heterozygous subjects	All participants heterozygous for the risk allele randomised to study intervention who received at least one dose of study intervention. Participants will be analysed according to the treatment they were randomised to.
PK analysis set	All participants homozygous for the risk allele randomised into the study who have at least one measured plasma concentration of AZD2693 full-length ASO. Participants will be analysed according to the treatment they actually received.
ADA set	All participants homozygous for the risk allele randomly assigned to study treatment and who take at least 1 dose of study treatment and had at least one ADA assessment available. Participants will be included in the analysis according to the treatment they actually received.

***Refers to ICF signed at Visit 1**

3.3 General Considerations

Study design:

This is a randomized Phase 2b, double-blind, placebo-controlled, multi-centre, multinational trial comparing once monthly SC administration of AZD2693 in participants with non-cirrhotic NASH (F2-F3) carrying the PNPLA3 rs738409 148M risk allele.

The study includes a screening period of up to 10 weeks (inclusive of 2-week early genotyping assessment) followed by randomisation, 52-week treatment period and about 12-week safety follow-up period.

Randomisation:

Randomisation is performed after screening period, on first day of receiving IP. All participants will be centrally assigned to randomised study intervention using an IRT/RTSM. The IRT/RTSM will provide to the investigator(s) or pharmacists the kit identification number to be allocated to the participant at the dispensing visit.

Participants who are eligible according to the inclusion/exclusion criteria will be randomised in a 2:1:2:1 ratio to receive AZD2693 25 mg, placebo matching AZD2693 25 mg, AZD2693 50 mg or placebo matching AZD2693 50 mg.

Block randomisation will be used. Randomisation will be stratified for T2DM (yes/no) and stage of fibrosis (F2/F3).

Primary objective and hypothesis:

The primary objective is to demonstrate superiority of AZD2693 compared to placebo on the resolution of NASH without worsening of fibrosis after Week 52.

Sample size:

The total planned sample size is 180 participants.

Table 2 Sample size per dose level

Dose group	Treatment	N (180 total)
25 mg	AZD2693	60
	Placebo	30

50 mg	AZD2693	60
	Placebo	30

For participants homozygous for the PNPLA3 rs738409 148M risk allele, about 60 participants in each AZD2693 dose arm and about 60 participants in pooled placebo arm will provide more than 90% power for the test of each dose versus placebo to detect a difference of 35% in NASH resolution without worsening of fibrosis assuming a response rate to placebo of 20%. The calculations were done for an independent samples chi-square test and assume a 2-sided alpha level of 0.05.

3.3.1 General Study Level Definitions

General considerations

- **Calculation of study day**
  - Study day will be calculated based on the actual assessment date. All data will be summarized in relation to date of first IP administration.
  - If actual assessment date is prior to the first IP administration, then study day will be:
    Study day = Actual assessment date – first IP administration date.
  - If actual assessment date is on or after the first IP administration, then study day will be:
    Study day = Actual assessment date – first IP administration date + 1.

- **Definition of baseline (apart from ALT for safety)**
  - In general, the last non-missing measurement prior to first application of IP will serve as the baseline measurement for efficacy and safety endpoints.
  - If multiple visits are equally eligible to assess participant status at baseline (for example assessments both on the same date/time prior to first dose of IP administration), the average (for numeric values)/best (for categorical values) of the measurements will be taken as the baseline value
  - If two assessments are available on the same day prior to the first dose of IP, of which one has available time and the other not, the assessment with available time will be counted as baseline, except for definition of baseline for the biopsy data. For the biopsy data, in case of two assessments on the same day prior to the first dose of IP, the record of the visit of resubmission

of screening will be used as baseline, irrespective of time being present or not.

- If several assessments are performed on the same day as first dose of IP administration, only those with time recorded prior to IP time of administration will be used to assess participant status at baseline (i.e. if time of assessment on the same day as IP administration is missing, the assessment will be assumed to be collected after administration).
- If no value exists before the first dose of any study drug, then the baseline value will be treated as missing.

- **Baseline ALT for safety**

  Baseline ALT will be calculated as an average of two last pre-treatment measurements. If only one measurement is available, this value will be used as baseline ALT.
  An AstraZeneca-appointed central laboratory should be used for all baseline ALT measurements (no local or historical tests permitted).

- **Baseline of ECG parameters**

  Baseline for each collected ECG parameter will be calculated as an average of available replicates at the last visit prior to IP. In case of only one available replicate at the last visit, this replicate will be taken as baseline values.

- **Derivation of absolute change from baseline**

  Absolute change from baseline = post-IP value – baseline value.

Missing data

Missing data are not imputed unless otherwise specified for statistical modelling. Only AE start and end date and concomitant medications (CMs) start and end date are imputed for the treated participants. Imputation rules are reported below.

- **Imputation of AE and of CM end date**

  Completely missing AE end dates are not imputed. Partially missing AE end dates are imputed as below:

  - If the AE/CM is ongoing, the end date is set to missing
  - If the AE/CM is not ongoing, and if only the day is missing: Assume the last day of the collected month

- If the AE/CM is not ongoing, and both the day and the month are missing: Assume 31-DEC of the collected year
- If the day is complete, but the month is missing, assume that the day is also missing and follow the above instructions.

After applying these rules, if the imputed AE/CM end date is after the end of study date, the AE/CM is set to ongoing and the AE/CM end date to missing.

- **Imputation of AE and of CM start date**

Before proceeding with the AE/CM start date imputation, the AE/CM end date is imputed as described in the previous section. Only partial AE/CM start dates are imputed. Dates which are completely missing are not imputed. Partial dates are imputed as described below:

- If the day is missing and the month and year are different from the month and year of the first dose of IP, assume 01-MMM-YYYY. If the month and year are the same as the first dose of IP month and year and the end date is on or after (including ongoing / missing) the first dose of IP, then assume the date of the first dose of IP. If the month and year are the same as the first dose of IP month and year and the end date is prior to the first dose of IP, then assume the end date.
- If the month is missing and the year is different from the year of first dose of IP, assume 01-JAN-YYYY of the collected year. If the year is the same as the first dose of IP year and the end date is on or after (including ongoing / missing) the first dose of IP, then assume the date of the first dose of IP. If the year is the same as the first dose of IP and the end date is prior to the first dose of IP, then assume the end date.

- If the day is complete, but the month is missing assume that the day is also missing and follow the above instructions.

After applying these rules, if the imputed AE/CM start date is after a complete AE/CM end date, then assume the same date as the complete AE/CM end date; if the end date is missing and the imputed AE/CM start date is after the end of study date then assume the same date as the study end date.

- **Missing anti-drug antibody (ADA) baseline**
  Participants with a missing baseline ADA assessment will be assumed to be ADA negative as a conservative approach to ensure that all subjects are included in all analyses.

Summary statistics for exploratory endpoint biomarkers

Result and change from baseline for laboratory parameters will be summarised using mean, SD, median, Min and Max. Percent change from baseline summary will include additionally geometric mean (gmean) and geometric coefficient of variance (%) (gCV). For gmean and gCV calculation, in case of zero records, regardless of treatment group assignment in all subjects at either baseline or post-baseline, impute half of the minimum of the non-zero records prior to log transformation and determine the lowest value present at either baseline or post-baseline.

Actual arm definition

Actual arm will be determined based on majority (>50%) of received doses. In the case of a tie, the randomized treatment will be used. In case if actual arm cannot be determined, randomised treatment will be used as actual.

Study periods

The study includes a screening period of up to 10 weeks (inclusive of 2-week early genotyping assessment) followed by randomisation, 52-week treatment period and about 12-week safety follow-up period. Due to the long half-life of the compound all measurements obtained after first dose will be considered on-study. Precise definition of each study period is given in Table 3.

Table 3 Study periods

Study period	Period window
Screening	Before first dose of IP*
On-treatment	From first dose of IP up to and including 30 days post last dose of IP**
On-study	From first dose of IP to last visit

* Includes all measurements collected before first dose of IP. If the measurement is collected on the day of first dose of IP but the time of collection (or the planned timepoint) cannot determine if the measurement was before or after first dose of IP, then it will be considered as collected after first dose of IP.
** Includes measurements up to 30 days post last dose of IP, irrespective of time.

3.3.2 Visit Window

For the purpose of the statistical analysis, collected data (apart from PK concentrations) are allocated to the analysis visit using analysis visit windows. Separate windows are defined for non-biopsy/non-imaging data, for biopsy and for imaging data. In PK analyses nominal visits are used (unscheduled visits will not be part of the Tables but only presented in listings). Study day 1 refers to the date the participant receives the first dose of IP (randomisation date). Analysis visit windows for the non-biopsy/non-imaging data are

constructed in an exhaustive manner, giving every measurement a potential to fall into an analysis visit window. The complete list of analysis windows for non-biopsy/non-imaging parameters is provided in Table 4.

After applying the analysis visit windows, analysis visits which were not planned in the Schedule of Activities for a specific parameter are classified as Unscheduled visits. For example, ECG is planned at Screening and Weeks 1, 4, 12, 24, 36, 52, 58, 64. Visits not belonging to this set will be considered Unscheduled.

The biopsy should occur in Week 54 (scheduled study day 372). However, every biopsy that takes place after 36 weeks of treatment will be classified as a scheduled biopsy. This is in line with the eligibility for an exit biopsy in case of withdrawal from the study. It is highly unlikely that early biopsies will occur if a participant continues with the treatment. A window of 40 days is accepted for delayed biopsies occurring beyond target day of Week 54.

If there is more than one assessment in the same visit window, the assessment closest to the scheduled visit day is selected for analysis.

If there are 2 records that are equidistant from the target day based on date, the later record will be selected. For continuous parameters, if there is more than 1 record on the selected day, the average will be taken. For categorical parameters, if there is more than 1 record on the selected day, the value with worse severity will be selected. In case of more than 1 record on the selected day for ECG, the continuous parameters attached to the measurement with worse overall assessment will be selected. If overall assessments are equal, the average of continuous ECG parameters will be taken.

In case if central and local values are available for the selected measurement in a particular visit window, central values will be used for the analysis.

Table 4 Analysis visit windows for non-biopsy/non-imaging parameters

CSP visit number	Analysis visit	Scheduled visit day	CSP window	Analysis window (days)
Visit 1 and Visit 2	Screening	1		Before first IP dose
Visit 3	Week 1	1		1-4
Visit 4	Week 2	8	5-11	5-14
Visit 5	Week 4	22	19-25	15-28
Visit 6	Week 6	36	33-39	29-42

Visit 7	Week 8	50	47-53	43-56
Visit 8	Week 10	64	61-67	57-70
Visit 9	Week 12	78	75-81	71-84
Visit 10	Week 14	92	89-95	85-98
Visit 11	Week 16	106	103-109	99-119
Visit 12	Week 20	134	131-137	120-147
Visit 13	Week 24	162	159-165	148-168
Visit 14	Week 26	176	173-179	169-182
Visit 15	Week 28	190	187-193	183-203
Visit 16	Week 32	218	215-221	204-231
Visit 17	Week 36	246	243-249	232-259
Visit 18	Week 40	274	271-277	260-287
Visit 19	Week 44	302	299-305	288-315
Visit 20	Week 48	330	327-333	316-343
Visit 21	Week 52	358	355-361	344-362
Visit 22	Week 54	372	365-379	363-387
Visit 23	Week 58	400	397-403	388-420
Visit 24	Week 64	442	439-445	421-last visit

Table 5 Analysis visit window for biopsy parameters

Analysis visit	Scheduled visit day	Analysis window (days)
Screening	1	Before first IP dose
Week 54	372	232-412 (start of Week 36 until Scheduled day of Week 54 +40 days)

Table 6 Analysis visit window for liver imaging parameters

Analysis visit	Scheduled visit day	Analysis window (days)
Screening	1	Before first IP dose
Week 26	176	169-222
Week 54	372	232-412 (start of Week 36 until Scheduled day of Week 54 +40 days)

3.3.3 Handling of Unscheduled Visits

Unscheduled visits will not be presented in Tables and Figures but will be part of the Listings.

For biopsy and imaging, measurements taken outside the analysis window will be classified as Unscheduled visits. For the analysis of the biopsy endpoints Unscheduled biopsy will be treated as missing. Detailed handling of missing data is described in Section 4.2. Unscheduled biopsies and imaging data will be presented in listings.

For summaries of extreme values all post baseline values are used to determine the maximum/minimum/worst value recorded on-study (regardless of where it falls in an interval or of whether the value is closest to the scheduled visit date), including those collected at unscheduled visits.

For ECG, the last on-study observation will be based on scheduled and unscheduled measurement.

3.3.4 Multiplicity/Multiple Comparisons

Adjustment for multiple comparisons (testing of 2 doses) against placebo will be applied to the primary and key secondary endpoint. To control familywise error at 5%, hierarchical testing will be applied.

The following hypotheses for the primary endpoint are defined:

- $H_{NULL_NRES_50\ mg}$: There is no difference in the proportion of participants achieving NASH resolution without worsening of fibrosis between AZD2693 50 mg and the pooled placebo

- $H_{NULL_NRES_25\ mg}$: There is no difference in the proportion of participants achieving NASH resolution without worsening of fibrosis between AZD2693 25 mg and the pooled placebo

The following hypotheses for the key secondary endpoint are defined:

- $H_{NULL_IFIB_50\ mg}$: There is no difference in the proportion of participants achieving at least one stage of liver fibrosis improvement with no worsening of NASH between AZD2693 50 mg and the pooled placebo

- $H_{NULL__IFIB_25\ mg}$: There is no difference in the proportion of participants achieving at least one stage of liver fibrosis improvement with no worsening of NASH between AZD2693 25 mg and the pooled placebo

Significance testing will be performed in the following order: $H_{NULL_NRES_50\ mg}$, $H_{NULL_IFIB_50\ mg}$, $H_{NULL_NRES_25\ mg}$, $H_{NULL__IFIB_25\ mg}$. Each time, if $p < 0.05$ then the next hypothesis will be tested at alpha=0.05 (2-sided). Thus, the 50mg dose will be tested first and 25mg last.

For all other comparisons (secondary and exploratory endpoints, subgroup analyses) no multiplicity correction will be applied.

3.3.5 Handling of Protocol Deviations in Study Analysis

Study specific important protocol deviations (IPD) are defined in the Protocol Deviation Plan. Important protocol deviations may include, but are not limited to the following:

- Inclusion Criteria Deviations
- Exclusion Criteria Deviations
- Discontinuation Criteria for study product met but participant not withdrawn from study treatment
- Subject received incorrect IP
- Subject received incorrect dose of IP
- Subject received concomitant medications defined as prohibited in the CSP.

Incidence of important protocol deviations are summarised by deviation categories. The number and percentage of participants in the full analysis set with at least one IPD per category are presented. Separate summary will be created for East Asia and China cohorts. A listing is provided with the important protocol deviation details.

None of the deviations result in subjects being excluded from any analysis populations described in the SAP, unless otherwise specified. If a deviation is serious enough to have a

potential impact on the primary analysis, sensitivity analyses may be performed. A list of all protocol deviations is reviewed and decisions regarding how to handle these deviations are documented by the study team physician, clinical pharmacology scientist and statistician prior to database lock.

3.3.6 Analysis of data from China, Japan and East Asia cohorts

To support ethnic sensitivity evaluation between China and the rest of world, East Asia and the rest of world, and Japan and the rest of the world evaluation in efficacy, safety and PK in China, East Asia and Japan population is planned in this study for exploratory purpose. These analyses will be reported separately from the CSR.

3.3.6.1 Definition of China cohort, Japan cohort and East Asia cohort

China cohort is defined as the participants enrolled in sites located in mainland China, in Hong Kong and Taiwan.

The East Asia cohort is defined as the participants enrolled at sites located in East Asian countries/regions. The list of countries/regions includes: mainland China, Hong Kong, Taiwan, South Korea and Japan.

Japan cohort is defined as the participants enrolled in sites located in Japan.

3.3.6.2 Analysis sets

The definition of the analysis sets will be the same as the global cohort as defined in Section 3.2 with the exception of including only participants from China cohort, East Asia cohort and Japan cohort, respectively.

3.3.6.3 Analysis considerations

The analysis for China cohort, East Asia cohort and Japan will be performed at the same time as the overall population.

All efficacy and safety variables will be derived in the same way as for the analyses of the global cohort. Analyses detailed in the section 4 of the study SAP will be repeated for the China cohort, East Asia cohort and Japan cohort using the analysis sets described in this section. The China cohort, East Asia cohort and Japan cohort are not powered subpopulations, so all statistical analyses for these cohorts will be considered exploratory. No subgroup analysis or sensitivity analysis will be performed for China cohort, East Asia cohort and Japan cohort. No separate listings of individual participants in the China cohort, East Asia cohort or Japan cohort will be produced.

3.4 Alternative approach for main and sensitivity analyses of all binary endpoints

In cases where the primary stratified analysis using the Cochran-Mantel-Haenszel (CMH) test is not feasible due to a low number of responders in one or more treatment groups-resulting in sparse or zero cell counts within strata-an unstratified analysis will be conducted as an alternative for all primary analyses and binary endpoints.

Specifically, the association between treatment and binary response will be evaluated using Fisher's Exact Test, without adjustment for stratification factors. Estimates of the odds ratio (OR) with exact 95% confidence intervals and difference in proportions with Miettinen-Nurminen (score) 95% confidence intervals will be provided, calculated from the 2×2 contingency table of treatment by response.

Results from this analysis will be reported, including the exact two-sided p-value from Fisher's Exact Test, odds ratios with exact confidence intervals, and differences in proportions along with their corresponding Miettinen-Nurminen confidence intervals.

No analyses will be conducted if there are no responders in either treatment group.

Sensitivity analyses mirroring this alternative approach for primary and key secondary endpoints will not be pre-specified in this SAP but may be conducted as ad hoc analyses, if deemed necessary, and documented separately.

4 STATISTICAL ANALYSIS

This section provides information on definitions, derivation and analysis/data presentation per domain.

Study results will be presented for AZD2693 25 mg, AZD2693 50 mg, AZD2693 Total, pooled Placebo (when applicable) and Total (when applicable).

Selected listings will be presented for Heterozygous subjects who were randomised before the CSP amendment.

4.1 Study Population

The domain study population covers subject disposition, analysis sets, protocol deviations, demographics, baseline characteristics medical history, prior and concomitant medication and study drug compliance.

4.1.1 Subject Disposition and Completion Status

4.1.1.1 Definitions and Derivations

Screen failures are defined as participants who consent to participate in the clinical study but are not subsequently randomly assigned to study intervention. Two screenings are part of this study: genotyping screening (to assess if participants are homozygous carriers of the PNPLA3 rs738409 148M risk allele) and main study screening. A minimal set of screen failure information is required to ensure transparent reporting of screen failure participants to meet the Consolidated Standards of Reporting Trials (CONSORT) publishing requirements and to respond to queries from regulatory authorities. Minimal information includes demography, screen failure details, eligibility criteria, medical history, and any SAE.

Subjects randomised will be defined as those who signed informed consent, were genotyped as homozygous for the risk allele and have been randomly allocated to one of the treatment arms.

Participants who started treatment are those who reported at least one administration of IP.
Participants who completed treatment are those who received 52 weeks of IP.
Participants who discontinued treatment are those who did not receive 52 weeks of IP administration due to a permanent discontinuation of IP administration.

Participants who completed the study are those who completed 52 weeks of treatment and 12 additional weeks of follow-up.

4.1.1.2 Presentation

Participant disposition will be summarised overall and by treatment arm (as number and percentage). Separate summary will be created for East Asia, China and Japan cohorts. The following milestones will be summarised:

- # participants screened
- # participants not screened at Visit 2 (incl. reasons)
- # participants screened at Visit 2
- # participants screened at Visit 2 not randomised (incl. reasons)
- # participants randomised
- # participants randomised, not treated (incl. reasons)
- # participants who started treatment
- # participants who completed the treatment
- # participants who discontinued treatment (incl. reasons)
- # participants who interrupted treatment
- # participants ongoing on treatment at data cut-off date
- # participants ongoing in study at data cut-off date
- # participants assessed for biopsy at Week 54
- # participants who completed the study
- # participants withdrawn from study (incl. reasons)

The table will be based on all enrolled participants. A listing including all standardized disposition terms will be also provided for all enrolled participants.

4.1.2 Analysis Sets

4.1.2.1 Definitions and Derivations

Analysis sets are defined in 3.2.

4.1.2.2 Presentation

The number and percentage of randomly assigned to study intervention homozygous participants belonging to each analysis set (excluding enrolled set) and excluded for each possible reason will be presented in a summary table. Any exclusions from analysis sets will be listed. Separate summary will be created for East Asia and China cohorts.

4.1.3 Protocol Deviations

4.1.3.1 Definitions and Derivations

Protocol deviations are described in 3.3.5. Deviations from the protocol will be assessed as important or non-important by blinded members of the study team before database hard lock and unblinding. A blind data review meeting will be held.

4.1.3.2 Presentation

The incidence of important protocol deviations (IPDs) is summarised for the Full analysis set by deviation categories. The number and percentage of subjects in the following categories are summarised:

- Number of subjects with at least 1 important protocol deviation.
- Number of subjects with at least 1 important protocol deviation related to global/country situation.
- Number of subjects with at least 1 important protocol deviation excluding global/country situation related IPDs.

A listing will be provided with the important protocol deviation details.

4.1.4 Demographics

4.1.4.1 Definitions and Derivations

The following demographics variables are collected: country, age, gender, race, and ethnicity. Age (as reported in the eCRF) will be grouped accordingly in the following categories: < 50, ≥ 50 to < 65, ≥ 65.

4.1.4.2 Presentation

Demographics will be listed and summarised for participants in the FAS. Countries will be clustered into the following regions: East Asia, South Asia, Europe, North America, LATAM (Central and South America), and Middle East. Separate summary will be created for East Asia, China and Japan cohorts.

4.1.5 Baseline Characteristics

4.1.5.1 Definitions and Derivations

Height is collected at screening. Weight and waist circumference are collected repeatedly according to the SoA. BMI (as reported in the eCRF) will be categorised as: normal weight (BMI < 25), overweight (BMI ≥ 25 and BMI < 30) and obese (BMI ≥ 30).

4.1.5.2 Presentation

Baseline variables will be summarized for each treatment group and overall, for the FAS. Continuous variables will be summarized using n, mean, median, SD, minimum, and maximum and categorical variables will be summarized using number and percent of participants. Separate summary will be created for East Asia, China and Japan cohorts.

4.1.6 Disease Characteristics at Baseline

4.1.6.1 Definitions and Derivations

Disease characteristics consists of the following variables: fibrosis stage and presence of T2DM at randomisation, NAS score (overall and its specific components: steatosis, lobular inflammation, and ballooning), ALT, AST, GGT, CK-18 (M30 and M65), APRI, NIS4, NFS, ELF score, FIB-4, HbA1c and lipid parameters (total cholesterol, LDLc, HDLc, triglycerides, FFA, ApoB). Separate summary will be created for East Asia, China and Japan cohorts.

4.1.6.2 Presentation

Disease characteristics will be summarized at baseline for each treatment group and overall, for the FAS. Continuous variables will be summarized using n, mean, median, SD, minimum, and maximum and categorical variables will be summarized using number and percent of participants.

4.1.7 Medical History and Concomitant Disease

4.1.7.1 Definitions and Derivations

Medical history is coded in Medical Dictionary for Regulatory Activities (MedDRA) version 25.0 or higher.

4.1.7.2 Presentation

Relevant medical history will be presented in summary tables as number and percentages of subjects by System Organ Class (SOC) and Preferred Term (PT) for each treatment group and overall. Subjects with multiple events in the same SOC/PT will be counted only once in that SOC/PT. Subjects with events in more than one SOC/PT will be counted once in each of those SOC/PT. The table will be sorted alphabetically by SOC and PT and will be based on the FAS. Separate summary will be created for East Asia and China cohorts.

4.1.8 Prior and Concomitant Medications

4.1.8.1 Definitions and Derivations

Prior medications, concomitant and post-treatment medications are defined based on imputed start and stop dates as follows:

- Prior medications are those taken prior to study treatment with a stop date prior to the first dose of study treatment.
- Concomitant medication is a medication meeting any of the conditions:
  - Medication with start date prior to treatment with IP and with stop date on or after the first IP dose, missing stop date or medication which is ongoing.
  - Medication which started during treatment.
- Post-treatment medications are those with a start date after the last dose date of study treatment.

4.1.8.2 Presentation

Summaries of all prior and concomitant medications (by ATC classification) will be summarised for FAS, for each dose group, pooled placebo and total population.

4.1.9 Study Drug Compliance

4.1.9.1 Definitions and Derivations

Compliance to IP is defined by how many of the planned doses the participant has received. This will be expressed in percentage with the following formula

$$\text{Compliance (\%)} = \frac{\text{Number of injections taken}}{\text{Number of expected injections}} \times 100.$$

Number of expected injections is defined as the number of expected injections between first dose and time of completed planned treatment period (equals to 14 injections). Number of injections taken are defined as the number of injections taken according to the EX eCRF.

The data will also be categorised in three different groups, these are

- < 80%
- ≥ 80% - < 120%
- ≥ 120%.

If the information on whether the planned dose has been administered is missing, it will be counted as if the injection was not taken irrespective of the drug was dispensed or not.

4.1.9.2 Presentation

Compliance to IP will be summarized as a continuous and categorical variable. It will be presented for FAS. Separate summary will be created for East Asia and China cohorts.

4.2 Endpoint Analyses

List of primary, secondary and exploratory endpoints is provided in Table 7.

The primary estimand of interest for the binary histological endpoints is based on the hybrid of treatment policy strategy with the composite strategy.

Intercurrent event strategies:

Treatment policy strategy is applied to the following intercurrent events: treatment discontinuation, deviations from the protocol dosing schedule, changes in background medication, use of prohibited medications, other protocol deviations.

Composite strategy is applied to the following adjudicated events: hepatic decompensation, clinical cirrhosis, DILI, hepatocellular carcinoma, and death from any cause. These events will be referred to as liver outcomes. A liver outcome occurring indicates worsening in a participant's condition. Consequently, the participant would be considered to be a non-responder.

Participants will be listed for each intercurrent event mentioned below:

- Treatment discontinuation
- Deviations from protocol dosing schedule as explained in section "Investigational Product (IP) Deviation" of Protocol Deviation Plan
- Use of prohibited medications
- Other protocol deviations as explained in section "Other Important Protocol Deviations" of Protocol Deviation Plan
- Liver event
- Death

For all binary biopsy endpoints, if a liver outcome occurs, the participant will be classified as a non-responder regardless of the biopsy result and regardless of whether a biopsy was obtained or not. In case if a participant is a responder based on an early discontinuation visit and the participant did not remain in the study until Week 54, it will be assumed that no liver outcome would have occurred in the missing time interval. If biopsy is collected

later than the end of the biopsy window and there is no liver outcome until the end of that window, the outcome of this participant will be considered missing.

Table 7 List of all study objectives

Statistical category	Endpoint	Population	Intercurrent event strategy	Population level summary (analysis)	Details in section
Objective 1: To assess the effect of AZD2693 versus placebo on histological resolution of NASH in participants with non-cirrhotic NASH with fibrosis (and homozygous for the *PNPLA3* rs738409 148M risk allele)					
Primary	NASH resolution without worsening of fibrosis based on histology after Week 52	FAS	Hybrid strategy with missing data imputed as non-responder	Difference in proportions of participants with NASH resolution without worsening of fibrosis based histology after Week 52	4.2.1
Objective 2: To assess the effects of AZD2693 versus placebo on histological fibrosis improvement					
Secondary	At least one stage of liver fibrosis improvement with no worsening of NASH based on biopsy after Week 52	FAS	Hybrid strategy with missing data imputed as non-responder	Difference in proportions of participants with at least one stage of liver fibrosis improvement with no worsening of NASH based on biopsy after Week 52	4.2.2
Objective 3: To assess the effect of AZD2693 versus placebo on ≥ 2-point improvement in NAS					
Secondary	≥ 2-point improvement from baseline in NAS based on biopsy after Week 52	FAS	Hybrid strategy with missing data imputed as non-responders	Difference in proportions of participants with ≥ 2-point improvement from baseline in NAS based on biopsy at after Week 52	4.2.3
Objective 4: To assess the effect of AZD2693 versus placebo on improvement in fibrosis by at least one stage					
Secondary	Improvement in fibrosis by at least one stage based on biopsy after Week 52	FAS	Hybrid strategy with missing-data imputed as non-responder	Difference in proportions of participants with improvement in fibrosis by at least one stage based on biopsy after Week 52	4.2.4

Statistical category	Endpoint	Population	Intercurrent event strategy	Population level summary (analysis)	Details in section
Objective 5: To assess safety and tolerability of AZD2693 compared with placebo in participants with non-cirrhotic NASH with fibrosis					
Safety	Safety and tolerability will be evaluated in terms of AEs, vital signs, ECG, and clinical and laboratory evaluation including platelet count	Safety set	Treatment policy	Descriptive statistics	4.6
Objective 6: To assess the effect of AZD2693 versus placebo on resolution of NASH and improvement in fibrosis					
Exploratory	Resolution of NASH and improvement in fibrosis by at least one stage based on biopsy after Week 52	FAS	Hybrid strategy with missing data imputed as non-responders	Difference in proportions of participants with both resolution of NASH and improvement in fibrosis by at least one stage based on biopsy after Week 52	4.2.6
Objective 7: To assess the effects of AZD2693 versus placebo on NASH-related histological parameters					
Exploratory	Change from baseline to post-Week 52 in: • NAS (0-8) • Stage of fibrosis according to Kleiner fibrosis classification (0-4) • Activity component of SAF score (0-4)	FAS	Treatment policy without imputation of missing data	LSmean difference in change from baseline based on ANCOVA	4.2.7
Objective 8: To assess the effect of AZD2693 versus placebo on circulating markers of hepatic injury and apoptosis or necrosis					
Exploratory	Change from baseline to post-Week 52 in • Liver enzymes (ALT, AST, GGT) • CK-18 (M65 and M30)	FAS	Treatment policy without imputation of missing data	LSmean difference in change from baseline based on repeated measures model	4.2.8

Statistical category	Endpoint	Population	Intercurrent event strategy	Population level summary (analysis)	Details in section
Objective 9: To assess the effect of AZD2693 versus placebo on circulating biomarkers of fibrosis and fibrogenesis					
Exploratory	Change from baseline to post-Week 52 in: • Collagen turnover (e.g. PRO-C3) • ELF score • FIB-4	FAS	Treatment policy without imputation of missing data	LSmean difference in change from baseline based on repeated measures model	4.2.9
Objective 10: To assess the effect of AZD2693 versus placebo on lipids parameters					
Exploratory	Change from baseline to post-Week 52 in: • Total cholesterol • LDLc • HDLc • Triglycerides • FFA • ApoB	FAS	Treatment policy without imputation of missing data	LSmean difference in change from baseline based on repeated measures model	4.2.10
Objective 11: To explore the effect of AZD2693 on different non-invasive algorithms developed as marker of NAFLD or NASH disease activity					
Exploratory	Change from baseline to post-Week 52 in: • NIS4 • APRI • NFS	FAS	Treatment policy without imputation of missing data	LSmean difference in change from baseline based on repeated measures model	4.2.11
Objective 12: To explore the effect of AZD2693 on biomarkers for NASH comorbidities					
Exploratory	Change from baseline to post-Week 52 in: Diabetes/metabolic syndrome: • HbA1c • Insulin • C-Peptide • Plasma glucose • HOMA-IR Renal: • eGFR	FAS	Treatment policy without imputation of missing data	LSmean difference in change from baseline based on repeated measures model	4.2.12
Objective 13: To evaluate pharmacokinetics of AZD2693					

Statistical category	Endpoint	Population	Intercurrent event strategy	Population level summary (analysis)	Details in section
Exploratory	1) Observed plasma concentrations of AZD2693 full-length ASOs 2) Population PK parameters to be reported separately	PK	Treatment policy without imputation of missing data	Summary statistics	4.4
Objective 14: To evaluate the immunogenicity of AZD2693					
Exploratory	Incidents of development of ADA and ADA titre (if participants are ADA-positive) during treatment and follow-up	ADA set	Treatment policy without imputation of missing data	Summary statistics	4.5
Objective 15: To evaluate effect of AZD2693 on patient-reported outcomes versus placebo					
Exploratory	Changes from baseline to post-Week 52 in • NASH-CHECK • SF-36 • PGIS	FAS	Treatment policy without imputation of missing data	• LSmean difference in change from baseline based on repeated measures model (for SF-36 and NASH-CHECK) • Summary statistics for PGIS	4.2.15
Objective 16: Not mandatory at all sites: To explore liver stiffness and liver steatosis using imaging analysis					
Exploratory	Change from baseline to Week 26 and Week 54 in HFF, liver fat volume, volume corrected HFF, liver stiffness	FAS	Treatment policy without imputation of missing data	LSmean difference in change from baseline based on repeated measures model	4.2.16

4.2.1 Primary Endpoint: NASH resolution without worsening of fibrosis based on histology after Week 52

4.2.1.1 Definition

Liver biopsy samples taken at baseline (either historical [collected within 6 months before randomisation] or taken during the screening period) and after Week 52 will be assessed centrally by a Central Pathology Review Committee to determine the NAS and fibrosis staging.

The resolution of NASH is defined as a ballooning score of 0, inflammation score of 0 to 1, and steatosis score of any degree (from 0 to 3), as assessed by NASH Clinical Research Network (CRN)/NASH Activity Score (NAS). Worsening of fibrosis is defined as an increase in the NASH CRN fibrosis score. Because resolution of NASH allows steatosis score of any degree (from 0 to 3), if the steatosis score is not evaluable, resolution of NASH will still be concluded if there is a ballooning score of 0 and an inflammation score of 0 to 1. If a ballooning score is not 0, or inflammation score is not 0 or 1 then such a case will be concluded as no NASH resolution.

4.2.1.2 Derivations

The resolution of NASH is defined as:

- Ballooning score of 0
- Inflammation score of 0 or 1
- Steatosis score of any degree (0 to 3)

The NIDDK NASH CRN system of grading disease activity and staging of fibrosis separately is the most widely used and validated system for histological assessment of NASH in clinical trials.

Description of NASH CRN scoring and fibrosis staging is provided in Table 8 and Table 9. Worsening of fibrosis is defined by an increase of at least one stage.

Table 8 Description of non-alcoholic fatty liver disease activity score

Steatosis		**S score**		**Lobular inflammation**	**L score**		**Hepatocyte ballooning**	**B score**
< 5%		0		None	0		None	0
5-33%		1		< 2	1		Few ballooned cells	1
34-66%		2		2-4	2		Many ballooned cells	2
> 66%		3		> 4	3			

Table 9 Description of fibrosis stage

NASH fibrosis stage	Stage	Score
None	0	0
Mild, zone 3 perisinusoidal fibrosis	1a	1
Moderate, zone 3 perisinusoidal fibrosis	1b	1
Portal/periportal fibrosis only	1c	1
Zone 3 perisinusoidal and portal/periportal fibrosis	2	2
Bridging fibrosis	3	3
Cirrhosis	4	4

4.2.1.3 Handling of Dropouts and Missing Data

- Post-baseline biopsy is expected for each participant. It should take place in Week 54.

- Participants who prematurely and permanently discontinue the study intervention will continue their study participation by following the original visit schedule. Therefore, those participants will be encouraged to have a biopsy at Week 54 (regardless of treatment discontinuation).

- Participants who withdraw from the study (expected to be uncommon) will have an early discontinuation visit. Participants who completed at least 36 weeks of the intervention at the time of choosing to withdraw are eligible to complete an exit biopsy 2 weeks (± 7 days) after the last dose received and should be counselled on this option.

- Participants who have not received at least 36 weeks of treatment or underwent biopsy outside of the visit window will have missing primary endpoint (if they did not experience a liver outcome)

In case ballooning score or inflammation score is missing and NASH/no NASH resolution cannot be confirmed based on the rest of the scores then response will be missing.

For the primary analysis participants with missing biopsy result will be considered non-responders.

In the sensitivity analyses multiple imputation or complete-cases analysis will be performed.

4.2.1.4 Primary Analysis of Primary Endpoint

Primary endpoint will be analysed for FAS in homozygous population. The summary measure for the primary endpoint will be proportion of responders. The treatment effects (of each dose compared to pooled placebo) will be summarised by difference in proportions and odds ratio. A 95% confidence interval will be computed for the treatment effects. The p-value will come from a Cochran-Mantel-Haenszel test with presence of T2DM and F2/F3 fibrosis stages as stratification factors. The method of Miettinen-Nurminen will be used to compute 95% confidence intervals for difference in proportions with stratification by T2DM and stage of fibrosis. A Breslow-Day test will be performed to report a p-value for homogeneity of ORs. Separate summary will be created for East Asia, China and Japan cohorts.

4.2.1.5 Sensitivity Analyses of the Primary Endpoint

Sensitivity 1: Multiple imputation

Multiple imputation will be performed using the FAS population for the primary endpoint. Observed histological responses in the placebo arm will be used to impute missing data in both active and placebo groups. Firstly, the probability of the histological endpoint occurring in the placebo arm for participants with the observed endpoint will be modelled using logistic regression with stratification variables (baseline fibrosis stage and T2DM status) as independent variables, obtaining parameter estimates and covariance matrix, β and $Var(\beta)$.

If a logistic regression model fails to converge, variables will be removed in the following order until a convergent model is obtained: T2DM status, then fibrosis stage. For each analysis, N=1000 imputations will be used. For each imputation, a vector of parameters will be simulated from the posterior distribution of the logistic regression, $N(\beta, Var(\beta))$, and used to calculate individual predicted probabilities for each participant. Using the predicted probability, a Bernoulli draw will be performed to impute the binary endpoint for participants with missing data. Imputation will be performed using a MONOTONE LOGISTIC statement with PROC MI in SAS with specification of a random seed 1305417.

A common odds ratio (OR) stratified by baseline fibrosis stage and T2DM status, and its associated standard error (within imputation variance) will be calculated from each

completed dataset. The OR will be estimated using the Cochran-Mantel-Haenszel (CMH) method via SAS's PROC FREQ procedure. By default, PROC FREQ computes the Mantel-Haenszel adjusted odds ratio and its corresponding confidence limits using a log transformation and returns the back-transformed estimates.

Given that the OR follow a log-normal distribution, a logarithmic transformation will be applied before pooling the estimates and the asymptotic standard error of the log(OR) will be computed following the method of Robins, Breslow, and Greenland. Rubin's rules (Rubin 1987) will be used to pool the analysis results of the imputed OR. Pooled p-values will be obtained using the Wilson-Hilferty method through PROC MIANALYZE. Specifically, the CMH chi-square statistic will be transformed using the Wilson-Hilferty approximation to yield a variable that is approximately normally distributed with mean 0 and variance 1, allowing valid pooling of p-values across imputed datasets.

To obtain an estimate and 95% confidence interval (CI) for the common difference in proportions, the CMH risk difference (RD) estimate and associated standard error (using the Sato estimator [(Sato, 1989)]) will be calculated for each completed dataset and combined using Rubin's rules.

The p-values, common odds ratio (OR) estimates, and their 95% confidence intervals will be computed by inputting the log(OR) and associated standard error into PROC MIANALYZE. The standard error (SE) of the common odds ratio will be derived from the 95% confidence interval provided by PROC FREQ, which estimates the common OR using the Mantel-Haenszel method on the logarithmic scale. Specifically, under the assumption that log(OR) is approximately normally distributed, the confidence interval is computed as:

$$\log(OR) \pm z_{1-\alpha/2} \times SE$$

Based on this, the standard error can be calculated as:

$$SE = \log(UCL/OR) / z_{1-\alpha/2}$$

The pooled estimate of OR from PROC MIANALYZE will be back-transformed by exponentiation. The pooled estimate of the standard error from PROC MIANALYZE will be back-transformed to its original scale using the Delta Method. The confidence intervals will be then computed as

$$OR_{back\text{-}transformed} \pm z_{1-\alpha/2} \times SE_{back\text{-}transformed}$$

The estimate of common difference in proportions and associated 95% confidence interval will be computed by inputting the estimates of RD and standard error (Sato) into PROC

MIANALYZE. The standard error will be derived from the upper confidence limit (UCL) of the 95% confidence interval of the common RD as:

$$SE = (UCL - RD) / z_{1-\alpha/2}$$

Sensitivity 2: Tipping point

A tipping point analysis will be performed to identify the assumptions about response rates among missing data that would alter the conclusion of the primary analysis comparing AZD2693 50 mg to Placebo. For this, 50 imputed datasets will be generated. Missing values will be imputed by drawing from a Bernoulli distribution with response probabilities p_0^i and p_1^i for the Placebo and AZD2693 50 mg groups, respectively, where i=1,…,50. The probabilities p_0^i and p_1^i will range from 0.05 to 0.95 to reflect a range of plausible response scenarios by treatment group.

For each imputation i, the number of responders among participants with missing data in each group will be drawn from these Bernoulli distributions, resulting in paired counts (n_1^i, n_0^i), for AZD2693 50 mg and Placebo, respectively. Imputations will be performed using a random seed of 1305417.

After completing imputations for each scenario, Rubin's rules (Rubin, 1987) will be applied to pool the estimated odds ratios (ORs) across the 50 datasets. The OR will be estimated using the Cochran-Mantel-Haenszel (CMH) method via the PROC FREQ procedure, which computes the estimates on the logarithmic scale and subsequently back-transforms them to the original scale. Pooled p-values will be obtained using the Wilson-Hilferty method through PROC MIANALYZE. Specifically, the CMH chi-square statistic will be transformed using the Wilson-Hilferty approximation to yield a variable that is approximately normally distributed with mean 0 and variance 1, allowing valid pooling of p-values across imputed datasets.

This multiple imputation approach using Bernoulli draws results in mean response rates approximately equal to the specified probabilities p_0 and p_1. A tipping point is defined as the combination of p_0 and p_1 values at which the superiority of AZD2693 50 mg over Placebo loses statistical significance.

Sensitivity 3: Complete cases

Complete-cases analysis where only participants with baseline biopsy and biopsy after Week 52 contribute to the analysis.

4.2.1.6 Supplementary Analyses of the Primary Endpoint

In a supplementary analysis each dose of AZD2693 will be compared to the matching placebo. Hybrid estimand strategy with missing outcomes imputed as non-responders will be applied.

4.2.1.7 Subgroup Analyses

Subgroup analyses will be performed for stratification factors (T2DM and stage of fibrosis) as given at randomisation, as well as for age, race, sex and ethnicity. Treatment effect will be summarized as a difference in proportions together with 95% CIs computed using Miettinen-Nurminen method (without stratification). Difference between treatment (for each dose separately) and placebo will be tested using chi-square or Fisher's exact test. Forest plot will be drawn to display the results.

4.2.2 Key Secondary Endpoint: At least one stage of liver fibrosis improvement with no worsening of NASH based on biopsy after Week 52

4.2.2.1 Definition

Proportion of participants with at least one stage of liver fibrosis improvement with no worsening of NASH (yes/no) based on biopsy after Week 52. Worsening of NASH is defined as an increase of at least one stage of either lobular inflammation or hepatocyte ballooning or steatosis according to NASH CRN criteria.

4.2.2.2 Derivations

NASH and fibrosis scoring is provided in the derivation of the Primary Endpoint.

4.2.2.3 Handling of Dropouts and Missing Data

See 4.2.1.3.

4.2.2.4 Primary Analysis of Secondary Endpoint

See 4.2.1.4. Separate summary will be created for East Asia, China and Japan cohorts.

4.2.2.5 Sensitivity Analyses of the Secondary Endpoint

Sensitivity analyses will incorporate the multiple imputation and complete-cases strategies described in 4.2.1.5.

No tipping point analysis will be performed for this endpoint.

4.2.2.6 Supplementary Analyses of the Secondary Endpoint

The same supplementary analyses as for the primary endpoint (see Section 4.2.1.6) will be implemented.

4.2.2.7 Subgroup Analyses

None.

4.2.3 Secondary Endpoint: ≥ 2-point improvement from baseline in NAS based on biopsy after week 52

4.2.3.1 Definition

Proportion of participants with ≥2-point improvement from baseline in NAS after 52 weeks.

4.2.3.2 Derivations

NAS score is the sum of the separate scores for the steatosis (0 to 3), hepatocellular ballooning (0 to 2), and lobular inflammation (0 to 3). It is an integer score between 0 and 8. For the definition of the separate scores see 4.2.1.2.

4.2.3.3 Handling of Dropouts and Missing Data

See 4.2.1.3

4.2.3.4 Primary Analysis of the Secondary Endpoint

See 4.2.1.4. Separate summary will be created for East Asia, China and Japan cohorts.

4.2.3.5 Sensitivity Analyses of the Secondary Endpoint

See 4.2.2.5

4.2.3.6 Supplementary Analyses of the Secondary Endpoint

None.

4.2.3.7 Subgroup Analyses

None.

4.2.4 Secondary Endpoint: improvement in fibrosis by at least one stage based on biopsy after Week 52

4.2.4.1 Definition

Proportion of participants with improvement in fibrosis by at least one stage based on biopsy after Week 52.

4.2.4.2 Derivations

Fibrosis stage is based on the NASH CRN grading system. See 4.2.1.2.

4.2.4.3 Handling of Dropouts and Missing data

See 4.2.1.3.

4.2.4.4 Primary Analysis of the Secondary Endpoint

See 4.2.1.4. Separate summary will be created for East Asia, China and Japan cohorts.

4.2.4.5 Sensitivity Analyses of the Secondary Endpoint

See 4.2.2.5

4.2.4.6 Supplementary Analyses of the Secondary endpoint

None.

4.2.4.7 Subgroup Analyses

None.

4.2.5 Safety endpoint

Safety is described in Section 4.6.

4.2.6 Exploratory endpoint: resolution of NASH and improvement in fibrosis by at least one stage based on biopsy after Week 52

4.2.6.1 Definition

Proportion of participants with both resolution of NASH and improvement in fibrosis by at least one stage based on biopsy after Week 52. Definitions are NASH resolution and staging of fibrosis are based on the NASH CRN criteria.

4.2.6.2 Derivations

See 4.2.1.2

4.2.6.3 Handling of Dropouts and Missing data

See 4.2.1.3

4.2.6.4 Primary Analyses of the Exploratory Endpoint

See 4.2.1.4 for the analysis methods.

4.2.6.5 Additional Analyses of the Exploratory Endpoint

Multiple imputation and complete-case analysis, see 4.2.1.5

4.2.6.6 Subgroup Analyses

None.

4.2.7 Exploratory endpoint: change from baseline to post-Week 52 in NASH-related histological parameters

4.2.7.1 Definition

The following outcomes will be analysed:

- Change in NAS score
- Change in stage of fibrosis according to the Kleiner fibrosis classification
- Change in activity component score (SAF)

4.2.7.2 Derivations

Definition of the NAS score and the staging of fibrosis is provided in the NASH CRN (See 4.2.1.2). Stages 1a, 1b, 1c in Table 9 are considered as fibrosis score equal to 1. Activity score of the SAF is the sum of hepatocyte ballooning and lobular inflammation. In case hepatocyte ballooning or lobular inflammation is missing then SAF score will be missing.

4.2.7.3 Handling of Dropouts and Missing data

In the descriptive statistics missing values will be presented separately. In the ANCOVA models no imputation will be performed.

4.2.7.4 Primary Analyses of the Exploratory Endpoint

An ANCOVA model will be fitted to each change from baseline adjusting for baseline value of NASH-related histological parameter, for value of stratification factors (T2DM and fibrosis stage) and for treatment group. For change from baseline in fibrosis score: i) adjusting for the second stratification factor (F2/F3) will not be considered, ii) the baseline fibrosis score will still be included in the model. The following results will be reported:

- LSmean change from baseline at Week 54, for placebo and each AZD2693 dose, together with a 95% confidence interval
- LSmean difference in change from baseline at Week 54 for each AZD2693 dose compared to placebo, together with a 95% confidence interval and two-sided p-value

4.2.7.5 Additional Analyses of the Exploratory Endpoint

Shift tables from baseline to Week 54 of all the scores will be presented.

4.2.7.6 Subgroup Analyses

None.

4.2.8 Exploratory endpoint: change from baseline to post-Week 52 in circulating markers of hepatic injury and apoptosis or necrosis

4.2.8.1 Definition

The following biomarkers will be collected: ALT, AST, GGT, CK-18 (M65, M30).

Change from baseline is defined as post baseline value minus baseline value.

Percent change from baseline is defined as:

$$\text{percent change} = \frac{(\text{post baseline} - baseline)}{\text{baseline}} * 100$$

4.2.8.2 Derivations

No derivations are necessary.

4.2.8.3 Handling of Dropouts and Missing data

All collected measurements will be analysed. Missing data will not be imputed.

4.2.8.4 Primary analyses of the Exploratory Endpoint

Change from baseline and percent change from baseline to post Week 52 in biomarkers will be analysed using a mixed model for repeated measures (MMRM). The dependent variable will be the change from baseline. The endpoint of interest is the change from baseline to post Week 52. Analysis of change from baseline for some parameters and of percent change

from baseline for others will be performed using the repeated measures obtained at weeks 4, 12, 24, 52, 54 and 64, respectively. The model will include intercept and fixed effects for treatment group (placebo/AZD2693 dose), for baseline value of the dependent variable, for value of stratification factors T2DM (yes/no) and fibrosis stage (F2/F3), for week as categorical variable and for treatment by week interaction. No random effects will be specified.

SAS® Proc Mixed with restricted maximum likelihood estimation (REML) and an unstructured within-participant covariance structure will be used. If this model fails to converge other covariance structures will be used: autoregressive (AR) and compound symmetry (CS) covariance structure.
For the parameters where change from baseline will be analysed, the results will include least squares means estimates, their 95% confidence intervals, treatment differences in least squares means estimates, and their 95% confidence intervals for each post week 52 time point. The null hypothesis is that the mean difference in the exploratory endpoint between the AZD2693 dose and placebo group is zero, versus the alternative hypothesis that this difference is not zero. The hypotheses can be expressed as follows: H0: μActive - μControl = 0 versus H1: μActive - μControl ≠ 0

Where μActive refers to the mean change from baseline to week 54 in biomarker score in the AZD2693 dose group and μcontrol refers to the mean change from baseline to week 54 in biomarker score in the placebo group.

Testing will be performed using the MMRM model-based t-test with a two-sided significance level of 0.05.

For the parameters where percent change from the baseline will be analysed, difference between post baseline and baseline value in logarithmic scale will be used. Results will include the estimates of ratio of post week 52 geometric mean to baseline geometric mean together with the estimates' 95% confidence intervals for each treatment group and each post week 52 time point. In addition, estimates for the difference in post week 52 geometric means as a ratio with their 95% confidence intervals and p value for the comparison between AZD2693 dose and Placebo will be presented. Back transformation of the results is shown in Table 10 Formula for MMRM for log-normally parameters.

Both change from baseline and percent change from baseline will be analysed for ALT, AST, GGT and CK-18.

Table 10 Formula for MMRM for log-normally parameters

Quantity	Calculation
Estimate for ratio of post Week 52 GeoMean to baseline GeoMean	exp(LSMEAN change from baseline in natural logarithm)
Lower Confidence Limit (LCL) for ratio of post Week 52 GeoMean to baseline GeoMean	exp(LCL for LSMEAN change from baseline in natural logarithm)
Upper Confidence Limit (UCL) for ratio of post Week 52 GeoMean to baseline GeoMean	exp(UCL for LSMEAN change from baseline in natural logarithm)
Estimate for Ratio Change from Baseline, AZD2693 relative to Placebo	exp(LSMD in mean change from baseline in natural logarithm)
LCL for Ratio Change from Baseline, AZD2693 relative to Placebo	exp(LCL for LSMD in change from baseline in natural logarithm)
UCL for Ratio Change from Baseline, AZD2693 relative to Placebo	exp(UCL for LSMD in change from baseline in natural logarithm)

4.2.8.5 Additional Analyses of the Exploratory Endpoint

None.

4.2.8.6 Subgroup Analyses

None.

4.2.9 Exploratory endpoint: change from baseline to post-Week 52 in circulating biomarkers of fibrosis and fibrogenesis

4.2.9.1 Definition

The following markers will be analysed: collagen turnover (Pro-C3), ELF score and FIB-4 score.

4.2.9.2 Derivations

No derivations are necessary for collagen turnover and ELF score. FIB-4 will be calculated according to the formula:

$$\frac{\text{Age (years)} * \text{ AST level (U/L)}}{\text{Platelet count } (10^9/\text{L}) * \sqrt{\text{ALT (U/L)}}}$$

In case any component is missing, the value of the biomarker will be set to missing.

Note that AST(IU/L)=AST(U/L) and AST(µkat/L) = 0.01667*AST(U/L). Similarly ALT(IU/L)=ALT(U/L) and ALT(µkat/L) = 0.01667*ALT(U/L).

4.2.9.3 Handling of Dropouts and Missing data

See 4.2.8.3

4.2.9.4 Primary Analyses of the Exploratory Endpoint

See 4.2.8.4. Change from baseline will be analysed for Pro-C3, ELF and FIB-4 scores.

4.2.9.5 Additional Analyses of the Exploratory Endpoint

None.

4.2.9.6 Subgroup Analyses

None.

4.2.10 Exploratory endpoint: change from baseline to post-Week 52 in lipids parameters

4.2.10.1 Definition

The following markers will be analysed: total cholesterol, LDLc, HDLc, triglycerides, FFA, ApoB.

4.2.10.2 Derivations

No derivations are necessary.

4.2.10.3 Handling of Dropouts and Missing data

See 4.2.8.3

4.2.10.4 Primary Analyses of the Exploratory Endpoint

See 4.2.8.4 Both change from baseline and percent change from baseline will be analysed.

4.2.10.5 Additional Analyses of the Exploratory Endpoint

None.

4.2.10.6 Subgroup Analyses

None.

4.2.11 Exploratory endpoint: change from baseline to post-Week 52 in different non-invasive algorithms developed as marker of NAFLD or NASH disease activity

4.2.11.1 Definition

The following markers will be analysed: NIS4, APRI, NFS.

4.2.11.2 Derivations

NIS4 does not require any derivations. APRI will be derived according to the formula:

$$\text{APRI} = \frac{\frac{\text{AST level}}{\text{AST (Upper Limit of Normal)}}}{\text{Platelet count } (10^9/\text{L})} * 100$$

NAFLD fibrosis score (NFS) will be calculated according to the formula:

NFS = -1.675 + (0.037 × age (years) + (0.094 × BMI (kg/m^2)) + (1.13 × IFG or T2DM (yes=1, no=0)) + (0.99 × AAR) - (0.013 × platelet count (10^9/L)) - (0.66 × albumin (g/dL)),

where AAR is the AST/ALT ratio, IFG is the impaired fasting plasma glucose and T2DM is the type 2 diabetes mellitus. If IFG is not recorded in medical history and T2DM is neither recorded in medical history nor in stratification factors at randomization, if baseline HBA1c is greater than 6.5% or if baseline fasting plasma glucose = 100 - 125 mg/dL then impaired fasting glucose/diabetes should be considered equal to Yes, otherwise equal to No. Serum Albumin is used.

Note that AST(IU/L)=AST(U/L) and AST(µkat/L) = 0.01667*AST(U/L). Similarly ALT(IU/L)=ALT(U/L) and ALT(µkat/L) = 0.01667*ALT(U/L). Albumin (g/L) = 1/10 *Albumin (g/dL). Plasma glucose (mg/dl)= (1/0.0555) * Plasma glucose (mmol/l)

In case any of the items is missing, the value of the biomarker will be set to missing.

4.2.11.3 Handling of Dropouts and Missing data

See 4.2.8.3

4.2.11.4 Primary Analyses of the Exploratory Endpoint

See 4.2.8.4 Change from baseline will be analysed for NFS. Both change and percent change from baseline will be analysed for NIS4, APRI.

4.2.11.5 Additional Analyses of the Exploratory Endpoint

None.

4.2.11.6 Subgroup Analyses

None.

4.2.12 Exploratory endpoint: change from baseline to post-Week 52 in biomarkers for NASH comorbidities

4.2.12.1 Definition

The following markers will be analysed: HbA1c, insulin, c-peptide, plasma glucose, HOMA-IR, eGFR.

4.2.12.2 Derivations

HOMA-IR will be calculated according to the formula:

$$\text{HOMA-IR} = \frac{Fasting\ glucose\ (mmol/L) * Fasting\ insulin\ (\mu U/mL)}{22.5}$$

using plasma glucose and conversion factor for insulin as below:
Insulin: pmol/L to (µU/mL), divide by (÷) 6.945

4.2.12.3 Handling of Dropouts and Missing data

See 4.2.8.3

4.2.12.4 Primary Analyses of the Exploratory Endpoint

See 4.2.8.4. Change from baseline will be analysed for insulin, c-peptide, plasma glucose and eGFR. Both change and percent change from baseline will be analysed for HbA1c and HOMA-IR.

4.2.12.5 Additional Analyses of the Exploratory Endpoint

None.

4.2.12.6 Subgroup Analyses

None.

4.2.13 Exploratory endpoint: pharmacokinetics of AZD2693

See section 4.4.

4.2.14 Exploratory endpoint: immunogenicity of AZD2693

See section 4.5.

4.2.15 Exploratory endpoint: change from baseline to post-Week 52 in patient-reported outcomes

4.2.15.1 Definition

Participants will be asked to digitally fill questionnaires at week 1 (before start of treatment), at Week 24 and at Week 52 (and at early discontinuation visit if such occurs). The following questionnaires will be used:

- SF-36v2
- NASH-CHECK
- PGIS

SF-36v2

The SF-36 health survey, version 2 (acute recall) (SF-36v2) is a 36-item self-report survey of functional health and well-being, with a 1-week recall period. Responses to 35 of the 36 items are used to compute an 8-domain profile of functional health and well-being scores. The remaining item, referred to as the 'Health Transition' item, asks participants to rate how their current state of health compared to their state of health 1 week ago, and is not used to calculate domain scores. The 8-domain profile consists of the following subscales: Physical Functioning, Role Limitations due to Physical Health, Bodily Pain, General Health Perceptions, Vitality, Social Functioning, Role Limitations due to Emotional Problems and Mental Health. Psychometrically-based Physical and Mental Component Summary scores are computed from domain scores to give a broader metric of physical and mental health-related quality of life.

NASH-CHECK

NASH-CHECK is a disease-specific PRO instrument assessing symptoms and health-related quality of life for participants with NASH. NASH-CHECK includes 28 items with a recall of the past 7 days. NASH-CHECK scores include 5 single-item symptom scores assessing Abdominal Pain, Abdominal Bloating, Fatigue, Sleep, and Itchy Skin; 1 symptom scale assessing Cognitive Symptoms (4 items); and 3 health-related quality of life scales assessing Activity Limitations (8 items), Emotional Impact (4 items) and Social Impact (7 items).

PGIS

The PGIS measures the participant's overall perceptions of their NASH related symptoms. The PGIS is a static, single-item measure: 'How would you describe the severity of the symptoms you have experienced with your fatty liver disease (known as NASH) over the past 7 days?' The PGIS has 4 response options: "no symptoms," "mild," "moderate-" and "severe."

4.2.15.2 Derivations

SF-36v2:

Scoring of SF-36v2 health domains and summary scores will be performed by QualityMetric, using norm based T-scores.

NASH-CHECK:

Text from User Manual Version 1.0: 9th February 2023 (Draft):

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 symptoms (scored 0) to worst possible or extreme symptoms (scored 10), with a higher score indicating more severe symptoms.

  The cognitive symptoms raw score is calculated as the mean of the nonmissing item scores for items 5 to 8 (i.e., the sum of items 5-8 divided by the number of nonmissing items). A minimum of 3 cognitive symptom items needs 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:

$$Cognitive\ symptoms\ raw\ score\ = \frac{sum\ of\ nonmissing\ scores\ for\ items\ 5\ to\ 8}{number\ of\ nonmissing\ items}$$

$$\boldsymbol{Cognitive\ Symptoms\ scale\ score}\ = cognitive\ symptoms\ raw\ score$$

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

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 nonmissing item scores for items 10 to 17 (i.e., the sum of items 10-17 divided by the number of nonmissing items). A minimum of 5 activity limitations items needs 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 from 0 to 10 to form the final scale score.
The formulas for the raw and scale scores are as follows:

$$Activity\ limitations\ raw\ score\ = \frac{sum\ of\ nonmissing\ scores\ for\ items\ 10\ to\ 17}{number\ of\ nonmissing\ items}$$

$$\boldsymbol{Activity\ Limitations\ scale\ score}\ = (\frac{\text{activity limitations raw score}}{4}) * 10$$

Activity Limitations scale scores range from 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 nonmissing item scores for items 18 to 21 (i.e., the sum of items 18-21 divided by the number of nonmissing 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:

$$Emotional\ Impact\ raw\ score\ = \frac{sum\ of\ nonmissing\ scores\ for\ items\ 18\ to\ 21}{number\ of\ nonmissing\ items}$$

$$\boldsymbol{Emotional\ Impact\ scale\ score} = (\frac{\text{emotional impact raw score}}{3}) * 10$$

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

- *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 nonmissing item scores for items 22 to 28 (i.e., the sum of items 22-28 divided by the number of nonmissing items). A minimum of 5 social impact items needs 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:

$$Social\ Impact\ raw\ score\ = \frac{sum\ of\ nonmissing\ scores\ for\ items\ 22\ to\ 28}{number\ of\ nonmissing\ items}$$

$$Social\ Impact\ scale\ score\ = (\frac{social\ impact\ raw\ score}{3}) * 10$$

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

PGIS:

No derivations are necessary for this questionnaire.

4.2.15.3 Handling of Dropouts and Missing data

Missing questions will not be imputed to derive scores.
In the analysis of change from baseline, missing outcomes will not be imputed.

4.2.15.4 Primary Analyses of the Exploratory Endpoint

SF-36v2 and NASH-CHECK:

Change from baseline to Week-52 in domain scores from NASH-CHECK, and domain and component summary scores from SF-36v2 will be analysed using a repeated measures model. The outcome will be difference from baseline in the PRO score. The following fixed effects will be entered in the model: intercept, treatment (placebo/AZD2693 dose), baseline PRO score, T2DM (yes/no), fibrosis stage (F2/F3, week, week*treatment interaction. Week will be entered as a categorical variable. No random effects will be specified. Errors will be assumed to be correlated within a participant with an unstructured covariance matrix. If the model will not converge other covariance structures will be used: autoregressive, compound symmetry.

The following results will be reported:

- LSmean change from baseline at Week 52, for placebo and each AZD2693 dose, together with standard errors
- LSmean difference in change from baseline at Week 52 for each AZD2693 dose compared to placebo, together with a 95% confidence interval and two-sided p-value

PGIS:

PGIS data will be presented using summaries and descriptive statistics. The proportion of participants in each response category at each visit will be presented by treatment arm. Change from baseline in the severity of NASH symptoms will be presented as the proportion of participants with no change, 1 category improvement, ≥ 2 category improvement, 1 category deterioration, and ≥ 2 category deterioration at each visit by treatment arm.

4.2.15.5 Additional Analyses of the Exploratory Endpoint

None.

4.2.15.6 Subgroup Analyses

None.

4.2.16 Exploratory endpoint: liver stiffness and liver steatosis using imaging analysis

4.2.16.1 Definition

The following markers will be analysed: HFF (hepatic fat fraction), liver fat volume, volume corrected HFF, liver stiffness.

4.2.16.2 Derivations

No derivations are necessary.

4.2.16.3 Handling of dropouts and Missing data

No imputations will be done.

4.2.16.4 Primary Analyses of the Exploratory endpoint

Percent change from baseline of liver fat volume and liver stiffness will be analysed using a model for repeated measures (see 4.2.8.4). Change from baseline will be analysed for hepatic fat fraction and for volume corrected liver fat fraction. Results will be provided for percent change from baseline to Week 26 and Week 54.

4.2.16.5 Additional Analyses of the Exploratory Endpoint

None.

4.2.16.6 Subgroup Analyses

None.

4.3 Pharmacodynamic Endpoint(s) (if not already covered as endpoint variables)

This section covers details related to pharmacodynamic endpoints and analyses.

4.3.1 Analysis

4.3.2 Definitions and Derivations

4.3.3 Presentation

4.4 **Pharmacokinetics** (if not already covered as endpoint variables)

4.4.1 Pharmacokinetic concentration data

Plasma concentrations for each scheduled time-point are summarized by treatment dose using appropriate descriptive statistics. The following descriptive statistics are presented for plasma concentrations:

- n
- n below LLOQ
- geometric mean (gmean)
- geometric standard deviation (gSD)
- geometric coefficient of variance (%) (gCV)
- arithmetic mean of non log-transformed data (mean)
- standard deviation of non log-transformed data (Std Dev)
- median
- minimum (min)
- maximum (max)

- The gmean is calculated as $\exp(\mu)$, where μ is the mean of the data on the natural log scale.
- The gSD is calculated as $\exp(\sigma)$, where σ is the standard deviation of the data on the natural log scale.
- The gCV is calculated as $100*\text{sqrt}[\exp(\sigma^2)-1]$.

Separate summary will be created for East Asia, China and Japan cohorts.

Handling of Non-Quantifiable Concentrations

Individual concentrations below the LLOQ of the bioanalytical assay are reported as NQ in the listings with the LLOQ defined in the footnotes of the relevant TFLs. Individual plasma concentrations that are Not Reportable are reported as NR and those that are missing are reported as NS (No Sample) in the listings. Plasma concentrations that are NQ, NR or NS are handled as follows for the provision of descriptive statistics:

- Any values reported as NR or NS are excluded from the summary tables and corresponding figures.

- At a time point where less than or equal to 50% of the concentration values are NQ, all NQ values are set to the LLOQ, and all descriptive statistics are calculated accordingly.

- At a time point where more than 50% (but not all) of the values are NQ, the gmean, mean, SD, gSD and gCV% are set to NC. The maximum value is reported from the individual data, and the minimum and median are set to NQ.

- If all concentrations are NQ at a time point, no descriptive statistics are calculated for that time point. The mean, gmean, minimum, median and maximum are reported as NQ and the SD, gSD and gCV% as NC.

- Three observations > LLOQ are required as a minimum for a serum concentration to be summarized. Two observations > LLOQ are presented as minimum and maximum with the other summary statistics as NC. One observation > LLOQ is presented as maximum with the other summary statistics as NC.

- The number of values below LLOQ (n < LLOQ) are reported for each time point together with the total number of collected values (n).

- Non-numeric result above ULOQ is reported as >ULOQ in the listing and excluded from summaries as no result.

4.4.2 Pharmacokinetic parameters data

AZD2693 PK in plasma will be summarized in a table by descriptive statistics described above. PK concentrations will be listed.

If data permit, a population PK model may be developed, possibly with the support of AZD2693 PK data from other studies, using nonlinear mixed effects modelling in NONMEM. The population PK model may be coupled with separate PD models. All PK/PD modelling will be described in a separate data analysis plan and the results of any such modelling will be provided in a separate population PK/PD report.

4.5 **Immunogenicity** (if applicable and if not already covered as endpoint variables)

Immunogenicity assessments will be conducted utilizing a tiered approach (screen, confirm, titer), and anti-drug antibody (ADA) data will be collected at scheduled visits shown in the CSP. The ADA result from each sample will be reported as either positive or negative. The following ADA definitions apply:

- ADA positive - A positive sample at any time during the study, including baseline and/or post-baseline

- ADA negative - Negative samples at all timepoints, including baseline and post-baseline

- TE-ADA positive - A positive post-baseline result and either of the following statements hold: treatment induced ADA positive or treatment boosted ADA positive

- Treatment-induced ADA positive – baseline is ADA negative and at least one post-baseline is ADA positive

- Treatment-boosted ADA positive - ADA positive at baseline and the baseline titre is boosted by greater than the variability of the assay (i.e. >=4 fold) at >=1 post-baseline timepoint

- Non-TE ADA positive - ADA positive but not fulfilling the definition of TE-ADA positive

- ADA prevalence - The proportion of subjects in a population who are ADA positive

- ADA incidence - The proportion of subjects in a population who are TE-ADA positive

- Only baseline positive - ADA positive at baseline and all post-baseline assessments are either negative or missing (or a mixture of negative and missing)

- ADA persistently positive - ADA negative at baseline and either of the following statements hold: ADA positive at >= 2 post-baseline assessments (with >= 16 weeks between first and last positive) or ADA positive at the last post-baseline assessment

- ADA transiently positive - ADA negative at baseline, having at least one post-baseline ADA positive assessment and not fulfilling the conditions of ADA persistently positive

- Both baseline and post-baseline positive – ADA positive at baseline and at least one post-baseline assessment is positive

ADA positive/negative status at the sample level:

- A sample is considered to be ADA positive if the titre is greater than or equal to the minimum required dilution (MRD) of the assay.

- A sample is considered to be ADA negative if the titre is <MRD of the assay.

Participants with missing baseline will be imputed per section 3.3.1. ADA results from samples collected post-dose instead of pre-dose on an investigational product administration day are considered unreliable and should be excluded from all derivations. The percentage of participants in each ADA-positive category will be summarised for ADA set (including separate summary for East Asia, China and Japan cohorts). For participants with positive ADA summary statistics of ADA titres will be provided. Plasma concentrations of AZD2693 over time by ADA categories will be summarised with descriptive statistics. Summary of adverse events per ADA category will be provided. Both these summaries will be separately created for East Asia and China cohorts. Additionally, the following Figures will be presented:

- A figure of proportion of subjects with positive ADA response over time with one line for each treatment group, and percentage of subjects with positive ADA response on the y-axis and time from first dose on the x-axis. At a given time t, the curve will show the percentage of subjects with at least one positive ADA sample up until time t.

- For subjects with positive ADA response, ADA titres on log scale over time will be presented in a Spaghetti plot.

- Boxplots for AZD2693 full length ASO concentrations for ADA positive vs ADA negative response per the specific visit will also be presented.

- Titres of AZD2693 by ADA category, timepoint and treatment group.

The above figures will be separately created for East Asia and China cohorts.

4.6 Safety Analyses (if not already covered as endpoint variables)

The domain safety covers exposure, adverse events, clinical laboratory, vital signs, and ECG.

Tables are provided for the safety set, listings are provided for all participants or the safety set, depending on the availability of data.

4.6.1 Exposure

4.6.1.1 Definitions and Derivations

Exposure duration is calculated only for participants in the Safety population as the total number of days on study drug and is calculated as:

Exposure duration (days) = Last date of IP - first day of IP + 30.

If any of the first or last dates are missing or partially missing, then imputed dates will not be used, and study drug exposure is set to missing.

Cumulative exposure will be computed, using the following duration (days) categories (corresponding to the visits with IP administration and exposure definition):

- ≥ 30 days
- ≥ 51 days
- ≥ 79 days
- ≥ 107 days
- ≥ 135 days
- ≥ 163 days
- ≥ 191 days
- ≥ 219 days
- ≥ 247 days
- ≥ 275 days
- ≥ 303 days
- ≥ 331 days
- ≥ 359 days
- ≥ 387 days

The total number of doses received will be calculated as number of days where at least one dose amount of > 0 is administered.

4.6.1.2 Presentation

The duration of exposure will be summarised using descriptive statistics for each study intervention group and overall.
Duration of exposure will also be presented categorically by cumulative categories as described above. These categories are cumulative, and participants will be included in all categories that apply to them.

The total time of exposure will be displayed for the safety population. Separate summary will be created for East Asia, China and Japan cohorts.

4.6.2 Adverse Events

4.6.2.1 Definitions and Derivations

Adverse events will be collected starting from the time of randomisation throughout the treatment period and during the follow-up period. SAEs will be recorded from the time of signing of the ICF.

The Medical Dictionary for Regulatory Activities (MedDRA) (latest version) is used to code the AEs. AEs are graded according to their severity (Mild, Moderate, Severe).

For rules on missing or partial dates, see 3.3.1. AEs with a missing start time (or where time is not collected) which occur on the same day as first IP administration are reported as on-treatment/on-study.

Adverse events of special interest are:

- Platelet count reduction
- Liver toxicity (transaminase elevations)
- Pro-inflammatory events
- Injection site reactions

4.6.2.2 Presentation

Analyses will be performed for safety analysis set and for on-treatment and on-study periods. AEs will be summarized for both on-treatment and on-study periods, except for summaries of key subject information, of AEs leading to discontinuation of IP and of AESI which will be summarised only for the on-study period.
AEs are counted once for each participant for calculating percentages of participants experiencing AE. In addition, if the same AE occurs multiple times within a particular participant, the highest severity and level of relationship observed are reported. For tables by MedDRA system organ class (SOC) and MedDRA preferred term (PT), participants with multiple AEs are counted once for each SOC/PT.

The overall table of AEs will summarise:

- # participants with any AE
- # participants with any AE considered related to IP
- # participants with any AE leading to death

- # participants with any SAE
- # participants with any AE leading to discontinuation of IP
- #participants with any AE leading to interruption of IP
- #participants with any AE leading to withdrawal of participant from the study

A separate summary for the overall table of AEs will be created for East Asia, China and Japan cohorts.

The following summary tables will be presented by SOC and PT:

- All AEs
- All AEs sorted in decreasing frequency on PT in AZD2693 Total
- AEs considered related to IP
- SAEs
- SAEs leading to death
- AEs leading to discontinuation of IP
- AEs by maximum reported intensity

For China, East Asia and Japan cohorts, separate summaries will be created for: All AEs, All AEs considered related to IP, AEs leading to discontinuation of IP and AEs by maximum reported intensity. For China and East Asia cohorts, separate summaries will be created for: All AEs sorted in decreasing frequency on PT in AZD2693 Total and for SAEs, SAEs leading to death, SAEs leading to discontinuation of IP and SAEs considered related to IP.

Injection site reaction (based on eCRF) and pro-inflammatory events (based on eCRF in combination with predefined PT list) will be summarised per SOC, PT (including separate summary for East Asia, China and Japan cohorts) and maximum intensity. Adverse events of injection site reactions will be summarized i) by symptom, ii) by symptom and injection site location and iii) by number of IP administrations and symptom. The other two AESI (platelet count reduction and liver toxicity) will not be summarised based on reported AEs.

Adjudicated adverse events will be summarised separately for the on-study period.

A list of key participant information for participants with SAEs with outcome of death, participants with SAEs, participants with AEs leading to discontinuation of IP and

participants with potential Hy's law will be provided. The durations reported in these tables, will be derived only for fully completed dates as below:

- Time from first dose of IP to AE (in days) will be calculated as the AE start date minus date of dose + 1.
- Time from first dose to death (in days) will be calculated as the date of death minus date of dose + 1.
- The same approach will be used for deriving time from start of treatment to AE becoming serious or discontinuation.
- Time from last dose prior to AE start and last dose prior to death will be calculated as the date of death minus the date of last dose prior to AE/death + 1.
- If an SAE occurred before the IP administration the time from IP to SAE will be negative and calculated as SAE start date – IP start date

All AEs and SAEs will be listed for the on-study period. SAEs occurring in the screening period will be listed separately. SAEs for enrolled but not randomised subjects and for subjects who were not exposed to treatment will be listed. Listings with details on injection site reactions based on eCRF and on each adjudicated adverse event base on transferred data will be provided.

4.6.3 Clinical Laboratory, Blood and Urine Samples

4.6.3.1 Definitions and Derivations

Haematology (whole blood)	
• B-Mean corpuscular haemoglobin (MCH)	• B-Mean corpuscular haemoglobin concentration (MCHC)
• B-Red blood cell (RBC) count	• B-Neutrophils absolute count
• B-Platelet count	• B-Lymphocytes absolute count
• B-Haemoglobin (Hb)	• B-Monocytes absolute count
• B-Haematocrit (HCT)	• B-Basophils absolute count
• B-Mean corpuscular volume (MCV)	• B-Reticulocytes absolute count
• B-White blood cell (WBC) count	• B-Eosinophils absolute count

Clinical chemistry (serum or plasma) • Creatinine • Bilirubin, total and direct • Alkaline phosphatase (ALP) • Aspartate transaminase (AST) • Alanine transaminase (ALT) • Gamma-glutamyl transferase (GGT) • Potassium • Calcium, total • Sodium • Magnesium	• Uric acid • Urea • Phosphate • Bicarbonate • Thyroid stimulating hormone (TSH; screening) • Glucose • Insulin • C-peptide • Haemoglobin A1C(HbA1C) • Glutamate dehydrogenase (GLDH)
Urinalysis (dipstick) • Blood • Protein	• Glucose • Microscopy (if positive for protein or blood)
Renal safety biomarkers • Total protein • Albumin	• Creatinine • UACR (derived)
Coagulation • International normalised ratio (INR) • Activated partial thrombin time (APTT)	• Prothrombin time (PT)

Additional laboratory variables • Estimated glomerular filtration rate (eGFR)	• Retinol • Retinol binding protein-4 (RBP-4)
Inflammatory panel • High-Sensitivity C-reactive protein (hs-CRP) • Albumin • Complement C3 and C5 (Total, not activated complement)	• Fibrinogen • Monocyte chemoattractant protein-1 (MCP-1) • Interleukin-6 (IL-6)
Thyroid panel • Free thyroxine (free T4)	
Lipid panel • Low density Lipoprotein (LDL, measured) • High Density Lipoprotein (HDL) • Free fatty acid (FFA)	• Triglycerides (TG) • Total cholesterol
Complement Activation Panel • C5a	• Bb

Other laboratory safety assessments:

- Pregnancy test: Human-beta chorionic gonadotrophin (b-HCG)
- Hormone testing: Follicle-stimulating hormone (FSH)

LLOQ and ULOQ

Laboratory parameters below the lower limit of quantification (LLOQ) for which the exact value cannot be determined, will be replaced with the LLOQ/sqrt(2) value. Similarly,

laboratory test results above the upper limit of quantification (ULOQ) for which the exact value cannot be determined, will be replaced with the ULOQ value. After the LLOQ/ULOQ replacement, change from baseline to each post-baseline visit is defined as the post-baseline visit value minus the baseline value. After the LLOQ/ULOQ replacement laboratory parameters are also classified as normal (if value is within normal reference range), low (if value is below the normal reference range), and high (if value is above the normal reference range) based on the reference range indicator. Tables will utilize the LLOQ/ULOQ replacement. However, listings will display the original LLOQ/ULOQ inequality (e.g. "< x" and "> x").

Abnormalities in laboratory parameters

The following laboratory parameters will be categorised into non-mutually exclusive thresholds:

- Platelet count: < LLN, >ULN, < 50 x 10^9/L, < 75 x 10^9/L, < 100 x 10^9/L, < 150 x 10^9/L, >30% decrease from baseline, < 150 x 10^9/L and > 30% decrease from baseline
- ALT: < LLN, >ULN, > 3 x ULN, > 5 x ULN, > 10 x ULN, > 20 x ULN
- AST: < LLN, >ULN, > 3 x ULN, > 5 x ULN, > 10 x ULN, > 20 x ULN
- Alkaline Phosphatase: < LLN, >ULN, > 2 x ULN, > 3 x ULN
- Total Bilirubin (TBL): < LLN, >ULN, > 2 x ULN, > 5 x ULN, > 8 x ULN
- Direct Bilirubin: < LLN, >ULN, > 2 x ULN, > 5 x ULN
- INR: < LLN, >ULN, > 1.5 x ULN, > 3 x ULN, > 5 x ULN
- eGFR: ≥ 25% decrease, ≥ 50% decrease and ≥ 75% decrease
- Creatinine: < LLN, >ULN, ≥ 1.5 x baseline, ≥ 2 x baseline and ≥ 3 x baseline

Potential Hy's Law

AST or ALT ≥ 3 × ULN together with TBL ≥ 2× ULN at any point during the study following the start of study intervention irrespective of an increase in ALP. The elevation in transaminases must precede or be coincident with (i.e., on the same day) the elevation in TBL, but there is no specified timeframe within which the elevations in transaminases and TBL must occur.

Shift tables

Shift tables will be presented for the following parameters: ALT, AST, ALP, total and direct bilirubin, GGT, platelet count, GLDH, retinol and INR. For ALT and AST baseline values will be categorised into: ≤ 1 ULN, > 1 ULN and ≤ 3 ULN, > 3 ULN and ≤ 5 ULN and > 5 ULN. On-treatment/On-study values will be categorised into: ≤ 1 ULN, > 1 ULN and ≤ 3 ULN, > 3 ULN and ≤ 5 ULN, > 5 ULN and ≤ 10 ULN, > 10 ULN and ≤ 20 ULN, > 20 ULN. Total and direct bilirubin will be categorised into: ≤ 1 x ULN, > 1 x ULN and ≤ 2 x ULN, > 2 x ULN. ALP will be categorised into ≤ 2 x ULN, > 2 x ULN. GGT and retinol will be categorised (both at baseline and on-treatment/on-study) as < LLN, ≥ LLN and ≤ ULN, > ULN. Platelet absolute count will be categorised into < 50, ≥ 50 and < 75, ≥ 75 and < 100, ≥ 100 and < 150, ≥ 150 x 10^9/L. GLDH will be categorised into: ≤ 3 ULN, > 3 ULN and ≤ 5 ULN, > 5 ULN. INR will be categorised into: INR < 1, INR ≥ 1 and INR < 1.5, INR ≥ 1.5 and INR < 2 and INR > 2. Shift from baseline to minimum on-treatment/on-study will be summarized for the following parameters: platelet count, retinol. Shift from baseline to maximum on-treatment/on-study will be summarized for the rest of the above-mentioned parameters. Shift from baseline category (low, normal, high) to maximum on-treatment/on-study category (low, normal, high) will be presented for ALT, AST, ALP, total bilirubin, direct bilirubin, GGT, GLDH, INR, serum creatinine, hs-CRP, aPTT, PT, complement C3, complement C5a, complement Bb, IL-6. Shift from baseline category (low, normal, high) to minimum on-treatment/on-study category (low, normal, high) will be presented for platelet count, eGFR and retinol.

Other tables

Participants with elevated post baseline ALT (> 2 x baseline, >2 x baseline for >= 12 weeks, > 3 x baseline, >3 x baseline for >= 12 weeks, > 5 x baseline, >5 x baseline for >= 12 weeks, > 3 x ULN, >3 x ULN for >= 12 weeks) will be summarized by their baseline category (normal/near normal ALT (<1.5xULN), elevated ALT (≥1.5xULN)).

Table 11 Normal reference ranges for INR, GLDH, Complement C3, Complement C5a, Complement Bb and IL-6.

Parameter (unit)	Normal reference range
International normalized ratio	0.8 - 1.2
Glutamate dehydrogenase (IU/L)	≤ 9.999999
Complement C3 (g/L)	0.90 - 1.80
Complement C5a (ng/mL)	≤ 74.33

IL-6 (pg/mL)	< 12.50
Complement Bb (ug/mL)	≤ 1.42

4.6.3.2 Presentations

All tables, graphs, and listings for laboratory data will be based on the baseline and on-study/on-treatment data in the safety population and presented for each treatment group. Laboratory test results will be split for haematology, chemistry, urinalysis and other laboratory parameters (per category). Continuous haematology, chemistry and other laboratory parameters will be summarized with n, mean, SD, median, minimum, and maximum by treatment and analysis visit and for change from baseline for the on-study period. Frequency tables will present number of participants reporting at least one treatment emergent abnormality for both on-treatment and on-study periods, except for key subject information reporting at least one treatment emergent abnormality which will be summarized only for the on-study period. For urinalysis shift tables from baseline to last and worst on-treatment/on-study value will be presented for both on-treatment and on-study periods. Maximum on-treatment/on-study ALT and AST versus maximum on treatment/on-study bilirubin will be presented in a table for both on-treatment and on-study periods. Number of subjects with post baseline ALT value >2x baseline, >3xbaseline and >5xbaseline will be summarized in a table. Average of ALT, AST, ALP, bilirubin, GGT, platelet count, eGFR, serum creatinine and retinol over time will be displayed in figures for on-study period. Evaluation of drug-induced serious hepatotoxicity will be displayed in eDISH (peak of ALT/AST versus peak of TBL as multiples of ULN). Shift tables for the selected clinical chemistry and haematology parameters will be presented for both on-treatment and on-study periods. For participants with ALT>3xULN (at any timepoint on-study) trajectories of liver parameters will be displayed in spaghetti plots. Key subject information for laboratory abnormalities and individual laboratory results will be summarized and listed, respectively, for the on-study period. For the participants who had post baseline ALT >3 x ULN, a figure will present all measurements of the liver parameters ALT, AST, ALP, GGT, BILI for each participant. ALT, AST, Platelet count and B-Haemoglobin will be presented in a boxplot by timepoint. Shift tables and frequency tables for participants reporting at least one treatment emergent abnormality will be separately produced for China and East Asia cohorts, except for key subject information.

4.6.4 Vital Signs

4.6.4.1 Definitions and Derivations

The following vital signs will be measured:

- Weight, BMI, waist circumference

- Systolic blood pressure (mmHg)
- Diastolic blood pressure (mmHg)
- Pulse (bpm)
- Temperature

Vital signs will be collected at the visits indicated in the SoA (see CSP).

Blood pressure and pulse rate will be collected after the participant has rested in the supine position for at least 10 minutes for screening visit and in subsequent visits after resting in a sitting position for 5 minutes.

Temperature will be measured; oral, temporal, and tympanic measurements are allowed but should be consistent throughout the study.

Additionally, vital signs values will be classified as normal (if value is within normal reference range), low (if value is below the normal reference range), and high (if value is above the normal reference range) according to the normal reference ranges given in Table .

Table 12 Vital signs normal reference ranges

Parameter	Normal reference range
Systolic blood pressure	80-130 mmHg
Diastolic blood pressure	50-80 mmHg
Pulse rate	50-100 bpm
Temperature	≤ 37°C

4.6.4.2 Presentations

Observed values and change from baseline in vital signs will be summarised by study intervention group and visit with descriptive statistics for the on-study period. The number and percentage of participants with at least one treatment emergent vital sign abnormalities (Low/High values) will be summarised by abnormality criteria and intervention group. Key subject information for participants with abnormalities and individual vital signs will be summarized and listed, respectively, for the on-study period. Shift from baseline to maximum on-study value will be presented. Mean change from baseline will be shown for diastolic and systolic blood pressure over time for the on-study period in a figure. This summary will be separately produced for China and East Asia cohorts.

4.6.5 Electrocardiogram

4.6.5.1 Definitions and Derivations

At the visits and timepoints specified in the SoA, triplicate 10-second, 12-lead safety ECG will be obtained after 10 minutes of supine rest, using the study site's own ECG machines. The following parameters will be collected in the eCRF: cardiac intervals of RR, PR, QRS, QT, QTcF and the ECG mean heart rate. The PI will interpret the overall findings of 3 replicates as normal, abnormal clinically nonsignificant, or abnormal clinically significant. This evaluation will be reported in the eCRF. Additionally, clinically significant findings should be documented on the AE page of the eCRF if applicable.

ECG recordings will be performed prior to study intervention administration if applicable/scheduled. Starting from Visit 9 (Week 12), ECGs will be recorded approximately every 12 weeks; extra ECGs may be performed at the request of an investigator. ECG recording should be performed prior to any blood draw.

4.6.5.2 Presentations

ECG tables and listings will be based on the baseline and on-study data in the safety population and presented for each treatment group for the on-study period. All ECG data will be listed but only heart rate and QTcF will be summarised. When possible based on the data, descriptive statistics per visit and timepoint as well as change from baseline will be presented for heart rate and QTcF. For each visit and timepoint the ECG parameters for every participant will be averaged over the available replicates. Abnormalities in QTcF at any visit on-study will be summarised using following criteria:

- QTcF value above specific ms at any time during treatment:
  - > 450 and ≤ 480 (ms)
  - > 480 and ≤ 500 (ms)
  - > 500 (ms)
- QTcF increase from baseline by more than specific ms at any time during treatment:
  - > 30 and ≤ 60 (ms)
  - > 60 (ms)
- QTcF value above specific ms and QTcF increase from baseline by more than specific ms at any time during treatment:
  - Value > 450 and ≤ 500 (ms) and increase > 30 and ≤ 60 (ms)

- Value > 500 (ms) and increase > 60 (ms)

Abnormality in ECG (split for clinically significant and clinically non-significant) for baseline versus last on-study and versus worst on-study value will be presented for each treatment arm. Separate summary of abnormality in ECG and separate shift table to worst on-study will be produced for China and East Asia cohorts.

In the listings measurements of individual replicates as well as averages will be presented.

4.6.6 Other Safety Assessments

4.6.6.1 Definitions and Derivations

None.

4.6.6.2 Presentations

None.

5 INTERIM ANALYSIS

No interim analyses are planned for the primary endpoint. Safety analysis will be produced throughout the study. No stopping criteria are specified.

6 REFERENCES

Robins. (1986). Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models. *Biometrics*, 42(2):311-23.

Rubin. (1987). *Multiple imputation for nonresponse in surveys.* New York: John wiley & Sons.

Sato. (1989). On the variance estimator of the Mantel-Haenszel risk difference. *Biometrics*, 45:1323-4.

7 APPENDIX

None.