Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT03987451
Sponsor trial ID:	NN9931-4492
Official title of study:	Investigation of Efficacy and Safety of Semaglutide s.c. Once-weekly Versus Placebo in Subjects With Non-alcoholic Steatohepatitis and Compensated Liver Cirrhosis
Document date*	17 June 2021

*Document date refers to the date on which the document was most recently updated.
Note: The date in the header of Page 2 is the date of compilation of the documents and not of an update to content.

16.1.9 Documentation of statistical methods

List of contents

Redacted statistical analysis plan
Includes redaction of personal identifiable information only.

.

Statistical Analysis Plan

Investigation of efficacy and safety of semaglutide s.c. once-weekly versus placebo in subjects with non-alcoholic steatohepatitis and compensated liver cirrhosis

Substance: Semaglutide

Table of contents

Page

Table of figures

Page

Table of tables

Page

Version History

This Statistical Analysis Plan (SAP) for study NN9931-4492 is based on the protocol version 3.0 dated 21-FEB-2020.

Table 1 SAP version history summary

SAP Version	Date	Change	Rationale
1.0	**17-JUN-2019**	**Not Applicable**	**Original version**
2.0	**17-JUN-2021**	Clarification of primary analysis. Specification of secondary supportive analyses not described in protocol.	Updated to reflect protocol amendment.

1 Introduction

This SAP elaborates on the statistical analyses described in the protocol section 10. Additional analyses for supportive secondary endpoints are specified, see section 4.

1.1 Objectives, Endpoints, and Estimands

1.1.1 Objectives

Primary objective
To investigate the effect of semaglutide s.c. 2.4 mg once-weekly on liver fibrosis compared with placebo in subjects with NASH and compensated fibrosis stage 4.

Secondary efficacy objective
To investigate the effect of semaglutide s.c. 2.4 mg once-weekly on NASH compared with placebo in subjects with NASH and compensated fibrosis stage 4.

Secondary safety objective
To evaluate the safety and tolerability of semaglutide s.c. 2.4 mg once-weekly compared with placebo in subjects with NASH and compensated fibrosis stage 4.

1.1.2 Endpoints

Primary and secondary endpoints
See protocol table 4-1. or 6.2

1.1.3 Estimands

In subjects with NASH and compensated fibrosis stage 4, the estimand addressing the primary objective is the proportion of subjects with liver fibrosis improvement with no worsening of NASH compared between semaglutide s.c. 2.4 mg once-weekly and placebo at 48 weeks regardless of adherence to treatment. The treatment policy strategy is applied for the intercurrent event of premature discontinuation of randomised treatment.

Results based on this estimand are expected to mirror the clinical practice scenario because the estimand considers both the efficacy and tolerability of subcutaneously administered semaglutide.

1.2 Study Design

This is a 48 week, randomised, double-blind, placebo-controlled, two-armed, parallel group, multi-centre, multi-national trial comparing once-weekly administration of 2.4 mg semaglutide s.c. with placebo in subjects with NASH and fibrosis stage 4.

A total of 71 subjects were randomised in a 2:1 ratio.
The trial has a 6-week screening period followed by a randomisation visit and a 48-week treatment period. The treatment period id divided into a dose escalation period of 16 weeks and a maintenance period of 32 weeks. The follow up period is 7 weeks. The total trial duration is approximately 61 weeks (see Figure 1). For further trial detail see protocol section 5.

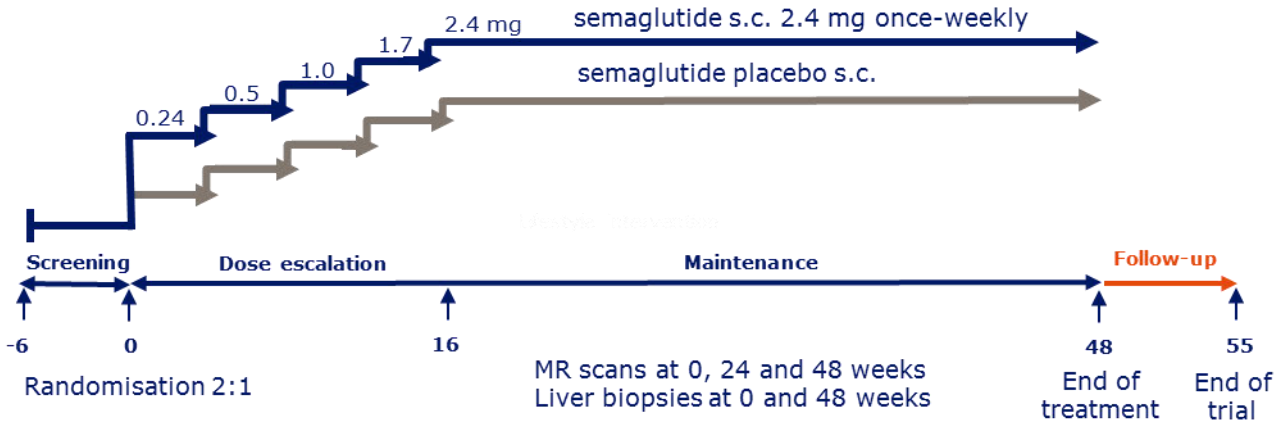

Figure 1 A schematic diagram of the trial design, with the duration of the trial periods including follow-up period.

2 Statistical Hypotheses

For the primary endpoint, at least one stage of liver fibrosis improvement with no worsening of NASH after 48 weeks, a confirmatory test comparing semaglutide 2.4 mg OW with placebo is planned. The comparison will test for superiority of semaglutide versus placebo using a one-sided significance level of 2.5% corresponding to a two-sided level of 5%. For remaining analyses, a two-sided 5% significance level will be used. For more detail see section 4.2.2.

2.1 Multiplicity Adjustment

Not applicable as only one confirmatory endpoint is stated.

3 Analysis Sets

Population sets

The following population analysis sets are defined in accordance with ICH-E9 guidance[1]

Participant Analysis Set	Description
Full analysis set (FAS)	Includes all randomised subjects. Subjects in the FAS will contribute to the evaluation "as randomised".
Safety analysis set (SAS)	Includes all subjects receiving at least one dose of randomised treatment. Subjects in the SAS will contribute to the evaluation "as treated".

Observation periods

Data will be evaluated based on different observation periods which will be derived individually for each subject. The following two observation periods are defined:

- In-trial: This period starts on the date of the randomisation visit and ends on the first of the following dates (both inclusive):
  a) follow-up visit (end-of-trial visit)
  b) withdrawal of consent
  c) last contact with subject (for subjects lost to follow-up)
  d) death
- On-treatment: For evaluation of adverse events (AEs), electrocardiogram (ECG) and hypoglycaemic episodes, this period starts on the date of first administration of trial product and ends on the date of whatever comes first of: a) last dose of trial product + 49 days (7 half-lives of semaglutide) or b) end of the in-trial period.
  For evaluation of MR scans and histology data the period ends at the date of the last dose of trial product +42 days.
  For evaluation of all other data, the period ends at the date of the last dose of trial product +7 days.

Data collected after the observation period in question will be excluded from any summary or analysis based on that observation period.

4 Statistical Analyses

4.1 General Considerations

The statistical analyses will in general consist of the pairwise treatment comparison between:

Semaglutide s.c. 2.4 mg OW versus Placebo

The results of the comparisons will be presented as estimated treatment contrasts together with two-sided 95% confidence intervals and p-values corresponding to two-sided alternative tests of the null hypothesis of no difference. In addition, for the primary analysis i.e. confirmatory test of superiority, the one-sided p-value will be reported additionally to the two-sided p-value, see section 4.2.2 for details.
The statistical analyses of the efficacy endpoints will primarily be based on the in-trial period. The on-treatment period is used for some supportive efficacy analyses and all statistical analyses of safety endpoints. Summary statistics will in general be presented for both observation periods at the end of treatment visit (visit 12/12A).

The baseline measurement is defined as the latest available measurement at or prior to the randomisation visit. An exception is made for identifying abnormal laboratory findings (including ALT, AST, GGT, bilirubin and INR) in which case the baseline value is defined as the mean of the available measurements at the screening and randomisation visits.

Laboratory values below the lower limit of quantification (LLOQ) will be set to ½LLOQ while values above upper limit of quantification (ULOQ) will be set to ULOQ.

The stratification variable given by type 2 diabetes (T2D) status (yes/no) at baseline will be included in the models as a factor.

When using body weight measurements, body weight will be used without consideration to fasting status.

For continuous endpoints the change from baseline of the response will be given by either absolute change from baseline, logarithm of the ratio to baseline or percental change from baseline.

4.2 Primary Endpoint/Estimand *Analysis*

4.2.1 Definition of Endpoint

The primary endpoint is the binary outcome defined as: At least one stage of liver fibrosis improvement with no worsening of NASH. Improvement in fibrosis is defined by a decrease from baseline to week 48 of at least one stage according to the Kleiner fibrosis classification. Worsening of NASH is defined as an increase of at least one stage of either lobular inflammation, hepatocyte ballooning or steatosis as defined by the NASH Clinical research network (CRN).

4.2.2 Main Analytical Approach

The primary analysis will be based on the Cochran-Mantel-Haenszel (CMH) test. The common odds ratio between semaglutide and placebo adjusting for baseline diabetes will be estimated along

with exact 95% confidence interval based on conditioning on the marginal 2×2 tables.
The exact one-sided p-value will be calculated for testing superiority as the sum of probabilities of outcomes being equal to or more in favour of semaglutide under the null hypothesis conditioning on the marginal 2×2 tables.
The two-sided p-value will be calculated as the sum of probabilities of outcomes having equal or lower probability than the observed outcome.

The primary analysis is based on FAS in-trial. Response data will consist of histological readings of the week 48 biopsies including biopsies taken after premature discontinuation of trial product. Missing or non-eligible data will be imputed as non-responders.

4.2.3 Sensitivity Analysis

To investigate the sensitivity of the result of the primary analysis with regard to the handling of missing data, the following sensitivity analysis will be performed:

An analysis where missing data will be handled by reference-based multiple imputation (MI) that is informed by data from placebo subjects. It is assumed that subjects in either treatment group without an observed outcome have the same chances of meeting the endpoint as subjects in the placebo group with an observed outcome. Subjects dying before week 48 visit will be handled as non-responders, i.e. these subjects will not be considered to have met the endpoint regardless of imputation.

Technically, the analysis will be performed in the following steps:

1. 500 replicates of the non-missing data in the placebo group are generated using a non-parametric resampling scheme with replacement within the two strata given by diabetes status (with or without T2D). For each replicated data set and each stratum the responder proportion (i.e. probability of meeting the endpoint) is estimated by the observed sample proportion.
2. The observed data including missing outcomes is replicated 500 times.
   For each replicate both treatment groups' missing outcomes are imputed by independent draws from a Bernoulli distribution using the corresponding estimated responder proportion as parameter within strata.
3. For each of the 500 complete data sets, the log common odds ratio adjusting for stratum is estimated using the Mantel-Haenszel estimator together with Robins, Breslow, and Greenland[2] estimate of the standard error of the log common odds ratio. The estimated log common odds ratio and standard errors for the 500 complete data sets are then pooled using Rubin's rule[3]:

$$m_{MI} = \frac{1}{N}\sum_{i=1}^{N} m_i\,, \qquad SE_{MI} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} SE_i^2 + \left(1+\frac{1}{N}\right)\left(\frac{1}{N-1}\right)\sum_{i=1}^{N}(m_i - m_{MI})^2}\,,$$

   where m_i and SE_i are the estimated log odds ratios and corresponding standard errors for the $N = 500$ data sets, and m_{MI} and SE_{MI} are the pooled estimates. From m_{MI} and SE_{MI}, the 95% confidence intervals for the odds ratios and associated p-values are calculated.

In step 3, zero counts in cells may not allow for calculation of the Mantel-Haenszel estimator, consequently all sample proportions will be equal to zero yielding with in-between variation of zero. In this case the above analysis will not be carried out.

4.2.4 Supplementary Analyses

The following two supplementary analyses will be made for the primary endpoint

1. A complete case on-treatment analysis: The same as the primary analysis but where subjects with missing week 48 data or for whom the data were collected after the on-treatment period are excluded from the analysis.
2. An analysis excluding subjects whose primary endpoint is missing due to COVID-19. Subjects have been identifies through protocol deviations (PD) where it has been deemed that attainment of primary endpoint has not been possible primarily due to COVID-19. The analysis-excludes subjects N = 3, listed by unique subject id: NN9931-4492/[redacted], NN9931-4492/[redacted], NN9931-4492/[redacted].

4.3 Secondary Endpoint *Analysis*

Not applicable

4.3.1 Supportive Secondary Endpoints

4.3.2 Efficacy endpoints

4.3.2.1 Histology

Histology scores

Analyses will be performed for the change from baseline in the following histological features, all of which are ordinal scores except hepatic collagen which is a proportion:

- Fibrosis stage according the Kleiner fibrosis classification: (0-4)
- Hepatocyte ballooning, CRN score: (0-2)
- Lobular inflammation CRN score: (0-3)
- Steatosis CRN score: (0-3)
- Total CRN NAS: (0-8)
- Activity component of SAF: (0-4)
- Ishak fibrosis score: (0-7)
- Hepatic collagen, proportionate area: (0-100).

The activity component of the SAF score is defined as the unweighted sum of hepatocyte ballooning and lobular inflammation. The definition of the lobular inflammation score is modified in this calculation so that the scores 2 and 3 on the original scale are merged to a score of 2. The possible range of the sum is thus 0 to 4. For all scores, a higher value indicates a more severe state of disease.

Ordinal histological feature scores (i.e. excluding hepatic collagen) will be analysed by ordered logistic regression (also known as proportional odds model) with the histological scores at week 48 as response, treatment, baseline diabetes status as factor; and baseline body weight and

corresponding histological score at baseline as covariate.
The results will be presented as the estimated cumulative odds ratio between semaglutide and placebo. The analyses will be based on FAS in-trial population and missing week 48 data will be imputed as no change from baseline in agreement with the analyses of the binary histological endpoints.

Change in hepatic collagen from baseline to week 48 will be analysed using ANCOVA and MMRM. See section 4.3.2.2 for description of ANCOVA and MMRM analyses.

Histology, derived binary endpoints
The binary endpoint based on histology NASH resolution after 48 weeks (yes/no) defined by NASH CRN value of hepatocyte ballooning of 0 and lobular inflammation score 0-1 will be analysed as the primary endpoint (see section 4.2.2) including the complete case on-treatment supplementary analyses as for the primary endpoint and the multiple imputation sensitivity analysis (see section 4.2.3).

The composite binary endpoint based on histology defined by NASH Resolution and improvement in fibrosis will be analysed following the same approach as for the primary analysis.

Further, binary histological endpoints defined as improvement in histological score from baseline to week 48 (Yes/No) will be derived and analysed. Improvement is defined as at least one stage decrease from baseline to week 48 in corresponding histological score:

- Improvement in Fibrosis stage according the Kleiner fibrosis classification (Yes/No)
- Improvement in Hepatocyte ballooning, CRN score (Yes/No)
- Improvement in Lobular inflammation CRN score (Yes/No)
- Improvement in Steatosis CRN score (Yes/No)
- Improvement in Total CRN NAS (Yes/No)
- Improvement in Activity component of SAF (Yes/No)
- Improvement in Ishak fibrosis score (Yes/No)

The derived binary endpoints will be analysed using a CMH test following the same approach as for the primary analysis (see section 4.2.2).

4.3.2.2 Biomarkers of NASH disease

Analyses will be performed for the change from baseline to week 48 in the following biomarkers for NASH disease:

- Algorithms
  - Fibrosis-4 (Fib-4) score.
  - NAFLD Fibrosis Score (NFS)
- Blood samples
  - Portal hypertension
    - Thrombocytes
  - Liver enzymes
    - Alanine aminotransferase (ALT)
    - Aspartate aminotransferase (AST)
    - Gamma glutamyl transferase (GGT)
  - Liver synthesis function
    - Albumin
    - International normalized ratio (INR)
    - Total Bilirubin
    - Direct Bilirubin
  - Exploratory biomarkers
    - Adiponectin
    - Released N-Terminal Pro-Peptide C3 (Pro-C3)
    - Enhanced liver fibrosis (ELF), including subcomponents:
      - Tissue inhibitor of metalloproteinase 1 (TIMP1)
      - Hyaluronic acid (HA)
      - Procollagen III Amino Terminal Peptidecollagen (PIIINP)
- Imaging
  - Liver stiffness (MRE) (kPa)
  - Liver fat content, steatosis (MRI-PDFF) (%)
  - Liver fat volume (MRI) (L)
  - Total liver volume (MRI) (L)

All above parameters except NFS and ELF will be logarithmically transformed in the statistical analysis. The estimated treatment differences will subsequently be back-transformed to the original scale as the estimated treatment ratios.

Fib-4 score will be derived according to the formula[4]:

$$\text{Fib-4} = \frac{\text{Age (years)} \times \text{AST (U/L)}}{\text{Thrombocyte count } (10^9/\text{L}) \times \sqrt{\text{ALT (U/L)}}}$$

The derivation will be performed at each visit where ALT, AST and thrombocyte count have been assessed. If any of the three laboratory parameters is missing at a specific visit, the Fib-4 score will be considered missing as well.

The NAFLD Fibrosis Score will be derived according to the linear regression formula[5]:

$$\text{NFS} = -1.675 + 0.037 \times \text{Age (years)} + 0.094 \times \text{BMI (kg/m}^2\text{)} + 1.13 \times \text{Hyperglycaemia (yes/no)} + 0.99 \times \text{AST/ALT} - 0.013 \times \text{Thrombocyte count (}10^9\text{/L)} - 0.66 \times \text{Albumin (g/dL)}$$

Hyperglycaemia (yes/no) is a binary variable defined as 1 if FPG ≥ 6.1 mmol/L (110 mg/dL) at the corresponding visit of assessment or the subject has been diagnosed with T2D at screening; otherwise, the variable is defined as 0. NFS will be derived at each visit where body weight, FPG, ALT, AST, thrombocyte count and albumin have been assessed. If any of the elements are missing, NFS will be considered missing as well.

The ELF discriminant score will be derived as a log-linear combination of the markers hyaluronic acid (HA), amino-terminal propeptide of type III collagen (PIIINP) and tissue inhibitor of metalloproteinase 1 (TIMP1). The derivation will be performed at the central laboratory according to the instructions of the equipment manufacturer.

The main analysis of each of the biomarkers will based on FAS in-trial using analysis of covariance (ANCOVA) with missing outcomes handled by unconditional reference-based imputation.

This will be done following the steps:

1. An ANCOVA model with baseline diabetes status as factor and baseline body weight and baseline value of the corresponding biomarker as covariates is fitted to the change from baseline to week 48 for the placebo group only.
2. Using the fitted model 500 sets of values of the model parameters (regression coefficients and residual variance) are drawn from the posterior distribution of the estimators.
   For each draw of parameter values, missing week 48 outcomes across treatment groups are imputed as predicted values from the linear combination of drawn model parameters and data. Subjects with missing covariates e.g. baseline measurement will not be included.
3. For each of the 500 complete data set, the treatment effect on change from baseline to week 48 is estimated using an ANCOVA model with treatment and baseline diabetes status as factors, and baseline body weight and baseline biomarker as covariates.
4. The 500 estimated treatment differences and standard errors are pooled using Rubin's rule. From the pooled estimates and standard errors the 95% confidence interval for treatment difference and associated two-sided p-value is calculated.

A supportive analysis based on FAS on-treatment using a mixed model for repeated measurements (MMRM) will be performed for each of the biomarkers. In this model, all scheduled post-baseline measurements taken during the individual subject's on-treatment period will enter as response; planned treatment and baseline diabetes status will be included as factors; and baseline body weight and baseline value of the corresponding biomarker will be included as covariates. All factors and covariates will be nested within visit and an unstructured covariance matrix for within subject residual errors will be employed. There will be no explicit imputation of missing values. As for the main analysis, the estimated treatment differences at week 48 with associated 95% confidence intervals and two-sided p-values will be presented.

Whereas the main analysis attempts to estimate the de-facto treatment effect (in agreement with the chosen estimand for the primary objective), the supportive analysis aims to estimate the de-jure effect that would have been observed if all subjects had remained on treatment and completed all

visits. The latter analysis relies on the assumption that data are missing at random, which means that given the observed data, the events that lead to data being missing are independent of the unobserved data.

4.3.2.2.1 Weight-related parameters

The endpoints related to weight are defined as change from baseline to week 48 in:

- Body weight (% and kg)
- Waist circumference (cm)
- Body mass index (BMI)

These endpoints will be analysed separately based on FAS in-trial using the same type of ANCOVA with MI as for the biomarkers with treatment, baseline diabetes status and corresponding baseline value as a single covariate. Supportive FAS on-treatment analyses will be performed based on MMRM in the same way as for the biomarkers with factors and covariates specified as in the ANCOVA model.

Weight-derived binary endpoints

In addition to the continuous endpoints, the following binary endpoints related to weight will be analysed separately:

- Weight loss of ≥ 5% of baseline body weight at 48 weeks (yes/no)
- Weight loss of ≥ 10% of baseline body weight at 48 weeks (yes/no)

The binary endpoints will be compared between treatment arms using an MI approach similar to the continuous endpoints but based on logistic regression:
The 500 data sets with imputed values generated in the ANCOVA for percent change in body weight will be reused to derive an equal number of complete data sets for the binary outcomes. For each data set, the binary outcomes will be analysed using a logistic regression model which includes treatment, diabetes status as factors and baseline body weight as a covariate. The estimated log odds ratios and standard errors for the 500 complete data sets are then pooled using Rubin's rule. From the pooled estimates and standard errors, the 95% confidence intervals for the odds ratios and associated p-values are calculated.

4.3.2.2.2 Glucose metabolism related parameters

The secondary endpoints related to glucose metabolism are defined as change from baseline to 48 weeks in:

- Glycosylated haemoglobin A1c (HbA_{1c})
- Fasting plasma glucose (FPG)
- Fasting C-peptide

Before analysis fasting C-peptide will be logarithmically transformed. The endpoints will be analysed using the same type of MMRM and ANCOVA with MI as for the biomarkers. Only subjects with type 2 diabetes at screening will be included consequently diabetes statues will not be

included as a factor in the models. The model will include treatment as factor; and baseline bodyweight and corresponding baseline value as covariates.

4.3.2.2.3 Cardiovascular risk factors

The secondary endpoints related to cardiovascular risk factors are defined as change from baseline to 48 weeks in:

- Systolic and diastolic blood pressure
- Lipids (total cholesterol, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, very low-density lipoprotein cholesterol, triglycerides, free fatty acids)
- High-sensitivity C-reactive protein (hsCRP)

These endpoints will be analysed separately using the same type of MMRM and ANCOVA with MI as for the biomarkers The lipids and hsCRP will be logarithmically transformed before analysis. The model will include treatment and diabetes status as factors; and baseline body weight and corresponding baseline value covariates.

4.3.2.3 Safety endpoints

Specifications of safety tables, figures and listings (TFL) and other specifications not included in this SAP will be described in the mock TFL.

The following secondary endpoints are used to support the safety objectives.

- Number of treatment-emergent adverse events during the trial
- Number of treatment-emergent hypoglycaemic episodes during the trial
- Change from baseline to 48 weeks in pulse

4.3.2.3.1 Adverse events

AEs will be coded using version 24.0 of the Medical Dictionary for regulatory Activities (MedDRA) coding. A treatment emergent adverse event (TEAE) is defined as an event that has onset date during the on-treatment period (see section 3).

AE data will be displayed in terms of the number of subjects with at least one event, the percentage of subjects with at least one event, the number of events and the event rate per 100 patient years of exposure. The main AE summaries will only contain TEAEs. Non-treatment emergent AEs will be included in listings and overview summaries.

Additional summaries will be made for select AEs based on predefined MedDRA searches. A list of MedDRA search terms will be specified and documented before database lock.

4.3.2.3.2 Hypoglycaemic episodes

For subjects with type 2 diabetes at randomisation, the severe or BG confirmed symptomatic hypoglycaemic episodes will be analysed as a binary outcome using Fisher's exact test for the comparison between semaglutide and placebo. The results will be presented as estimated odds ratio with two-sided p-value.

For details on the classification of hypoglycaemia, see protocol section 9.2.10.

4.3.2.3.3 Pulse

Pulse will be summarised by descriptive statistics and analysed using the same type of MMRM as for the on-treatment analyses of the biomarkers with actual treatment, baseline diabetes status as factors; baseline bodyweight and baseline pulse as covariates.

4.4 Exploratory Endpoints/Estimand Analysis

Exploratory analysis to evaluate the performance of the biomarkers as predictors of NASH, fibrosis and/or NAS components are not included since they do not concern the trail objectives.

4.5 Other Safety Analysis

Not applicable

4.6 Other Analyses

4.6.1 Other Variables and/or Parameters

Not applicable

4.6.2 Subgroup Analyses

Not applicable

4.7 Interim Analyses

No interim analyses or other analyses of unblinded data will be performed before the database lock.

4.8 Changes to Protocol-planned Analyses

Estimand
The estimand stated in the protocol is updated in the SAP (see section 1.1.3) to align with the specified primary analysis approach in the protocol.

Primary endpoint
An additional analysis of the primary endpoint has been added that excludes subjects from primary analysis if endpoint is missing mainly due to COVID-19 (see section 4.2.4)

Supportive secondary endpoints
Addition of new binary endpoints given as improvement in the different histological ordinal outcomes that are analysed separately using the method of primary endpoint (see section 4.3.2.1).

5 Sample size determination

Refer to protocol section 10.1 for sample size considerations.

6 Supporting Documentation

6.1 Appendix 1: List of abbreviations

AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
BG	blood glucose
BMI	body mass index
CAP	controlled attenuation parameter
CK-18	cytokeratin 18
CMH	Cochran-Mantel-Haenszel
ECG	electrocardiogram
eCRF	electronic case report form
ELF	enhanced liver fibrosis
FAS	full analysis set
FGF-21	fibroblast growth factor 21
Fib-4	fibrosis-4 score
FPG	fasting plasma glucose
GGT	gamma glutamyltransferase
HbA_{1c}	glycosylated haemoglobin A1c
HOMA-IR	homeostatic model assessment - insulin resistance
hsCRP	high-sensitivity C-reactive protein
ICH	International Council on Harmonization
IL-1R	interleukin-1 receptor
INR	international normalized ratio
LLOQ	lower limit of quantification
MCP-1	monocyte chemoattractant protein 1
MedDRA	medical dictionary for regulatory activities
MI	multiple imputation
miR-122	microRNA 122
MMRM	mixed model for repeated measurements
NAFLD	non-alcoholic fatty liver disease
NAS	NAFLD activity score
NASH	non-alcoholic steatohepatitis
NFS	NAFLD fibrosis score
SAP	statistical analysis plan
SAS	safety analysis set
s.c.	subcutaneous
T2D	type 2 diabetes
TEAE	treatment emergent adverse event

6.2 Appendix 2: Definition and calculation of endpoints, assessments and derivations

Refer to protocol section 4.2 Table 4-1.

Endpoints analysed but not specified in protocol are:

- Improvement in Fibrosis stage according the Kleiner fibrosis classification
- Improvement in Hepatocyte ballooning, CRN score
- Improvement in Lobular inflammation CRN score
- Improvement in Steatosis CRN score
- Improvement in Total CRN NAS
- Improvement in Activity component of SAF
- Improvement in Ishak fibrosis score
- Resolution of NASH and improvement in fibrosis

Where improvement is defined as at least one stage decrease from baseline to week 48. Resolution of NASH and improvement in fibrosis is as defined in section 4.3.2.1.

7 References

1. ICH E9 Expert Working Group. Statistical principles for Clinical Trials: ICH Harmonised Tripartite Guideline. Statistics in Medicine. 1999. p. 1905-42.
2. A A. Categorical Data Analysis. 2nd ed. New York ed: John Wiley & Sons; 2002.
3. Little R, Rubin D. Statistical analysis with missing data. Sons. JW, editor. New York.: John Wiley & Sons. 1987.
4. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology. 2006;43(6):1317-25.
5. Angulo P. Long-term mortality in nonalcoholic fatty liver disease: is liver histology of any prognostic significance? Hepatology. 2010;51(2):373-5.