Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT03914326
Sponsor trial ID:	EX9924-4473
Official title of study:	Semaglutide cardiovascular outcomes trial in patients with type 2 diabetes (SOUL)
Document date:	14 December 2022

Statistical Analysis Plan

Trial ID: EX9924-4473

SOUL – Semaglutide cardiovascular outcomes trial in patients with type 2 diabetes

Table of contents

Page

List of abbreviations

ANCOVA	analysis of covariance
CI	confidence interval
CKD	chronic kidney disease
CKD-EPI	chronic kidney disease - epidemiology collaboration
CV	cardiovascular
DBL	data base lock
DMC	data monitoring committee
EAC	event adjudication committee
eGFR	estimated glomerular filtration rate
FAS	full analysis set
HbA_{1c}	glycosylated haemoglobin
HF	heart failure
HHF	hospitalisation for heart failure
HR	hazard ratio
LTFU	lost to follow-up
MACE	major adverse cardiovascular event
MALE	major adverse limb event
MI	myocardial infarction
PAD	peripheral arterial disease
SAP	statistical analysis plan
TTE	time-to-event

1 Introduction

1.1 Trial information

1.1.1 Rationale

To evaluate the hypothesis that oral semaglutide lowers the risk of cardiovascular events in patients with type 2 diabetes at high risk for cardiovascular disease.

1.1.2 Objectives, endpoints and estimand

Primary objective

To demonstrate that oral semaglutide lowers the risk of major adverse cardiovascular events compared to placebo, both added to standard of care in patients with type 2 diabetes and at high risk of cardiovascular events.

Key secondary objectives

To compare the effects of oral semaglutide versus placebo, both added to standard of care in patients with type 2 diabetes and at high risk of cardiovascular events with regards to:

- Chronic kidney disease
- Cardiovascular events
- Peripheral artery disease

Primary endpoint

The primary endpoint is time from randomisation to first occurrence of a major adverse cardiovascular event, a composite endpoint consisting of: cardiovascular death, non-fatal myocardial infarction or non-fatal stroke.

Confirmatory secondary endpoints

Time from randomisation to:

- first occurrence of a composite chronic kidney disease endpoint consisting of: cardiovascular death, renal death, onset of persistent ≥50% reduction in estimated glomerular filtration rate (CKD-EPI) compared with baseline, onset of persistent eGFR (CKD-EPI) <15 mL/min/1.73 m^2 or initiation of chronic renal replacement therapy (dialysis or kidney transplantation)
- cardiovascular death
- first occurrence of a major adverse limb events, a composite endpoint consisting of: acute limb ischemia hospitalisation or chronic limb ischemia hospitalisation

Primary estimand

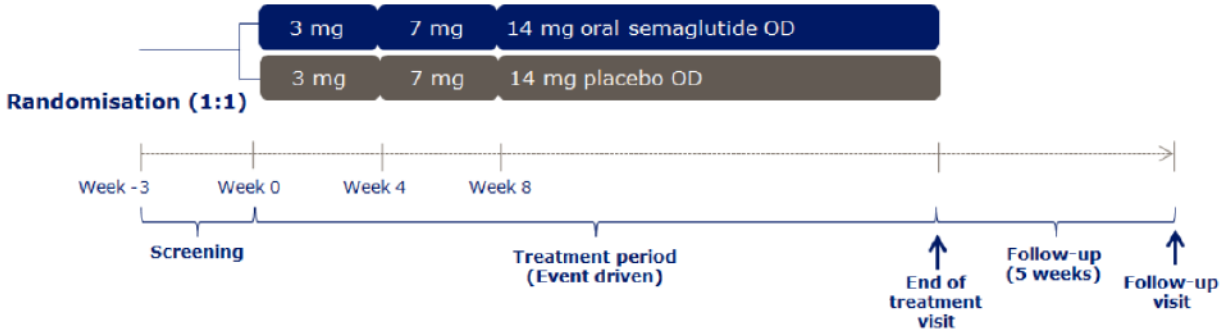
The estimand for all objectives is the intention-to-treat estimand evaluating the effect of the randomised treatment intervention irrespective of adherence to treatment and changes to background medication.

1.1.3 Design

This is a randomised, double-blind, parallel-group, placebo-controlled trial comparing oral semaglutide versus placebo both administered once daily and added to standard of care in patients with type 2 diabetes at high risk of cardiovascular events. Patients are randomised 1:1 to receive either oral semaglutide or placebo.

The trial is event driven; therefore, end of trial is scheduled according to accrual of events. The trial will employ a group sequential design with one interim testing for superiority. Under the design assumptions, the trial duration is approximately 61 months following randomisation of the first patient. 9,642 patients are planned to be randomly assigned to trial products. A schematic overview of the trial design is shown in Figure 1.

Figure 1 Trial design



1.2 Scope of the statistical analysis plan

The SAP includes elaborations on statistical analyses outlined in the protocol for SOUL trial (EX9924-4473) as well as details on the interim testing for superiority. Any changes to the SAP after first patient first visit are documented in a change log.

An external independent statistical service provider will conduct the interim analysis, see also section 3. Novo Nordisk is responsible for all other statistical analyses and reporting of data but will remain blinded to treatment allocations until data base lock (DBL). Additionally, a statistician independent of trial conduct, DMC analyses, interim analysis, and external to Novo Nordisk will independently confirm the statistical primary analyses of the primary endpoint and secondary confirmatory endpoints. This statistician will also be blinded until DBL.

2 Statistical considerations

2.1 Sample size determination

The trial is designed with 90% power to confirm superiority for the primary endpoint, i.e., reject the null-hypothesis of a hazard ratio (HR) ≥ 1.0 against the one-sided alternative of HR < 1.0, where HR is the hazard ratio of oral semaglutide versus placebo.

The trial is designed with one interim testing for superiority of the primary endpoint when two thirds of the total planned number of primary endpoint events has been accrued. Testing for futility

is not included. The Lan-DeMets alpha spending function, approximating the O'Brien-Fleming's stopping boundaries, is used to test superiority at a study-wise one-sided type I error rate of 2.5%. The one-sided alpha spending function is given by

$$f(t) = min\{2 - 2 \cdot \Phi(z_{\frac{\alpha}{2}}/\sqrt{t}), \alpha\}$$

where t is the proportion of information included in the interim analysis (accrued primary endpoint events relative to the total planned primary endpoint events), Φ denotes the standard normal cumulative distribution function, α is the overall one-sided alpha of 2.5% and $z_{\alpha/2}$ is the 98.75% quantile of the standard normal distribution. Based on a randomisation ratio of 1:1 and a design HR of 0.83 a total of 1,225 primary endpoint events are required.

For calculating the number of randomised patients, the following is assumed:

- annual primary endpoint rate in the placebo group of 3.5%
- uniform recruitment in 18 months
- annual lost to follow-up rate in both treatment groups of 1%
- maximum trial duration of five years and five weeks

Under these assumptions, a total of 9,642 patients are needed for randomisation.

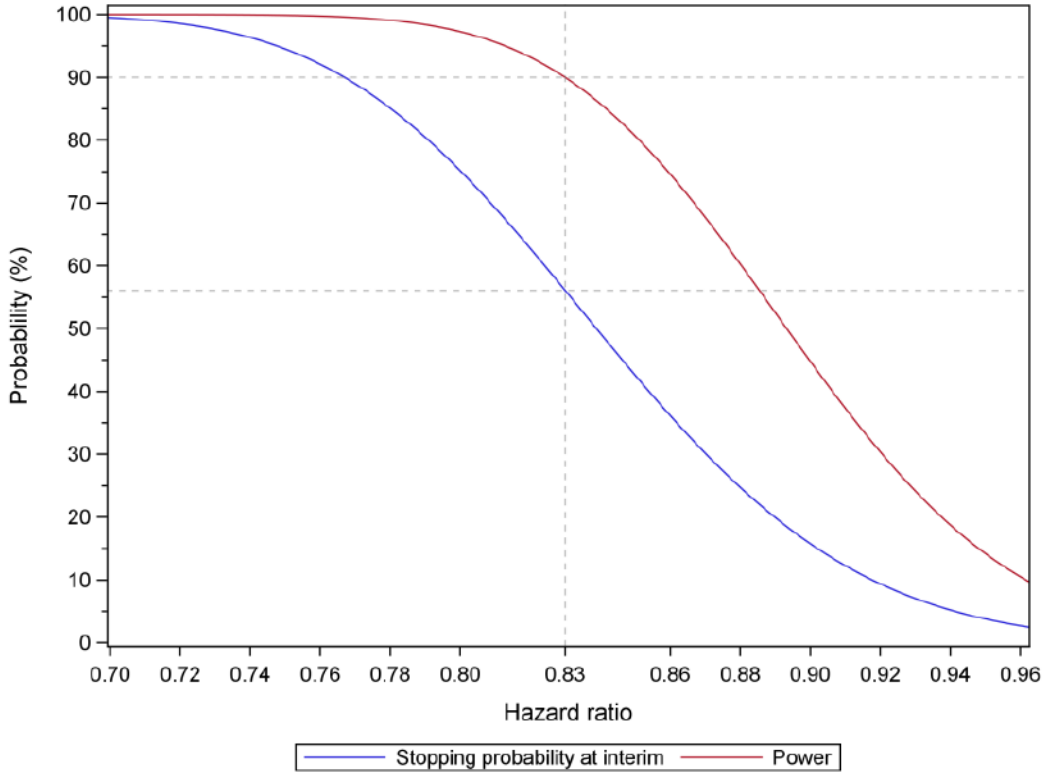
A Cox model as described in section 2.3 is used for the interim testing using the fixed sample one-sided lower p-value from the score test. Only a fixed sample p-value below the boundary specified by the error spending function will allow the DMC to recommend early trial termination for superiority. Table 1 provides the boundaries based on analyses performed after 817 and 1,225 events, along with the approximate hazard ratio estimates that correspond to those boundaries if the analyses are timed exactly to that schedule. The actual stopping boundaries will be based on the exact number of events available for the interim analysis.

Table 1 Stopping boundary scales at interim and scheduled termination

Stopping boundary scale	Interim 817 events	Scheduled termination 1,225 events
Hazard ratio	0.8389	0.8924
Nominal significance level	0.00605	0.02314

Figure 2 shows the probability of stopping the trial early at the interim (blue curve) and the overall power for confirming superiority for the primary endpoint (red curve) as a function of alternative values for the true HR. The design HR of 0.83 is marked with a dashed vertical reference line. The stopping probability at the interim and overall power for the design HR of 0.83 can be seen to be 56% and 90%, respectively.

Figure 2 Stopping probability at interim and overall power as a function of true hazard ratio



Confirmatory secondary endpoints

If superiority is confirmed for the primary endpoint the confirmatory secondary endpoints are controlled for multiplicity through a hierarchical testing strategy. The marginal powers below are calculated under the assumptions that the trial continues to the scheduled termination, a significance level of 2.5% (one-sided) and 9,642 randomised patients.

The marginal power for superiority in favour of oral semaglutide for the 5-component chronic kidney disease (CKD) endpoint is 94% based on an assumed HR of 0.80 and an annual event rate of 2.8% in the placebo group.

The marginal power for superiority in favour of oral semaglutide for CV death is 56% based on an assumed HR of 0.83 and an annual event rate of 1.4% in the placebo group.

The marginal power for superiority in favour of oral semaglutide for the major adverse limb event(s) (MALE) endpoint is 44% based on an assumed hazard ratio of 0.75 and an annual event rate of 0.44% in the placebo group.

The assumptions for annual event rate of primary endpoint and confirmatory secondary endpoints, lost to follow-up rates and the assumed HRs are based on the LEADER and SUSTAIN 6 CV outcomes trials.

2.2 General considerations

For confirmatory endpoints controlled for multiplicity, estimated treatment effects are presented together with two-sided 95% confidence intervals (CIs) and one-sided p-values for tests of the hypothesis of superiority. For reporting of results, the hazard ratio and the 95% CI are accompanied by the two-sided p-value.

For non-confirmatory endpoints, the estimated treatment effects are reported together with two-sided 95% CIs and two-sided p-values.

Baseline value is defined as the eligible measurement associated with the randomisation visit, if this measurement is taken before or at the date of first dose. If a randomisation assessment is missing or if it is taken after the date of first dose, then the assessment from screening is used as the baseline assessment, if available.
If more than one measurement is associated with the same visit, the earliest measurement is considered eligible.

Missing data are defined as data that are planned to be collected and could have been collected but are not present in the database. This implies that data that are structurally missing due to death or administrative censoring are not considered missing. Unless explicitly stated, unobserved data pertaining to subjects who are lost to follow-up or withdrawn and would not have been administratively censored at the time point in question are considered missing, irrespectively of vital status as collected at end of trial.

Assessments of eGFR taken after initiation of chronic renal replacement therapy will not be used for analyses or summary tables.

2.2.1 Definition of analysis set

The full analysis set (FAS) is defined as all unique randomised subjects and grouped according to the treatment assigned at randomisation.

If a subject is randomised more than once, only the subject ID and treatment corresponding to the first randomisation will be included in FAS. The additional randomised subject IDs will be excluded from FAS. The list of subject ID's to exclude will be part of the DBL minutes.

2.2.2 Definition of observations period

A trial completer is defined as a subject who either attends the follow-up visit or who dies while active in the trial.

A patient is considered lost to follow-up (LTFU) if the patient does not complete the trial and does not withdraw consent. The date and status for LTFU are determined by investigator at trial completion, either following interim testing or after accrual of the total planned number of primary endpoint events.

In-trial observation period

The in-trial observation period for a subject is defined as the period from date of randomisation to the first of (both inclusive):

- date of follow-up visit
- date when subject withdrew consent
- date of last contact with subject for subjects who are LTFU
- date of death

On-treatment observation period

A time-point in the in-trial observation period is considered as on-treatment if any dose of trial product has been administered within the previous 5 weeks (35 days). The on-treatment observation period is defined as all times which are considered on-treatment and may consist of several time intervals with gaps between.

First on-treatment observation period

The first on-treatment observation period is defined as the on-treatment observation period until first time being off treatment for 5 consecutive weeks (35 days). Thus it is the first time interval in the on-treatment period.

2.2.3 Estimands

Primary estimand (intention-to-treat)

The estimand for all objectives is an intention-to-treat estimand, evaluating the effect of the randomised treatment intervention irrespective of adherence to treatment and changes to background medication. The estimand is addressed using FAS and the in-trial observation period.

Secondary estimand (on treatment)

This estimand covers the primary and confirmatory secondary endpoints and is evaluating the effect of the randomised treatment intervention in all randomised subjects had they remained on their randomised treatment for the entire trial. The estimand is addressed using FAS and the first on-treatment observation period.

2.2.4 Intercurrent events

Intercurrent events, including but not limited to events of or associated with:

- randomised treatment adherence
- change in background medication modifying CV risk
- initiation of chronic renal replacement therapy
- withdrawal
- lost-to follow up
- death (if not part of endpoint)

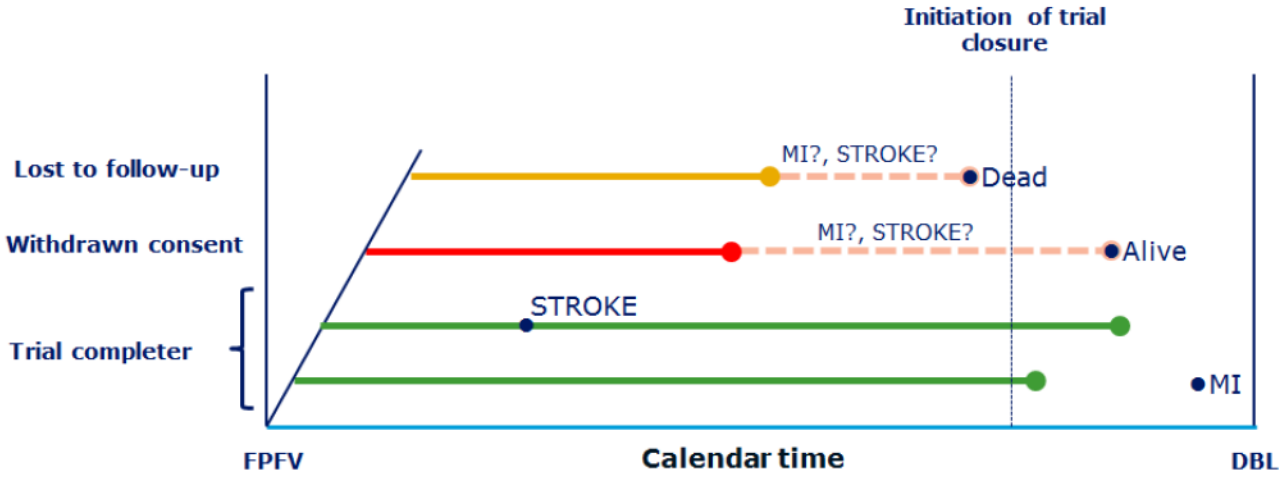
These are reported using descriptive statistics. Handling of intercurrent events for the statistical analyses of the confirmatory endpoints is described in Table 4.

2.2.5 Time-to-event endpoints, censoring and competing risks

Time-to-event endpoints are in general time-to-first-event endpoints but will for simplicity be denoted time-to-event (TTE) endpoints.

If adjudicated, the TTE endpoints are defined based on outcomes of the EAC evaluations. While vital status is ascertained systematically throughout the trial until DBL, other event types cannot be systematically collected after withdrawal, lost-to-follow-up, or after end-of-trial visit as illustrated in Figure 3. For this reason, any event occurring after the in-trial observation period is not included in analyses, unless otherwise stated.

Figure 3 In-trial observation periods used in analysis of time-to-event endpoints reflected by bold limes for four different subjects examples

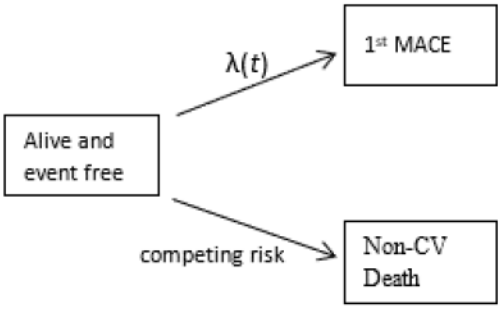


If a subject experiences the event of interest during the in-trial observation period, the observation of the TTE is the time from randomisation to the date of event.

The observation of the TTE is censored if the event of interest does not happen during the in-trial observation period and if the subject is still alive at the end of the observation period. The general assumption for censored observations is that the risk of experiencing an event is not changed by censoring, i.e. an assumption of independent censoring. This is a reasonable assumption for administrative censoring at end-of-trial visit but may not be for subjects withdrawing or subjects lost to follow-up. Sensitivity analysis addressing the assumption of independent censoring is planned for the primary endpoint, see section 2.3.2.

The observation of the TTE is terminated if the event of interest does not happen before the death of the patient unless death is part of the endpoint. Terminating events (competing risks) is potentially present for all TTE endpoints except for all-cause death; for the primary endpoint, non-CV death is a competing risk terminating the observation for the event of interest (MACE). Figure 4 illustrates competing risk as a multi-state model for the primary endpoint. The hazard rate of interest in this trial is denoted by $\lambda(t)$, t being time since randomisation.

Figure 4 Multi-state model illustrating competing risk for primary endpoint



Unless otherwise specified, the statistical analyses of TTE endpoints are done by using a Cox proportional hazards model with treatment group (oral semaglutide, placebo) as fixed factor under the assumption of independent censoring. Terminated observations (due to competing risks) are technically treated as censored observations but are not part of the independent censoring assumption. The population-level summary measure for TTE endpoints is the HR for oral semaglutide versus placebo. The assumption of proportional hazards is investigated by residuals. Tied event times are handled using the exact method and confidence intervals are based on the profile likelihood.

Cumulative incidence functions for TTE endpoints are estimated by the Aalen-Johansen estimator which accounts for competing risks.

Table 5 provides an overview of the TTE endpoints including any competing risk and whether the TTE endpoint is EAC-confirmed.

2.2.6 Continuous and binary endpoints

The population-level summary measure for continuous endpoints is the mean difference for oral semaglutide versus placebo. The population-level summary measure for binary endpoints is the odds ratio for oral semaglutide versus placebo.

2.3 Primary endpoint

Time from randomisation to first occurrence of a composite MACE endpoint consisting of

- CV death
- Non-fatal myocardial infarction (MI)
- Non-fatal stroke.

Fatal MI is defined as an EAC-confirmed MI occurring within (≤) 30 days of an EAC-confirmed CV death classified as cause of death being MI. All other MIs are defined as non-fatal. A similar definition is applied for fatal/non-fatal stroke, where MI is replaced with stroke.

Deaths attributed to the category "undetermined cause of death" are presumed cardiovascular death.

2.3.1 Primary analysis

The primary analysis will address the primary estimand (intention-to-treat). The HR for comparing oral semaglutide versus placebo is estimated from a Cox proportional hazards model with treatment group (oral semaglutide, placebo) as fixed factor together with the 2-sided 95% CI and one-sided

fixed design p-value for hypothesis testing. The score test from the Cox model is used for testing. The following superiority hypothesis is tested:

$$H_0: HR \geq 1.0 \text{ against } H_a: HR < 1.0.$$

Superiority of oral semaglutide versus placebo is considered confirmed if the associated H_0 is rejected. The nominal significance level is calculated using the alpha spending function and the actual observed number of events available for the analysis. Final inference on termination is adjusted for the group sequential design by using the likelihood ratio ordering.

Competing risk from non-CV death is handled as censorings in the Cox analysis as described in section 2.2.5. Please, refer to Table 4 for handling of other intercurrent events.

2.3.2 Sensitivity analyses

If superiority is established for the primary endpoint, the following sensitivity analysis is performed. The primary analysis assumes independent censoring for patients who have withdrawn consent or are lost to follow-up. To investigate the impact of this assumption on the primary analysis, a 2-way tipping point analysis based on the approach described in Zhao et al[1] is performed. In this analysis, subjects in the two treatment groups who have withdrawn or are lost to follow-up will have event times imputed from the conditional event distribution with a penalty in the sense that the risk (hazard) of MACE is changed following censoring compared to while under observation. Multiple imputed data sets are analysed with separate Cox regressions and results are combined using Rubin's rule. The tipping points are then defined as the combination of penalties (in each of the treatment groups) needed to turn around the superiority conclusion.

Two additional sensitivity analyses will be performed by multiple imputation of event times for subjects who are withdrawn or lost to follow-up. If the imputed event time occurs after the subjects planned end-of-trial time the subject will be censored at the planned end-of-trial time

The first will be done by treatment arm using an estimated annual event rate from subjects who discontinue treatment permanently but remain in the trial. The event rate will be based on events and time while these subjects are permanently off-treatment. A time-point in the in-trial observation period is considered as belonging to the permanently off-treatment period if any dose of trial product has been administered more than 5 weeks (35 days) ago and the subject remains off-treatment for the remainder of the trial. This analysis condition on the future in the sense that subjects are only known to be permanently off treatment by the end of the trial.

The second analysis avoids conditioning on the future by using an estimated annual event rate for subjects who discontinue treatment at any point in the trial. The imputations are done by treatment arm. The event rate will be based on events occurring from the first time subjects are off treatment corresponding to when their first on-treatment period ends (section 2.2.2) and until end of the in-trial observation period. This may include time periods where the subjects actually went back on trial treatment.

Technically, the first of the two sensitivity analyses will be performed in the following steps:

1. For the purpose of estimating the off-treatment event rates, a set of retrieved dropouts are selected. The selection criteria are that the subject shall be a trial completer, have their date of last dose during the trial reported as a treatment discontinuation, have ended the on-treatment observation period before the end of the in-trial observation period, and not having had an event before the end of the on-treatment observation period. For each selected subject, the off-treatment event time is calculated from a start date set to the day after the end of the on-treatment observation period. The event time is considered censored at the end of the in-trial observation period.
2. The off-treatment event time data are fitted within treatment arms to an exponential distribution using Bayesian analysis and accounting for censoring. A noninformative prior distribution is used for the rate parameter in each treatment arm. 500 replicates of the two off-treatment event rates are then randomly sampled from the posterior distribution.
3. To prepare the imputation, 500 copies of the original data set are created and linked to the corresponding replicate of the off-treatment event rates. For each subject who is censored due to withdrawal or being lost to follow up, the event time is imputed by adding a random variable to the original censoring date. The random variables are generated from an exponential distribution using the off-treatment event rate for the corresponding replicate and treatment arm, and rounded up to whole days. If the imputed event time lies beyond the planned date of end of trial for the subject, it is considered censored at this date. There will now be 500 complete data sets.
4. Each complete data set is analysed using the same Cox regression as in the primary analysis. The analysis gives the estimated log hazard ratio and associated standard error.
5. The log hazard ratios and standard errors from the 500 data sets are pooled using Rubin's rule to obtain a single point estimate, confidence interval and p-value.

The procedure for the second sensitivity analysis is identical to the first analysis except for step 1. The selection criteria for a retrieved dropout are instead that the subject shall have their date of last dose during the first on-treatment observation period reported as a treatment discontinuation, have ended the first on-treatment observation period before the end of the in-trial observation period and not having had an event during the first on-treatment observation period. The selection may include subjects who are later withdrawn or lost to follow-up. The off-treatment event time is calculated from a start date set to the day after the end of the first on-treatment observation period.

2.3.3 Supplementary analyses

The following supplementary analyses are planned:

- Absolute risk difference: Estimation of the absolute risk difference (and 95% CI) at year 3 between oral semaglutide and placebo based on the Aalen-Johansen estimator for the cumulative incidence functions for each treatment group. If the trial is stopped early for superiority, year 2 will be used.
- On-treatment: Analysis addressing the secondary estimand using a Cox proportional hazards model using the first on-treatment observation period.
- Additionally, an analysis of non-CV death using the same Cox model as for the primary endpoint will be done to evaluate the influence of the competing risk non-CV-death on the primary results.

2.3.4 Subgroup analyses

The consistency in the treatment effect for the primary endpoint is explored by subgroup analyses based on the below baseline information:

- Sex: Female, Male
- Age < 65 years (yes/no)
- Age: <65, 65≤ to <75, ≥75 years
- Race: White, Black or African-American, Asian, Other
- Ethnicity: Hispanic/Latino, Not Hispanic or Latino
- $HbA_{1c} \leq 8.0\%$ (yes/no)
- BMI ≤ 30 kg/m^2 (yes/no)
- Established CV disease only, Chronic kidney disease only, both
- Prior MI or stroke (yes/no)
- Metformin use (yes/no)
- Insulin use (yes/no)
- SGLT-2i use (yes/no)
- Chronic heart failure (yes/no)
- eGFR < 60 ml/min/1.72 m^2 per CKD-EPI (yes/no)
- eGFR: < 30, 30 ≤ to < 45, 45 ≤ to < 60, ≥ 60 mL/min/1.73m^2 (CKD-EPI)
- Peripheral artery disease (yes/no)
- Region: Europe, North America, Asia, Other. The regions are defined as
    - **Europe**: Austria, Belgium, Croatia, Czech Republic, Denmark, Finland, France, Germany, Italy, Netherlands, Romania, Slovakia, Spain, United Kingdom
    - **North America**: Canada, United States
    - **Asia**: China, Hong Kong, India, Israel, Japan, Malaysia, South Korea, Taiwan, Thailand, Turkey
    - **Other**: Argentina, Brazil, Colombia, Mexico, Russia, Serbia, South Africa, Ukraine

The subgroup analyses are based on Cox proportional hazards models with an interaction between treatment group (oral semaglutide, placebo) and the specific subgroup as a factor.

2.3.5 Supplementary analyses evaluating impact of the COVID-19 pandemic

The following supplementary analyses will be made to assess the impact of the COVID-19 pandemic on the primary endpoint. The analyses address two scenarios: one where MACEs are impacted by an increased MACE rate and potentially a different treatment effect for events occurring concurrently with COVID-19 infection; the other with a reduced MACE rate due to concurrent COVID-19 infection leading to fewer CV deaths as the subjects die (prematurely) of COVID-19 infection and not their underlying atherosclerotic disease.

- Time from randomisation to first MACE without concurrent COVID-19 SAE. The definition of MACE is modified so any MACE occurring concurrently with a COVID-19 SAE in a subject is not considered a MACE. The observation period and censoring are not changed. Any subsequent MACE can then qualify to be the first MACE for the subject.
- Time from randomisation to first MACE without concurrent COVID-19 AE. The definition of MACE is modified so any MACE occurring concurrently with a COVID-19 AE in a

subject is not considered a MACE. The observation period and censoring are not changed. Any subsequent MACE can then qualify to be the first MACE for the subject.

- Time from randomisation to first MACE or non-CV death occurring concurrently with a COVID-19 SAE. The definition of MACE is modified to include non-CV deaths potentially related to COVID-19. The observation period and censoring are not changed.

The analyses will be done with the same Cox regression model as for the primary analysis. An event is considered concurring with a COVID-19 AE if the event occurs in the time period from the start day of the COVID-19 AE and until 30 days after the last of the following two dates: the stop date of the COVID-19 AE or the end of hospitalisation date for a hospitalisation reported together with the COVID-19 AE

2.3.6 Supplementary analyses evaluating impact of co-participation in COVID-19 treatment or prevention trials

To assess the potential impact on the primary analysis of subjects being allowed to co-participate in trials with primary objective of evaluating an approved or non-approved investigational medical product for treatment or prevention of COVID-19 disease the following supplementary analysis will be done: An analysis of time to first MACE where all subjects co-participating in a COVID-19 treatment or prevention trial are censored at the day they receive the first trial treatment for preventing or treating COVID-19. This will reduce the observation time.
The analysis will be done with the same Cox regression model as for the primary analysis.
The analysis corresponds to the situation where patients withdraw from the trial when they start co-participation. If less than 10 subjects have co-participated in COVID-19 treatment or prevention trials then this analysis will not be performed.

2.4 Secondary endpoints

Confirmatory secondary endpoints are analysed under multiplicity control.

2.4.1 Confirmatory secondary endpoints

If superiority is established for the primary endpoint, the superiority hypothesis stated in section 2.3.1 is tested for each of the confirmatory secondary endpoints under multiplicity control via a stagewise hierarchical testing scheme using the below order:

- Time from randomisation to first occurrence of 5-component composite CKD endpoint
- Time from randomisation to CV death
- Time from randomisation to first occurrence of MALE endpoint

For the type I error to be strongly controlled at one-sided level of 2.5% (Glimm et al[2]) the same alpha-spending function as for the primary endpoint (section 2.1) is used for the confirmatory secondary endpoints.
No adjustments of results for the confirmatory secondary endpoints due to the group sequential design will be done.

Table 2 provides an example of the nominal significance levels at interim and scheduled termination when the interim testing is conducted at exactly 2/3 of the planned number of primary endpoint events and where the number of events for the secondary endpoint at scheduled termination is 3/2 times the number of secondary endpoint events at the interim.

The actual nominal significance level will be based on the exact number of events available at the interim analysis.

Table 2 Nominal significance level for confirmatory secondary endpoints at interim and scheduled termination – an example

Stopping boundary scale	Interim	Scheduled termination
Nominal significance level	0.00605	0.02314

2.4.1.1 5-component composite CKD endpoint

Time from randomisation to first occurrence of a composite endpoint consisting of:

- Onset of persistent ≥ 50% reduction in eGFR (CKD-EPI) compared with baseline
- Onset of persistent eGFR (CKD-EPI) < 15 mL/min/1.73 m2
- Initiation of chronic renal replacement therapy (dialysis or kidney transplantation)
- Renal death
- CV death

The 5-component composite CKD endpoint is analysed using a Cox proportional hazards model as for the primary endpoint and addressing the primary estimand.

For the eGFR components, a persistent outcome in eGFR is defined as having two consecutive central laboratory assessments at least 4 weeks apart meeting the criteria. When classifying the events based on consecutive laboratory assessments, the date of the event is the date of the first sample meeting the definition.

If eGFR at baseline is persistent below 15 ml/min/1.73m2 the subject will still be included in the analysis as the subject still can experience one of the other four component events.

When classifying chronic renal replacement therapy or kidney transplantation, the date of event is the date of initiation of the therapy or surgery, respectively.

Missing data for eGFR values due to e.g. missing blood samples while patients are still being followed are not imputed, assuming no eGFR component events observed during in-trial observation period with missing eGFR values.

In the case that only a single eGFR value fulfils the criteria of ≥50% reduction in eGFR compared with baseline or eGFR <15 ml/min/1.73 m^2 Table 3 provides data handling rules for defining endpoint events. Any eGFR assessment made after initiation of chronic renal replacement therapy will not qualify as a confirmatory eGFR value.

Table 3 Data handling rules where only one initiating eGFR measurement meets the criteria of ≥50% eGFR reduction or eGFR <15 ml/min/1.73 m2

Rule number	Event: One eGFR value fulfilling the criteria * with a subsequent event without any confirmatory eGFR value measured >=4 weeks after the first eGFR measurement being available.	eGFR component	Date of event
1	CV or renal death	No	Date of death
2	Non-CV and non-renal death	No	Not applicable
3	Initiation of chronic renal replacement therapy	No	Date of initiation
4	Lost-to-follow-up or withdrawal of consent	No	Not applicable
5	One eGFR value fulfilling the criteria* at planned end-of-treatment visit	No	Not applicable

* ≥50% reduction in eGFR compared with baseline or eGFR <15 ml/min/1.73 m^2

A persistent outcome in eGFR in the in-trial period is considered on-treatment if the first of the two consecutive measurements falls within the on-treatment period, irrespective of whether the confirmatory eGFR value falls within the on-treatment period or not.

2.4.1.2 CV death

Time from randomisation to CV death is analysed using a Cox proportional hazards model as described for the primary endpoint and addressing the primary estimand.

2.4.1.3 MALE endpoint

Time from randomisation to first occurrence of MALE, a composite endpoint consisting of

- acute limb ischemia hospitalisation
- chronic limb ischemia hospitalisation

The MALE endpoint is analysed using a Cox proportional hazards model as described for the primary endpoint and addressing the primary estimand.

2.4.1.4 Sensitivity analyses

The two sensitivity analyses with imputation from subjects off treatment will be done for all confirmatory secondary endpoints (section 2.3.2).

Sensitivity analysis for the 5-component composite CKD endpoint

In the analysis for the 5-component composite CKD endpoint, missing data for scheduled central laboratory eGFR values are not imputed. The following sensitivity analysis using multiple imputation is planned. Prior to analysis, missing data are imputed using multiple imputation generating 500 data sets to account for the inherent uncertainty. The imputation is performed separately for each treatment group. In the first step, intermittent missing values are imputed using the Markov Chain Monte Carlo method based on an assumption of multivariate normality. In the second step, imputation of monotone missing values is done within subject groups defined by the treatment group and based on a sequential univariate regression approach. At each scheduled visit starting with the first post-baseline visit the imputation model includes baseline eGFR value and the previous post-baseline scheduled values (observed and imputed) prior to the visit being imputed as covariates. For each eGFR component of the 5-component composite CKD endpoint it is evaluated whether an eGFR event has occurred (yes/no) within the in-trial observation period. Intermittent imputed data are excluded from this evaluation.

After imputation of missing eGFR data, the 5-component composite CKD endpoint is derived, and the 500 multiple-imputed data sets are analysed with the primary Cox proportional hazards model described above. Patients that do not experience an event during the in-trial observation period are censored at the in-trial observation period end date. The resulting estimates of the log(HR) are combined using the methods of Rubin and back transformed to HR scale to draw inference.

2.4.1.5 Supplementary analyses

Supplementary analyses described in sections 2.3.3 for the primary endpoint, will similarly be done for the confirmatory secondary endpoints.

Supplementary analyses described in section 2.3.5 with regards to the impact of the COVID-19 pandemic and section 2.3.6 with regards to the potential impact of co-participation in COVID-19 trials will be done for the primary endpoint and the confirmatory secondary endpoints.

Supplementary analysis using CKD-EPI equation not including race

The equation used for estimating eGFR (CKD-EPI) for the confirmatory secondary endpoint, component composite CKD endpoint, incorporates age, sex and race. To fight racial bias in clinical trials a new equation where race is not incorporated has been introduced, as described in Inker et al. (2021)[3]. A supplementary analysis will be made using this new equation where race is not included to estimate eGFR using a Cox proportional hazards model as for the primary endpoint and addressing the primary estimand.

2.4.2 Supportive secondary endpoints

2.4.2.1 Time-to-event supportive secondary endpoints

Each of the TTE supportive secondary endpoints are analysed using the same Cox proportional hazards model as described for the primary endpoint and addressing the primary estimand. Competing risks for the relevant TTE endpoints can be seen in Table 5.

For the two time-to-event endpoints "Onset of persistent ≥50% reduction in eGFR" and "Onset of persistent eGFR <15 mL/min/1.73 m^2" an event of initiation of chronic renal replacement therapy acts as competing risk. For these two endpoints subject will be censored in the Cox model at time of initiation of chronic renal replacement therapy. In case of one eGFR measurement fulfilling the criteria and no available confirmatory test, the data handling rules are as described for the confirmatory secondary CKD endpoint in section 2.4.1.1

For the time-to-events endpoints involving eGFR, the analyses will exclude patients who already have met relevant renal components at baseline, except for the composite CKD endpoint where if eGFR at baseline is persistent below 15 ml/min/1.73m2 the subject will still be included in the analysis as the subject still can experience one of the other four component events.

For the endpoint "Onset of persistent ≥50% reduction in eGFR" only subjects with a valid baseline value will be included in the analysis.

Supplementary analyses

All-cause death: In addition, all-cause death is analysed using FAS and an extended in-trial observation period including the follow-up for vital status for subjects who withdraw consent or are LFTU. The relative risk for the binary endpoint (death/alive) will be compared between the two treatment groups using the likelihood ratio method. The model is chosen because it doesn't depend on the observation time which is only extended for subjects withdrawn or LTFU.

5-component MACE and 2-component HF: For the two endpoints 5-component expanded MACE and 2-component HF supplementary analyses will be done by replacing the CV death component with all-cause death.

MI and stroke: In the analyses of MI and stroke supplementary analyses including fatal MI and fatal stroke are performed. Thus, the supplementary analysis will analyse endpoints defined as:

- Time from randomisation to first MI (fatal or non-fatal)
- Time from randomisation to first stroke (fatal or non-fatal)

2.4.2.2 Continuous supportive secondary endpoints

The continuous supportive secondary endpoints (change from baseline to 2 years) are analysed using multiple imputation for missing values.
An imputation model (linear regression) is estimated separately for each treatment group including baseline value as a covariate and fitted to subjects having an observed data point (irrespective of adherence to randomised treatment) at year 2. Subjects without a baseline measurement will not be a part of the model. The fitted model is used to impute values for all subjects with missing data at 2 years to create 500 complete data sets. The complete data sets are analysed by an ANCOVA with treatment as fixed factor and baseline value as covariate. Rubin's rule is used to combine the results.

Additionally, change from baseline in hsCRP at year 2 is analysed using the above model. hsCRP is logarithmic transformed and the mean difference on the logarithmic scale is back-transformed to original scale and reported as geometric mean ratio.

2.4.2.3 Annual rate of change in eGFR

The annual rate (slope) of change in eGFR is compared between treatment groups based on a linear random regression model on eGFR values with treatment, time (as a continuous variable) and treatment time interaction as fixed effects, and including subject effect as a random intercept and time as a random slope. The random intercept and slope are assumed to be bivariate normal distributed with mean zero and an unstructured covariance matrix. The independent error term is assumed to be identical univariate normal distributed with mean zero. The model is fitted to observed scheduled eGFR data at baseline and post-baseline. The parameter of interest is the regression coefficient for the treatment and time interaction term, which measures the slope difference between oral semaglutide and placebo.

2.4.2.4 Analysis of recurrent events

Repeated occurrence of the same type of event over time for the same patient may happen. The following recurrent event endpoints are compared between the treatment arms from randomisation to end of trial:

- Number of MACE events
- Number of MI events (fatal or non-fatal)
- Number of stroke events (fatal or non-fatal)
- Number of MI and stroke events (fatal or non-fatal)
- Number of heart failure events (heart failure requiring hospitalisation or urgent heart failure visit)

Mean number of events is plotted as a function of study time and analysed using a marginal mean regression model for recurrent events accounting for competing risk of dying as described by Ghosh & Lin[4, 5]. Treatment effect is reported as a mean ratio and corresponding 95% robust CI to account for the dependency of within-subject of recurrent events.

2.4.2.5 Number of severe hypoglycaemic episodes

Number of severe hypoglycaemic episodes is analysed using the same marginal recurrent event regression model described above.

2.5 Exploratory endpoints

Change from baseline to year 2 and 3 in cognitive function: Montreal Cognitive Assessment (MoCA) score is compared between oral semaglutide and placebo using the same analysis as outlined above for continuous endpoints. If the trial is terminated earlier than planned (stopping after interim testing) change from baseline to 3 years may not be evaluated.
Refer to Table 8 for more details on the MoCA score.

The smoking endpoint is analysed at year 2 using a logistic regression model with treatment (oral semaglutide, placebo) and baseline smoking status (yes/no) as fixed factors. Missing data are handled by multiple imputation. The imputation model (logistic regression) is done separately for each treatment arm and includes baseline smoking status as a fixed factor and fitted to subjects having an observed data point (irrespective of adherence to randomised treatment) at year 2. The fitted model is used to impute values for all subjects with missing data (see section 2.2) at year 2.

Change from baseline to year 2 for the PROTECT endpoints are compared between oral semaglutide and placebo using the same analysis as outlined above for continuous endpoints.
Refer to Table 9 and Table 10 for more details on the PROTECT endpoints.

Table 7 gives an overview of planned analyses for all endpoints.

2.6 Other assessments

All systematically collected AEs, i.e. serious AEs and non-serious events requiring additional data collection as well as COVID-19 related AEs are summarised as number of subjects with events, proportion of subjects with events, number of events and rate of events according to treatment group. Summaries of SAEs are categorised by severity, relation to treatment, and outcome.

3 Interim testing

The trial design includes *one* pre-planned interim testing for superiority of the primary endpoint. The planned timing is when 817 events (two thirds of the planned total events) of the primary endpoint have been accrued. The interim testing is performed based on a locked snapshot of the study database. The date of the snapshot defines the interim analysis cut-off date for the interim testing.
Subjects without an EAC-confirmed primary endpoint event prior to the date of analysis cut-off are considered censored with the censoring date defined as the first of:

- in-trial observation period end-date
- analysis cut-off date

The same Cox model as described in section 2.3 is used for the interim testing addressing the primary estimand.

3.1 Role of DMC

Blinded and un-blinded data analyses during trial conduct are evaluated by the DMC, as described in the DMC charter. Trial integrity is ensured by using an external independent statistical service provider (independent of trial conduct and external to Novo Nordisk) to prepare these data and analyses for the DMC.

The DMC will evaluate the interim result and make recommendation to terminate the trial early for superiority if appropriate. The DMC evaluates the un-blinded interim results using the group sequential stopping boundary as guidance. Stopping the trial early for superiority is only allowed if the stopping boundary is crossed and the DMC makes the decision to recommend early trial termination based on this and other considerations as specified in the DMC charter.

Recommendations from the DMC back to Novo Nordisk and any other party will exclude any details of the interim results as to maintaining trial integrity.

3.2 Stopping boundary for superiority at interim

The exact number of primary endpoint events used for the interim testing is only known at the time of analysis, and the exact boundary is re-calculated using the Lan-DeMets alpha spending function.

3.3 Analysis on termination

If the trial is terminated early for superiority following the interim testing, definitive evaluation of superiority for the primary endpoint is performed based on all the available data at the end-of-trial, including overrun data. Overrun data include events happening between the cut-off date for the DMC interim testing and end-of-trial as well as additional confirmed events that were undergoing adjudication at the analysis cut-off time point. If the trial is not terminated early for superiority following the interim testing, the analysis at scheduled termination is performed when the planned number of 1,225 events has been accrued. The exact number of primary endpoint events used for the analysis on termination is only known at the time of analysis, and nominal significance level is updated based on the exact number of total accrued events and the Lan-DeMets alpha spending function. Similarly, the significance levels for the confirmatory secondary endpoints are updated

based on the exact number of events and all available data at end-of-trial are used for analyses of both secondary and exploratory endpoints.

For reporting of results for the primary endpoint (p-value, HR and 95% CI), the analysis on termination (either early or at scheduled termination) are adjusted for the group sequential design using the likelihood ratio ordering.

4 Changes to the statistical analyses planned in the protocol

In general, this SAP describes in more details the statistical analyses planned in the protocol.

SAP version 1.0 dated 12-MAR-2019

For the confirmatory secondary endpoint 5-component composite CKD the primary analysis is changed not to include multiple imputation for missing values of eGFR. Instead multiple imputation is done in a sensitivity analysis (section 2.4.1).

Sensitivity analysis for the primary endpoint updated to a 2-way tipping point analysis (section 2.3.2).

Additional analyses added in the SAP and not described in the protocol (section 2.3.3 and 2.4.2):

- Supplementary analyses for primary endpoint and confirmatory secondary endpoints:
  - Risk difference at year 3
  - Secondary estimand (on-treatment)
- Supplementary analyses for 5-component MACE and 2-component HF by replacing CV death component with all-cause death
- Analyses of MI and stroke using both fatal and non-fatal events
- Analysis of all-cause death using an extended in-trial observation period
- Recurrent event analyses for selected endpoints

SAP version 2.0 dated 14-DEC-2022

- Author has been deleted from front page
- It has been clarified that the independent statistician confirming the confirmatory analyses will also be blinded until DBL.
- The definition of baseline value has been updated, see Section 2.2.
- It has been clarified how missing data from subjects who are withdrawn or lost to follow-up are handled, see Section 2.2.
- Assessments of eGFR taken after initiation of chronic renal replacement therapy will not be used for analyses or summary tables (section 2.2)
- It has been clarified how subjects randomised more than once will contribute to the FAS, see Section 2.2.1.
- A definition of the first on-treatment observation period has been added, see Section 2.2.2. The use of it has been added in Sections 2.2.3 and 2.3.3.
- Figure 3 has been added illustrating the in-trial observation period used in analysis of time-to-event endpoints (section 2.2.5)
- It has been clarified that tied event times are handled using the exact method and risk limits are based on the profile likelihood, see section 2.2.5

- It has been specified that the absolute risk difference analysis will be based on the Aalen-Johansen estimator (section 2.3.3).

- Two sensitivity analyses have been added. The analyses impute event times for those that are lost to follow-up or withdrawn, based on the observed event rates for subjects that either go permanently off treatment (first analysis) or at least go off treatment once. (Section 2.3.2 and 2.4.1.5)

- Analyses added in the SAP version 2 that is not described in the protocol

  - Analysis of change from baseline in hsCRP, see Section 2.4.2.2.

  - Subgroup analysis based on peripheral artery disease (yes/no); age: <65, 65≤ to <75, ≥75 years; eGFR: < 30, 30 ≤ to < 45, 45 ≤ to < 60, ≥ 60 mL/min/1.73m^2 (CKD-EPI); EU has been changed to Europe and it has been clarified which region each country belongs to see Section 2.3.4.

  - A supplementary analysis of non-CV death has been added to evaluate the effect of competing risk on the primary endpoint, see Section 2.3.3.

  - Supplementary analyses evaluating the impact of the COVID-19 pandemic on the primary and confirmatory secondary endpoints have been added (section 2.3.5. 2.4.1.5).

  - Supplementary analyses evaluating the impact of co-participation in COVID-19 trials on the confirmatory endpoints have been added (section 2.3.6, 2.4.1.5).

  - A supplementary analysis using a new equation (CKD-EPI) where race is not included to estimate eGFR has been added (section 2.4.1.5)

  - Recurrent event analysis for MACE (section 2.4.2.4)

- For confirmatory secondary endpoints, the significance level used in the hierarchical testing procedure has been changed to follow a alpha spending function in order to control the type 1 error, see Section 2.4.1.

- The definition of eGFR components of confirmatory secondary endpoint in case of competing events (section 2.4.1.1) has been adjusted and a description of on-treatment for eGFR components has been added.

- The supplementary analysis for all cause death has been changed to be a comparison of relative risks for the binary endpoint death/alive (section 2.4.2.1).

- Clarification as to how the ANCOVA for continuous supportive secondary endpoints should be done has been added (section 2.4.2.2).

- For the analysis of MOCA it has been clarified when year 3 will be used (section 2.5)

- For the analysis of smoking, details of the imputation model has been added (section 2.5)

- Section 2.6 has been added
- It has been added that the significance level for the confirmatory secondary endpoints will be updated based on the exact number of events and that all available data at end-of-trial are used for both secondary and exploratory endpoints (section 3.3).
- The part describing that Novo Nordisk will replicate the interim analysis has been removed (section 3.3).
- Details on the endpoints MoCA and PROTECT have been added in Appendix 6.5 6.6.
- In general, editorial changes for alignment with SAPs for trials EX9536-4388 (SELECT) and NN9535-4321 (FLOW) has been made.

5 References

1. Zhao Y, Herring AH, Zhou H, Ali MW, Koch GG. A multiple imputation method for sensitivity analyses of time-to-event data with possibly informative censoring. J Biopharm Stat. 2014;24(2):229-53.
2. Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. Stat Med. 2010;29(2):219-28.
3. Inker LA, Eneanya ND, Coresh J, Tighiouart H, Wang D, Sang Y, et al. New creatinine- and cystatin C-based equations to estimate GFR without race. N Engl J Med. 2021;385(19):1737-49.
4. Ghosh D, Lin DY. Nonparametric analysis of recurrent events and death. Biometrics. 2000;56(2):554-62.
5. Ghosh D, Lin D. Marginal regression models for recurrent and terminal events. Statistica Sinica. 2002;12:663-88.

6 Appendix

6.1 Handling of intercurrent events for the confirmatory endpoints

Table 4 Handling of intercurrent events for the confirmatory endpoints

Endpoint	Intercurrent event	Handling
Time to first occurrence of MACE	Treatment discontinuations Medication modifying cardio-renal risk Initiation of chronic renal replacement therapy	Events and follow-up time are collected after intercurrent events and used in the analysis
	Trial discontinuation (withdrawal of consent or lost-to follow-up)	Censoring at time of trial discontinuation
	Non-CV death (competing risk)	Censoring at time of non-CV death in the Cox model
Time to first occurrence of composite CKD	Treatment discontinuations Medication modifying cardio-renal risk	Events and follow-up time are collected after intercurrent events and used in the analysis
	Trial discontinuation (withdrawal of consent or lost-to follow-up)	Censoring at time of trial discontinuation
	Non-renal and Non-CV death (competing risk)	Censoring at time of non-renal or non-CV death in the Cox model
Time to occurrence of CV death	Treatment discontinuations Medication modifying cardio-renal risk Initiation of chronic renal replacement therapy	Events and follow-up time are collected after intercurrent events and used in the analysis
	Trial discontinuation (withdrawal of consent or lost-to follow-up)	Censoring at time of trial discontinuation
	Non-CV death (competing risk)	Censoring at time of non-CV death in the Cox model
Time to first occurrence of MALE	Treatment discontinuations Medication modifying cardio-renal risk Initiation of chronic renal replacement therapy	Events and follow-up time are collected after intercurrent events and used in the analysis
	Trial discontinuation (withdrawal of consent or lost-to follow-up)	Censoring at time of trial discontinuation
	All cause death (competing risk)	Censoring at time of death in the Cox model

6.2 List of time-to-event endpoints

Table 5 List of time-to-event endpoints

Endpoint	Composite order/details*	EAC**	Competing risk
Primary			
3-component MACE	- CV death - Non-fatal MI - Non-fatal stroke	Yes Yes Yes	Non-CV death
Confirmatory secondary			
5-component CKD	- Renal death - CV death - Initiation of chronic renal replacement therapy - Onset of persistent eGFR<15 ml/min/1.73 m2 - Onset of persistent ≥50% reduction in eGFR	Yes Yes Yes No No	Non-CV and non-renal death
CV death	–	Yes	Non-CV death
MALE	- Acute limb ischemia hospitalisation - Chronic limb ischemia hospitalisation	Yes Yes	All cause death
Supportive secondary			
3-component MACE with all cause death	- All-cause death - Non-fatal MI - Non-fatal stroke	Yes Yes Yes	None
5-component MACE	- CV death - Non-fatal MI - Non-fatal stroke - Coronary revascularisation - Unstable Angina hospitalisation	Yes Yes Yes No Yes	Non-CV death
2-component HF	- CV death - HHF or urgent HF Visit	Yes Yes	Non-CV death
4-component CKD	- Renal death - Initiation of chronic renal replacement therapy - Onset of persistent eGFR<15 ml/min/1.73 m2 - Onset of persistent ≥50% reduction in eGFR	Yes Yes No No	Non-renal death
All cause death	–	Yes	None
Non-fatal MI	–	Yes	All-cause death
Non-fatal stroke	–	Yes	All-cause death
HHF or urgent HF Visit	–	Yes	All-cause death
Coronary revascularisation	–	No	All-cause death
Unstable angina hospitalisation	–	Yes	All-cause death
Renal death	–	Yes	Non-renal death
Onset of persistent ≥50% reduction in eGFR	–	No	All-cause death and initiation of chronic renal replacement therapy
Onset of persistent eGFR<15	–	No	All-cause death and initiation of chronic renal replacement therapy
Initiation of chronic renal replacement therapy	–	Yes	All-cause death
Acute limb ischemia hospitalisation	–	Yes	All-cause death
Chronic limb ischemia hospitalisation	–	Yes	All-cause death
First severe hypoglycaemic episodes	–	No	All-cause death
No of severe hypoglycaemic episode	Recurrent events	No	All-cause death

* For composite endpoints this defines the hierarchy of components when reporting events contributing to a composite endpoint in the situation of ties of date of events of the components

** EAC-confirmed event

6.3 List of continuous and binary endpoints

Table 6 List of continuous and binary endpoints

Endpoint	Type	Details
Supportive secondary		
Annual rate of change in eGFR (total eGFR slope)	Continuous	From randomisation to end-of-trial
Change in HbA1c	Continuous	Change from randomisation to year 2
Change in Body weight	Continuous	Change from randomisation to year 2
Exploratory		
Change in MoCa score	Continuous	Change from randomisation to year 2
Change in MoCa score	Continuous	Change from randomisation to year 3
Smoking (yes/no)	Binary	Smoker at year 2
Change in Working Memory Index*	Continuous	Change from randomisation to year 2
Change in Verbal Reasoning*	Continuous	Change from randomisation to year 2
Change in Attentional Intensity Index*	Continuous	Change from randomisation to year 2
Change in Cognitive Reaction Time*	Continuous	Change from randomisation to year 2
Change in Sustained Attention Index*	Continuous	Change from randomisation to year 2

* Only relevant for English and/or Spanish speaking patients in Argentina, Canada, Colombia, Mexico, Spain, United Kingdom and United States

6.4 Overview of planned analyses for all endpoints

Table 7 Overview of planned analyses for all endpoints

Endpoint	**Model/method**	**Summary measure**	**Sensitivity analysis**	**Supplementary analysis**
Primary				
3-component MACE	Cox	Hazard ratio	2-way tipping point Analysis with imputations from subjects permanently off treatment Analysis with imputations from subjects off treatment	-Risk difference -First on-treatment -Subgroup analyses -Non-CV death - Excluding MACEs concurrent with COVID-19 SAE -Excluding MACEs concurrent with COVID-19 AE -Including non-CV deaths with concurrent COVID-19 SAE -censoring subjects who co-participate in COVID-19 trials
Confirmatory secondary				

	5-component CKD	Cox	Hazard ratio	Imputation of missing eGFR Analysis with imputations from subjects permanently off treatment Analysis with imputations from subjects off treatment	-Risk difference -First on-treatment - Excluding CKDs concurrent with COVID-19 SAE -Excluding CKDs concurrent with COVID-19 AE -Including non-CV, non-renal deaths with concurrent COVID-19 SAE -censoring subjects who co-participate in COVID-19 trials -Using eGFR equation where race was taken out of the equation
	CV deaths	Cox	Hazard ratio	Analysis with imputations from subjects permanently off treatment Analysis with imputations from subjects off treatment	-Risk difference -First on-treatment - Excluding CV deaths concurrent with COVID-19 SAE -Excluding CV deaths concurrent with COVID-19 AE -Including non-CV deaths with concurrent COVID-19 SAE -censoring subjects who co-participate in COVID-19 trials
	MALE	Cox	Hazard ratio	Analysis with imputations from subjects permanently off treatment Analysis with imputations from subjects off treatment	-Risk difference -First on-treatment - Excluding MALEs concurrent with COVID-19 SAE -Excluding MALEs concurrent with COVID-19 AE -Including all-cause deaths with concurrent COVID-19 SAE -censoring subjects who co-participate in COVID-19 trials
Supportive secondary					
	3-component MACE with all cause death	Cox	Hazard ratio	–	–
	5-component MACE	Cox	Hazard ratio	–	Including all-cause death
	2-component HF	Cox	Hazard ratio	–	Including all-cause death
	4-component CKD	Cox	Hazard ratio	–	

	All cause death	Cox	Hazard ratio	–	Relative risk extended in-trial period
	Non-fatal MI	Cox	Hazard ratio	–	-Including fatal MI -Recurrent events analysis (including fatal MI)
	Non-fatal stroke	Cox	Hazard ratio	–	-Including fatal stroke -Recurrent events analysis (including fatal stroke)
	MI and stroke (fatal or non-fatal)	Marginal mean regression	Mean ratio		
	HHF or urgent HF Visit	Cox	Hazard ratio	–	Recurrent events analysis
	Coronary revascularisation	Cox	Hazard ratio	–	–
	Unstable angina hospitalisation	Cox	Hazard ratio	–	–
	Renal death	Cox	Hazard ratio	–	–
	Onset of persistent ≥50% reduction in eGFR	Cox	Hazard ratio	–	–
	Onset of persistent eGFR<15	Cox	Hazard ratio	–	–
	Initiation of chronic renal replacement therapy	Cox	Hazard ratio	–	–
	Acute limb ischemia hospitalisation	Cox	Hazard ratio	–	–
	Chronic limb ischemia hospitalisation	Cox	Hazard ratio	–	–
	First severe hypoglycaemic episodes	Cox	Hazard ratio	–	–
	No of severe hypoglycaemic episodes	Marginal mean regression	Mean ratio	–	–
	Annual rate of change in eGFR (total slope)	Random regression model	Mean slope difference	–	–
	Change in HbA_{1c} to year 2	ANCOVA w MImp	Mean difference	–	–
	Change in body weight to year 2	ANCOVA w MImp	Mean difference	–	–
	Change in hsCRP at year 2*	ANCOVA w MImp, logarithmic scale	Mean difference		
Exploratory					
	Change in Montreal Cognitive Assessment score	ANCOVA w MImp	Mean difference	–	–
	Smoker at 2 years	Logistic regression and multiple imputation	Odds ratio	–	–
	Change in Working Memory Index**	ANCOVA w MImp	Mean difference	–	–
	Change in Verbal Reasoning**	ANCOVA w MImp	Mean difference	–	–
	Change in Attentional Intensity Index**	ANCOVA w MImp	Mean difference	–	–
	Change in Cognitive Reaction Time**	ANCOVA w MImp	Mean difference	–	–

	Change in Sustained Attention Index**	ANCOVA w MImp	Mean difference	–	–

ANCOVA w MImp = Analysis of covariance with multiple imputation. *Change from hsCRP at year 2 is not a defined endpoint in the protocol, but analysis is added in SAP.

** Only relevant for English and/or Spanish speaking patients in Argentina, Canada, Colombia, Mexico, Spain, United Kingdom and United States

6.5 MoCA

The MoCA is a cognitive test used for early detection of mild cognitive impairment.
The test checks different types of cognitive/thinking abilities: orientation, short-term memory, executive function/visuospatial ability, language, abstraction, animal naming, attention and clock-drawing test.
The endpoint used is the total score which is the sum of the score breakdowns specified in Table 8.
One point is added to the total score if the participant have 12 years or less of formal education (at baseline).
Lower MoCA score indicates greater cognitive impairment.

Table 8 MoCA score

Score breakdown	**Range**
Visuospatial and executive functioning	0-5
Animal naming	0-3
Attention	0-6
Language	0-3
Abstraction	0-2
Delayed recall (short-term memory)	0-5
Orientation	0-6
Total score	**0-30**

6.6 PROTECT

The PROTECT Cognitive Test Battery consists of 6 tests which produce 12 individual data points. Details can be seen in Table 9.
Based on the PROTECT tests 5 clinical endpoints are defined. More details on the 5 endpoints can be seen in Table 10.

Table 9 PROTECT tests and individual data points

Tests and data points	Unit	Outcome	Range
Self-ordered search			
Self Ordered Search	count	Total correct responses	0 to 40
Paired associate learning			
Paired Associate Learning	count	Total correct responses	0 to 16
Verbal reasoning			
Verbal reasoning score	count	Total correct responses minus total incorrect responses	-88 to 88
Simple Reaction Time			
Simple Reaction Time Median	msec	Median speed of individual correct responses	100 to 30000
Simple Reaction Time SD	msec	Standard Deviation of individual correct responses	0 to 30000
Digit vigilance			
Digit Vigilance Accuracy	%	Percentage of targets responded to within time window	0 to 100
Digit Vigilance Speed	msec	Mean speed of individual responses to targets within time window	150 to 1500
Digit Vigilance False Alarms	count	Number of responses falling outside of specified time window	0 to 999
Digit Vigilance Standard Deviation	msec	Standard Deviation of individual responses to targets within time window	0 to 1500
Choice reaction time			
Choice Reaction Time Accuracy	%	Percentage of stimuli responded to correctly	0 to 100
Choice Reaction Time Median	msec	Median speed of individual correct responses	150 to 30000
Choice Reaction Time SD	msec	Standard Deviation of individual correct responses	0 to 30000

Table 10 PROTECT endpoints

Endpoint	Derivation	Interpretation
Working memory index score	Self Ordered Search + Paired Associate Learning	• *Recalling the position of hidden objects* • A higher score indicates better cognitive function
Verbal reasoning score (count)	Verbal reasoning score	• *Understand and provide an answer* • A higher score indicates better cognitive function
Attentional intensity index (msec)	Simple Reaction Time Median + Digit Vigilance Speed + Choice Reaction Time Median	• *Intensity of concentration* • A lower score indicates better cognitive function
Cognitive reaction Time (msec)	Choice Reaction Time Median - Simple Reaction Time Median	• *Decision making time* • A lower score indicates better cognitive function
Sustained attention index (%-points)	((Choice Reaction Time Accuracy – 50) * 2 + (Digit Vigilance Accuracy* 0.45 - Digit Vigilance False Alarms) * 100/45) / 2	• *Stay focused* • A higher score indicates better cognitive function