Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT05891496
Sponsor trial ID:	NN6535-7519
Official title of study:	A randomised double-blind placebo-controlled clinical study investigating the effects of semaglutide s.c. once-weekly versus placebo on central and peripheral inflammation in participants with Alzheimer's disease
Document date*:	06 December 2024

*Document date refers to the date on which the document was most recently updated.

Statistical Analysis Plan

A randomised double-blind placebo-controlled clinical study investigating the effects of semaglutide s.c. once-weekly versus placebo on central and peripheral inflammation in participants with Alzheimer's disease

Semaglutide s.c. once-weekly

Redacted statistical analysis plan includes redaction of company confidential information.

Table of contents

Page

Table of tables

Page

Version History

This Statistical Analysis Plan (SAP) for study NN6535-7519 is based on the protocol version 4.0 dated 06Nov2024.

SAP Version	Date	Change	Rationale
1.0	06May2024	Not Applicable	Original version
2.0	xxDec2024	Modification of unit for exploratory endpoints	Alignment with protocol version 4.0
		Minor modifications to section 3.2.1	The quality of single cell samples was lower than expected giving minor modifications of the pipeline for processing of data
		Addition of QC step for proteomics data	The proteomic data requires additional QC
		Removal of one biomarker for the first CSR	Due to vendor issues data will not be available
		Description of the analyses of scTCR data added	Clarification of scTCR endpoint in endpoint table and addition of one more scTCR endpoint to aid interpretation of results
		CSF endpoint added to section 3.4	Data will be available for CSR 1
		Minor editorial changes and corrections	

List of abbreviations

AD	Alzheimer's Disease
AE	adverse event
ANCOVA	analysis of covariance
BCL	binary base call
CSF	cerebrospinal fluid
CSR	clinical study report
DBL	data base lock
DPS1	in-trial data set
DPS2	on-treatment data set
DEGs	differentially expressed genes
FAS	full analysis set
hsCRP	high sensitivity C-reactive protein
MAP	Modelling Analysis Plan
NPX	Normalized Protein Expression
PK	pharmacokinetics
QC	quality control
RMSD	Root Mean Square Distance
s.c.	subcutaneous
scRNAseq	single-cell ribonucleic acid sequencing
scTCRseq	single-cell T cell receptor sequencing
SAP	Statistical Analysis Plan
SAS	safety analysis set
SDTM	Study Data Tabulation Model
TCR	T cell receptor

TEAEs	treatment emergent adverse events
UMAP	uniform manifold approximation and projection

1 Introduction

This statistical analysis plan (SAP) is based on the protocol: A randomised double-blind placebo-controlled clinical study investigating the effects of semaglutide s.c. once-weekly versus placebo on central and peripheral inflammation in participants with Alzheimer's disease, version 4.0 (dated 06 November 2024). The statistical analyses of the co-primary endpoints presented in this SAP are identical to what is described in the protocol. Details for how to derive the co-primary endpoints have been added to this SAP. This SAP also includes a detailed description of the analyses of the secondary and exploratory endpoints, related to both efficacy and safety data.

1.1 Objectives and Endpoints

Objectives and Endpoints

Study objectives and endpoints are summarised in Table 1-1.

Table 1-1 Objectives and endpoints

Objectives	Endpoints		
Primary	**Title**	**Time frame**	**Unit**
To investigate the effect of semaglutide s.c. 1.0 mg once-weekly versus placebo on central and peripheral inflammation in participants with Alzheimer's disease	*Co-primary:*		
	Change in gene expression assessed by scRNAseq (cells in CSF)	From baseline (week 0) to visit 5 (week 12)	Number of differentially expressed genes
	Change in gene expression assessed by scRNAseq (cells in blood)	From baseline (week 0) to visit 5 (week 12)	Number of differentially expressed genes
Secondary	**Title**	**Time frame**	**Unit**
To compare the effects of semaglutide s.c. 1.0 mg once-weekly versus placebo on safety and tolerability in participants with Alzheimer's disease	*Supportive secondary:*		
	Number of treatment emergent adverse events (TEAEs)	From baseline (week 0) to visit 5 (week 12)	Number of events
To evaluate the effects of semaglutide s.c. 1.0 mg once-weekly on safety and tolerability in participants with Alzheimer's disease	Number of treatment emergent adverse events (TEAEs)	From baseline (week 0) to end of treatment (week 64)	Number of events
To evaluate the steady state pharmacokinetics of semaglutide s.c. 1.0 mg once-weekly in participants with Alzheimer's disease	Weekly average semaglutide concentration (C_{avg}) based on population PK analysis	From visit 3 (week 4) to end of treatment (week 64)	nmol/L

Objectives	Endpoints		
Exploratory	**Title**	**Time frame**	**Unit**
To explore the effect of semaglutide s.c. 1.0 mg once-weekly versus placebo on neuroinflammatory and neurodegenerative biomarkers in participants with Alzheimer's disease	*Exploratory:*		
	Change in biofluid-based (blood or CSF) neuroinflammatory and neurodegenerative biomarkers[a]	From baseline (week 0) to visit 5 (week 12)	-
To compare the effects of semaglutide s.c. 1.0 mg once-weekly versus placebo on the proteomic profiles in CSF and plasma in participants with Alzheimer's disease	Changes in proteome in CSF	From baseline (week 0) to visit 5 (week 12)	NPX, Normalized Protein eXpression
	Changes in proteome in plasma	From baseline (week 0) to visit 5 (week 12)	NPX, Normalized Protein eXpression
To compare the effects of semaglutide s.c. 1.0 mg once-weekly versus placebo on the T cell receptor profiles in CSF and plasma in participants with Alzheimer's disease	Changes in T cell clonal landscape assessed by scTCRseq (cells in CSF)	From baseline (week 0) to visit 5 (week 12)	Morisita-Horn index
	Changes in T cell clonal landscape assessed by scTCRseq (cells in blood)	From baseline (week 0) to visit 5 (week 12)	Morisita-Horn index
To measure semaglutide concentration in the CSF in participants with Alzheimer's disease	Semaglutide concentration in CSF[b]	At week 12	-

Notes: [a]Biomarkers is defined in this (SAP). [b]Analysis will be performed only if a validated reliable assay is available at the end of study.

Abbreviations: C_{avg} = average concentration; CSF = cerebrospinal fluid; PK = pharmacokinetics; s.c. = subcutaneous; scRNAseq = single-cell ribonucleic acid sequencing; scTCRseq = single-cell T cell receptor sequencing.

1.2 Study Design

For information on the study design see the protocol section 4.

A Data Base Lock (DBL) will be performed after the last randomised participant has had the opportunity to attend visit 5, which is the end of Study intervention period 1, the double-blind phase, of the study. Data up to and including visit 5 will be included in the first DBL for the first Clinical Study Report (CSR). During the study conduct of Study intervention period 1, study participants and investigators will remain blinded to treatment allocation. Novo Nordisk personnel involved in data handling will remain blinded until after the first DBL. An additional DBL will be performed after the end of the follow up period following the Study intervention period 2, the open label phase of the study, where all participants will be allocated to semaglutide s.c. 1.0 mg and neither study participants nor investigators are blinded to treatment allocation. Based on the second DBL a second CSR will be prepared.

Statistical Hypotheses

Analysis and reporting will be done and reported separately after *study intervention period 1* and *study intervention period 2*.

The primary aim is to investigate if semaglutide s.c. changes inflammation-related gene expression in immune cells from baseline to 12 weeks in CSF and blood compared to placebo in participants with Alzheimer's disease. The following two statistical tests of the co-primary endpoints will be performed jointly on a one-sided 2.5% significance level.

- Greater number of differentially expressed genes since baseline in immune cells in CSF after 12 weeks of treatment with semaglutide s.c. versus placebo.
- Greater number of differentially expressed genes since baseline in immune cells in blood after 12 weeks of treatment with semaglutide s.c. versus placebo.

Formally, the two null hypotheses, ($H_{0,1}$) and ($H_{0,2}$), will be tested against the corresponding alternative hypotheses, ($H_{a,1}$) and ($H_{a,2}$):

$$H_{0,1}: \mu\_{CSF} \leq 0.0 \text{ against } H_{a,1}: \mu\_{CSF} > 0.0$$

$$H_{0,2}: \mu\_{blood} \leq 0.0 \text{ against } H_{a,2}: \mu\_{blood} > 0.0$$

where μ_{CSF} denotes the treatment difference (s.c. Semaglutide minus placebo) in number of differentially expressed genes since baseline in immune cells in CSF after 12 weeks of treatment, and μ_{blood} denotes the treatment difference (s.c. Semaglutide minus placebo) in number of differentially expressed genes since baseline in immune cells in CSF after 12 weeks of treatment.

1.3 Multiplicity Adjustment

For the co-primary endpoints, the family-wise error rate will be controlled in the strong sense at $\alpha = 0.025$ by requiring the test for each associated hypothesis to be significant on a one-sided 2.5% level. For secondary and exploratory analyses testing will be done at a nominal significance level of 0.05 with no adjustment for multiple testing across analyses unless specified below.

For the exploratory analyses of biomarker data, p values will be adjusted using the Bonferroni-Holm method separately in plasma and CSF, respectively.

For the exploratory analyses of proteomics data, p values will be adjusted using the Bonferroni-Holm method separately in plasma and CSF, respectively.

2 Analysis Sets

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

Table 2-1 Participant Analysis Set

Participant Analysis Set	Description
Full analysis set (FAS)	• All randomised participants.
Safety analysis set (SAS)	• All participants who are exposed to investigational intervention.

The full analysis set will be used to analyse the endpoints and assessments not related to safety and the safety analysis set will be used to analyse the endpoints and assessments related to safety.

For analyses not related to safety, participants will be included in the analyses according to the planned investigational intervention; whereas for safety analyses, participants will be included in the analyses according to the investigational intervention they actually received.

The following data points sets are defined:

Table 2-2 Data Points Set

Data Points Sets	Description
DPS1 (in-trial)	All data points obtained at or after randomisation up to and including end of study visit. For participants who do not attend the end of study visit, only data points up to and including the date of the last study site contact (site or phone visit) are included.
DPS2 (on-treatment)	This data point set is a subset of DPS1 including all data points obtained while the participants are considered exposed to randomised treatment. The data point set includes all data points obtained at or after randomisation until date of last administration of randomised treatment + 35 days or last study site contact, whichever comes first.

DPS: Data Points Set

- FAS and DPS1 will be used to estimate the efficacy endpoints
- SAS and DPS2 will be used to present safety data with an acute onset
- SAS and DPS1 will be used to present safety data with a long lag-time

3 Statistical Analyses

3.1 General Considerations

For the co-primary endpoints controlled for multiplicity, one-sided p values will be used for the associated tests. For reporting of results, the estimated treatment effects will be accompanied by two-sided 95% confidence intervals and p values.

The latest available measurement, at or prior to the randomisation visit, will be used as the baseline measurement. If no measurement(s) has been obtained, at or prior to randomisation, the baseline value will be left missing.

If no statistical analysis is specified, data will be presented descriptively and, if relevant, using relevant summary statistics.

3.2 Primary Endpoints Analysis

3.2.1 Definition of Endpoints

The co-primary endpoints are the number of differentially expressed genes (DEGs) in immune cells in respectively CSF and blood at 12 weeks compared to baseline. Immune cells are defined as cells that, in uniform manifold approximation and projection (UMAP), belong to a cluster of cells with at least one transcript (gene expression count) of *PTPRC* (CD45), an established marker gene of nucleated hematopoietic cells, in more than 2% of cells in the cluster.

The co-primary endpoints will be derived by 1) applying standard preprocessing methodologies and 2) calculation of DEGs in immune cells in CSF and blood, respectively, for each study participant. The methodologies are identical for CSF and blood cells.

Note that the processing of data described in this section will be done before transfer of data to Study Data Tabulation Model (SDTM).

The following preprocessing steps [1] will be performed for all samples analysed and not flagged as ineligible for further analyses by the external vendors. The preprocessing will also include data for samples for participants only profiled at either baseline or week 12:

- In case only binary base call (BCL) file is received from the vendor: Conversion of BCL file output to FASTQ files using Cell Ranger with the version compatible with the applied library construction kit as specified by 10xGenomics
- Count matrix files will be generated by alignment to the human GRCh38 reference transcriptome and V(D)J libraries.Multiplexed samples (blood) will be demultiplexed using Cell Ranger multi as specified for demultiplexing and analysis of 5' immune profiling libraries pooled with hashtags by 10xGenomics. The final output files contain single cell gene expression counts and T cell receptor (TCR) counts.
- Samples containing 500 or more cells will be used for downstream analysis. Samples with less than 500 cells will be excluded based on the 10xGenomics guidelines for minimum number of cells.[2]

- The count matrices files will be further processed using Seurat V5 together with DropletUtils (version 1.18.1) using the emptyDrop function to exclude empty cell barcodes, as well as SoupX (version 1.6.2) to remove ambient RNA.
- Seurat V5 [3] will be applied for genes/cell filtering to minimize mitochondrial and ribosomal read contamination and relevant quality control (QC) metrics, and plots will be generated representing the sample level cell count and gene counts (i.e. feature counts).
- Normalization will be performed using the NormalizeData function and the FindVariablesFeatures function in Seurat V5
- As a QC step, to remove technical variation originating from scRNASeq library preparation the RNA assay in Seurat V5 will be applied. This will further allow inspection of cell cycle scoring, identification of highly variable features and scaling.
- A metadata matrix (SeuratObject) will be generated using Seurat V5, which includes the normalized and QC'ed data, followed by dimensionality reduction including optimal principal component identification, integration and batch correct using Harmony (version 1.2.0), and cell clustering (using Louvain algorithm), as well as generation of embedding and visualization of the high-dimensional non-linear data in 2-D space (UMAP) with default parameters as per best practises [4].

To calculate DEGs for each participant the following steps will be performed in Seurat V5:

- All immune cell clusters as defined above will be included to calculate DEGs
- A pseudobulk aggregated expression matrix using the AggregateExpression function will be generated from the SeuratObject, the cells will be grouped by cell cluster, participant and visit, and renormalized using NormalizeData.
- If any cell clusters has less than 5 cells for any of the participant across any visit, that cell cluster will be dropped.
- For each immune cell cluster in the pseudobulk aggregated expression matrix the 2000 most variable features (genes with highest expression variance) will be selected using the FindVariableFeatures(cell_data, selection.method = "vst", nfeatures = 2000), where cell_data is the pseudobulk data subset to a given immune cell cluster.
- For these 2000 genes, the difference between baseline and visit 5 (week 12) will be calculated for each cell cluster in each participant. If, for a given participant, any of the samples for baseline or week 12 is missing, the difference will be set to missing.
- For each participant, the number of genes with absolute log10 expression above 0.5 will be counted across all immune cell clusters and gene combinations to obtain the number of DEGs per participant which will be used for the analyses of the co-primary endpoints (see 3.2.2). Note that a log10 expression above 0.5 corresponds to about 3 times as high expression at week 12 compared to baseline and -0.5 is approximately 1/3 of the expression.

3.2.2 Main Analytical Approach

The analyses addressing the hypotheses for the co-primary endpoints will be based on the full analysis set (FAS). There will be no imputation of missing data and analyses will be performed using data for all participants profiled at baseline and week 12.

For CSF and blood respectively, the number of differentially expressed genes in immune cells at 12 weeks compared to baseline will be analysed using Welch's unequal variances t-test with treatment as grouping factor.

As stated in Section 1.3, the family-wise error rate will be controlled in the strong sense at $\alpha = 0.025$ by requiring the test for each associated hypothesis to be significant on a one-sided 2.5% level.

3.3 Secondary Endpoints Analysis

3.3.1 Treatment emergent adverse events (TEAEs)

All analyses of TEAEs will be made on the SAS.

All AEs will be coded using the newest version of the Medical Dictionary for Regulatory Activities (MedDRA) coding.

A TEAE is defined as an AE with onset included in DPS2 (see Table 2-2).

TEAEs will be summarised in terms of the number of participants with at least one event (N), the percentage of participants with at least one event (%), the number of events (E) and the event rate per 100 patient years of exposure time (R) for DPS2. The development over time in GI AEs will be presented graphically.

Results will be presented for the Study intervention period 1, and also after the Study intervention period 2 of the study.

3.3.2 Pharmacokinetic and/or pharmacodynamic modelling

Population PK analysis based on the semaglutide concentration data from the study will be performed.

The objective of the population PK analysis is to evaluate exposure levels of semaglutide s.c. in Alzheimer's disease patients. More technical and detailed elaboration of the population PK analysis will be given in a modelling analysis plan (MAP), which will be prepared before the second DBL.

The population PK analysis will be reported in a separate modelling report and will not be part of the biostatistical deliverables.

3.4 Exploratory Endpoints Analyses

3.4.1 Analyses of biomarker data

The following list of biomarkers will be analysed after Study intervention period 1:

CSF: Amyloidβ 42/40 ratio, p-Tau181, YKL-40, sTREM2, GFAP, CCL2, SMOC1, IL-6, IL-10, SPP1, IL-1β, CXCL10, IL-18, IL-1RA, complement factor 3, complement factor 4, TNF, NG, NfL, total Tau, sPDGFRB, 8-OHdG, and CSF albumin/serum albumin ratio, of which CCL2, SMOC1, IL-10, SPP1, IL-1β, CXCL10, IL-18, IL-1RA, complement factor 3, complement factor 4, TNF, sPDGFRB will be extracted from the CSF proteomics dataset.

Plasma: p-Tau181, GFAP, NFL, Amyloidβ 42/40 ratio, TNF, of which TNF will be extracted from the plasma proteomics dataset.

For each of the above biomarkers in CSF and each of the above biomarkers in plasma, the change from baseline in biomarker level after 12 weeks will be analysed using an ANCOVA model with treatment as fixed factor and baseline biomarker level as covariate and allowing for different residual variances between treatments. If deemed necessary, the endpoints will be log-transformed prior to analysis with the associated log-transformed baseline values as covariates.

For the following list of biomarkers only descriptive summaries will be made for the first CSR:

CSF: Amyloidβ 40 and Amyloidβ 42

Plasma: Amyloidβ 40 and Amyloidβ 42

After the Study intervention period 2 of the study for the second CSR, only descriptive summaries of biomarker data after visit 5 will be presented. In case measurements of additional biomarkers at baseline and visit 5 are available at the second DBL, explorative statistical analyses of change from baseline will be performed using the same method as described above.

The following list of biomarkers will be analysed after Study intervention period 2 if data for the biomarker is available:

Plasma: p-Tau181, GFAP, NFL, Amyloidβ 42/40 ratio, Amyloidβ 40, Amyloidβ 42, p-Tau217, isoprostanes.

3.4.2 Analyses of proteomics data

Olink performs two levels of QC, one for each sample and one from each protein where they can be categorized as PASS/WARN/FAIL. Any sample or protein that has the FAIL flag will not be included in any of the statistical outputs. In addition to the QC done by Olink, the median NPX values will be calculated for each protein, and also the Root Mean Square Distance (RMSD) for each sample to the median protein abundance. This will be normalized (divided) by the median RMSD such that most samples should have a score close to 1. Any sample with a score above 2 will be flagged as non eligible and will not be included in any of the statistical outputs.

Olink proteins are split into Blocks. For the majority of proteins there will be only one Block containing the protein. In the case where a protein is found in multiple Blocks, then the protein in the Block with highest number will be used for downstream analyses and the proteins in the other Blocks will be flagged as non eligible and will not be included in any of the statistical outputs.

Note that the processing of proteomic data described above will be done before transfer of data to SDTM.

Proteins that are included in the biomarker outputs will be excluded from the proteomics data before analyses of the remaining proteomic data and hence will not be included in the proteomics outputs.

For each of CSF and plasma proteomics, respectively, and for each protein the change from baseline to week 12 will be analysed using an ANCOVA model with treatment as fixed factor and baseline level as covariate and allowing for different residual variances between treatments.

All results (estimated treatment difference and associated un-adjusted p value) will be shown on a volcano plot. For selected proteins potentially related to AD pathophysiology, the results of the statistical analyses will be presented in a table. In addition, for the 100 proteins with most statistically significant differences between semaglutide s.c. 1.0 mg and placebo, the results of the statistical analyses will be presented in a table. In case of more than 100 results being statistically significant when adjusting for multiplicity, all statistically significant results when adjusting for multiplicity will be presented in the table. Descriptive summaries will only be shown for the proteins included in the statistical results tables.

3.4.3 Analysis of scTCRseq data

scTCRSeq of cells in CSF and blood, respectively, will be performed to estimate the change between baseline and visit 5 (week 12) in T cell repertoire composition to compare the effects of semaglutide s.c. 1.0 mg once weekly versus placebo.

The minimum number of T cells required for quantifying T cell repertoire changes is 10 cells, so samples with less cells will be excluded from further analyses.

The T-cell repertoire change will be quantified by the dissimilarity index, Morisita-Horn[5], which accounts for both the number of common clonotypes and the distribution of clone sizes and is sensitive to the clone sizes of the dominant clonotypes. A Morisita-Horn index of 1 represents two identical repertoires. The hypothesis is that the placebo group will experience no or minor change in the T cell repertoire composition, while the treatment group should experience a reduction in inflammation leading to a decrease in dominant T cells, and thereby a change in repertoire composition.

In a state of inflammation, the T cell repertoire will likely consist of some highly expanded clones, resulting in a repertoire with low diversity. When inflammation is reduced the dominant clones will shrink or disappear, likely leading to a repertoire of larger diversity. To quantify whether the repertoire is becoming more or less diverse, the Shannon diversity index[5] will be used, where increasing values indicate an increasingly even distribution of clone sizes in the repertoire. The methodologies are identical for cells in CSF and blood to count clonally expanded T cells.

TCR count matrices will be generated as a part of the initial preprocessing and demultiplexing (see section 3.2.1). TCR repertoire will be analysed using the NextFlow pipeline scrnaseq version 2.7.0 (https://nf-co.re/scrnaseq/2.7.0) github pull request 365.

Note that the calculation of Morisita-Horn index and Shannon index described above will be done for each participant before transfer of data to SDTM.

For each of CSF and blood, respectively, the Morisita-Horn index comparing baseline and visit 5 (week 12) will be analysed using Welch's unequal variances t-test with treatment as grouping factor.

For each of CSF and blood, respectively, the change from baseline to week 12 in Shannon index will be analysed using an ANCOVA model with treatment as fixed factor and baseline level as covariate and allowing for different residual variances between treatments.

3.4.4 Semaglutide concentration in CSF

Semaglutide concentration in CSF will be summarized using descriptive statistics.

3.5 Other Safety Analyses

All safety analyses will be made on the SAS.

3.5.1 Extent of Exposure

The extent of exposure will be summarised by treatment arm using relevant descriptive statistics.

3.5.2 Adverse Events

All AEs will be coded using the newest version of the Medical Dictionary for Regulatory Activities (MedDRA) coding. All AEs in the relevant periods will be included in the outputs.

Supportive summaries of AEs will be made based on DPS2. In case DPS2 differs from DPS1 additional summaries based on DPS1 will be made.

Results will be presented for the Study intervention period 1, and also after the Study intervention period 2 of the study.

3.5.3 Additional Safety Assessments

- C-SSRS

Change from baseline to visit 5 (week 12) for Study intervention period 1 and to visit 11 (week 64) for end of Study intervention period 2 in:

- Vital signs
  - pulse rate
  - systolic blood pressure
  - diastolic blood pressure
- Glucose metabolism
  - HbA1c
- Haematology
  - Erythrocytes
  - Haemoglobin blood
  - Leucocytes
  - Differential count (eosinophils, neutrophils, basophils, monocytes and
  - lymphocytes)
  - Thrombocytes
  - Haematocrit
- Coagulation
  - Prothrombin time
    Activated partial thromboplastin time

- Biochemistry
  - Alanine aminotransferase (ALT)
  - Alkaline phosphatase
  - Aspartate aminotransferase (AST)
  - Creatinine
  - Potassium
  - Sodium
  - Total bilirubin
- Inflammation
  - High sensitivity C-reactive protein (hsCRP)
- Renal function
  - eGFR
- Body weight
- Physical examination
- ECG

The additional safety assessment will be summarized using relevant descriptive statistics.

Antibodies

Anti-semaglutide antibodies will only be analysed by Novo Nordisk or a special laboratory appointed by Novo Nordisk if deemed necessary for clarification of unexpected drug exposure or other safety issues that may be related to antibody formation. If anti-semaglutide binding antibodies are measured, data as well as assay method description will be reported outside the CSR for this study.

3.6 Other Analysis

Not applicable.

3.7 Interim Analysis

No interim analysis is planned.

3.8 Changes to Protocol-planned Analysis

Not applicable.

4 Sample size determination

For sample size determination please refer to the protocol version 4.0 dated 06NOV2024, section 9.5.

5 Supporting Documentation

5.1 Appendix 1: Definition and calculation of endpoints, assessments and derivations

Type	Title	Time frame	Unit	Details
Primary endpoint	Change in gene expression assessed by scRNAseq (cells in CSF)	From baseline (week 0) to visit 5 (week 12)	Number of differentially expressed genes	Calculated as specified in section 3.2.1
Primary endpoint	Change in gene expression assessed by scRNAseq (cells in blood)	From baseline (week 0) to visit 5 (week 12)	Number of differentially expressed genes	Calculated as specified in section 3.2.1
Secondary endpoint	Number of treatment emergent adverse events (TEAEs)	From baseline (week 0) to visit 5 (week 12)	Number of events	
Secondary endpoint	Number of treatment emergent adverse events (TEAEs)	From baseline (week 0) to end of treatment (week 64)	Number of events	
Secondary endpoint	Weekly average semaglutide concentration (Cavg) based on population PK analysis	From visit 3 (week 4) to end of treatment (week 64)	nmol/L	
Exploratory endpoint	Change in biofluid-based (blood or CSF) neuroinflammatory and neurodegenerative biomarkers[a]	From baseline (week 0) to visit 5 (week 12)		
Exploratory endpoint	Changes in proteome in CSF	From baseline (week 0) to visit 5 (week 12)	NPX, Normalized Protein eXpression	
Exploratory endpoint	Changes in proteome in plasma	From baseline (week 0) to visit 5 (week 12)	NPX, Normalized Protein eXpression	
Exploratory endpoint	Changes in T cell clonal landscape assessed by scTCRseq (cells in CSF)	From baseline (week 0) to visit 5 (week 12)	Morisita-Horn index	Calculated as specified in section 3.2.1 and 3.4.3
Exploratory endpoint	Changes in T cell clonal	From baseline (week 0) to visit 5 (week 12)	Morisita-Horn index	Calculated as specified in section 3.2.1 and 3.4.3

Type	Title	Time frame	Unit	Details
	landscape assessed by scTCRseq (cells in blood)			
Exploratory endpoint	Semaglutide concentration in CSF[b]	At week 12		

Notes: [a] Biomarkers of interest is listed in section 3.4. Analysis will be performed only if a validated reliable assay is available at the end of study. [b] Analysis will be performed only if a validated reliable assay is available at the end of study.

6 References

1. 10xGenomics. Demultiplexing and Analyzing 5' Immune Profiling Libraries Pooled with Hashtags [internet] 2022 [cited 2024 01 May].
2. 10xGenomics. What is the minimum number of cells that can be profiled? [Internet] 2018 [cited 2024 22 Nov].
3. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol. 2024;42(2):293-304.
4. Amezquita RA, Lun ATL, Becht E, Carey VJ, Carpp LN, Geistlinger L, et al. Orchestrating single-cell analysis with Bioconductor. Nat Methods. 2020;17(2):137-45.
5. Chiffelle J, Genolet R, Perez MA, Coukos G, Zoete V, Harari A. T-cell repertoire analysis and metrics of diversity and clonality. Curr Opin Biotechnol. 2020;65:284-95.