Title page

Statistical Analysis Plan

Protocol title
Young adults with early-onset obesity treated with semaglutide – The RESETTLE study

Trial registration
EudraCT Number: 2019-002274-31
ClinicalTrials.gov ID: NCT05574439

SAP version
2.0

Date
9 October 2025

Protocol version
13

Authors

Simon B K Jensen[1], Sarah Byberg[1], Joachim Holt[1], Rasmus M Sandsdal[1], Louise A Holm[2,3], Lærke B Madsen[1], Bodil J Christensen[1], Axel I Poggi[2], Brice Ozenne[4,5], Morten A V Lund[1,3,6], Cilius E Fonvig[2,3,7], Torben Hansen[2], Jens-Christian Holm[2,3,7], Signe S Torekov[1]

Affiliations

[1]Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark; [2]Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark; [3]The Children's Obesity Clinic, Department of Paediatrics, Copenhagen University Hospital Holbæk, Holbæk, Denmark; [4]Neurobiology Research Unit, Rigshospitalet, Copenhagen, Denmark; [5]Department of Public Health, Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark; [6]Department of Cardiology, Copenhagen University Hospital - Rigshospitalet, Copenhagen, Denmark; [7]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.

Principal investigators

Sarah Byberg, Jens-Christian Holm

Sponsor-investigator

Signe S Torekov

Signatures

Date: 9 OCT 2025
Signe S Torekov
Sponsor-investigator

Date: 9 OCT 2025
Simon B K Jensen
SAP author

Date: 9 OCT 2025
Brice Ozenne
Statistician

Contents

1 Introduction

This document is the Statistical Analysis Plan (SAP) for the RESETTLE study. The SAP summarizes the design and objectives and provide detailed definitions of the endpoints and a detailed description of the planned statistical analyses. The SAP will be approved and signed prior to the last participant's last visit. The SAP is related to the primary trial report, which will include primary and key secondary endpoints, and harms.

1.1 Objectives and endpoints

Table 1 outlines the primary, secondary, and exploratory objectives with descriptions of the corresponding endpoints.

Table 1. Objectives and endpoints

Objectives	**Endpoints**		
Primary objective	**Primary endpoint[a]**		
	Title	**Time frame**	**Unit**
To demonstrate superiority of semaglutide 2.4 mg versus placebo as adjunct to non-pharmacological obesity care on *Body Mass Index (BMI)* in young adults with childhood-onset obesity who are non-responders (subgroup A) or insufficient responders (subgroup B)[b] to non-pharmacological childhood obesity treatment	Change in BMI	Week 0 (baseline) to week 68 (end of treatment)	Kg/m^2
Key secondary objectives	**Key secondary endpoints[a]**		
	Title	**Time frame**	**Unit**
To demonstrate superiority of semaglutide 2.4 mg versus placebo as adjunct to non-pharmacological obesity care on *body fat distribution* in young adults with childhood-onset obesity who are non-responders or insufficient responders[b] to non-pharmacological childhood obesity treatment	Change in fat mass[c]	Week 0 to week 68	Kg
	Change in fat percentage[c]	Week 0 to week 68	%-points
	Change in waist-to-height ratio	Week 0 to week 68	Ratio
	Relative change in visceral fat[d]	Week 0 to week 68	%
	Relative change in liver fat[d]	Week 0 to week 68	%
To demonstrate superiority of semaglutide 2.4 mg versus placebo as adjunct to non-pharmacological obesity care on *metabolic syndrome severity* in young adults with childhood-onset obesity who are non-responders or insufficient responders[b] to non-pharmacological childhood obesity treatment	Change in metabolic syndrome SDS	Week 0 to week 68	SDS
Additional secondary objectives	**Additional secondary endpoints**		
	Title	**Time frame**	**Unit**
To compare the effect of semaglutide 2.4 mg versus placebo on *body weight*	Change in body weight	Week 0 to week 68	Kg
	Percentage change in body weight	Week 0 to week 68	%

	Achievement of ≥5, ≥10, ≥15, and ≥20% weight loss	Week 0 to week 68	Count
To compare the effect of semaglutide 2.4 mg versus placebo on *fat-free mass*	Change in fat-free mass	Week 0 to week 68	Kg
To compare the effect of semaglutide 2.4 mg versus placebo on *markers of metabolic health*	Change in systolic blood pressure	Week 0 to week 68	mmHg
	Change in diastolic blood pressure	Week 0 to week 68	mmHg
	Change in waist circumference	Week 0 to week 68	Cm
	Change in blood lipids: • Triglycerides • Total cholesterol • Low-density lipoprotein-cholesterol (LDL) • High-density lipoprotein-cholesterol (HDL)	Week 0 to week 68	mM
	Change in fasting plasma glucose	Week 0 to week 68	mM
	Change in glycated hemoglobin (HbA1c)	Week 0 to week 68	%
To compare the effect of semaglutide 2.4 mg versus placebo on *health-related quality of life*	Changes in SF-36 scores: • Physical functioning • Role Physical • Bodily Pain • General Health • Vitality • Social Functioning • Role Emotional • Mental Health • Physical Component Summary • Mental Component summary	Week 0 to week 68	Score
Exploratory objectives	**Exploratory endpoints**		
	Title	**Time frame**	**Unit**
To explore the effect of semaglutide 2.4 mg versus placebo on *metabolic syndrome prevalence*	Proportion of participants who have metabolic syndrome	Week 0 and week 68	Count
To explore how *physical activity* changes during treatment with semaglutide 2.4 mg and placebo as adjunct to non-pharmacological obesity care	Change in daily steps	Week 0 to week 68	Steps/day
	Changes in moderate- and vigorous-intensity physical activity	Week 0 to week 68	Min/week
	Change in sedentary time	Week 0 to week 68	Min/day
To explore how *eating behavior and appetite* change during treatment with semaglutide 2.4	Changes in TFEQ-R18 scores: • Cognitive restraint	Week 0 to week 68	Score

mg and placebo as adjunct to non-pharmacological obesity care	• Emotional eating • Uncontrolled eating		
	Change in subjective appetite ratings	Week 0 to week 68	VAS (mm)
To explore longitudinal changes in *adiposity and metabolic health* in response to childhood treatment through early adulthood	• BMI SDS • Fat mass index (FMI) SDS • Waist circumference SDS	At first visit in childhood (subgroup A-D); last visit after childhood obesity treatment (subgroup A-C); and at baseline (subgroup A-D)	SDS
	Endpoints, for which data was not obtained in childhood	Baseline comparisons, group A-D	N/A
To explore the effect of semaglutide 2.4 mg in young adults with childhood-onset obesity defined as *non-responders*[b] to childhood obesity treatment versus young adults defined as *insufficient responders*[b]	Similar to primary and secondary endpoints for comparisons between subgroup A on semaglutide versus subgroup B on semaglutide		
To explore the effect of semaglutide 2.4 mg plus non-pharmacological obesity care in young adults with childhood-onset obesity defined as non-responders or insufficient responders[b] to non-pharmacological childhood obesity treatment *relative to reference groups without obesity (with and without childhood obesity)*	Similar to primary and secondary endpoints for subgroup A and B on semaglutide at week 68 versus baseline values for subgroup C and D.		
Safety objective	**Safety endpoints**		
	Title	**Time frame**	**Unit**
To describe the safety of semaglutide and placebo both as adjunct to non-pharmacological obesity care	Number of adverse events	Week 0 to week 71 (follow up safety visit)	Counts
	Number of serious adverse events	Week 0 to week 71 (follow up safety visit)	Counts

Baseline corresponds to week 0.
[a] Primary and key secondary endpoints will be controlled for multiplicity.
[b] Non-responders refer to subgroup A (BMI SDS reduction <0.1) and insufficient responders refer to subgroup B (BMI SDS reduction >0.25).
[c] Assessed by dual energy x-ray absorptiometry (DXA)
[d] Assessed by magnetic resonance imaging (MRI)
Abbreviations: BMI, body-mass index; SDS, standard deviation score; TFEQ-R18, three-factor eating questionnaire; VAS, visual analog scale.

1.1.1 Calculations

The primary endpoint is change in BMI from baseline (week 0) to end of treatment (week 68).
BMI is calculated as:

$$\frac{body\ weight\ (kg)}{height\ (m)^2}$$

BMI will be measured in fasted state at the following time points (weeks from randomization): 0, 17, 35, 52, and 68.

The following calculations are made for other endpoints:
Body fat percentage is calculated as:

$$\frac{total\ fat\ mass\ (kg)}{body\ weight\ (kg)} * 100$$

Waist-to-height ratio is calculated as:

$$\frac{waist\ circumference\ (cm)}{height\ (cm)}$$

Metabolic syndrome prevalence is scored (yes/no) according to the harmonized metabolic syndrome definition.[1] That is, three or more of the following:

- Waist circumference >94 cm (males) and >80 cm (females)
- HDL-c <1.0 mmol/L (males) and <1.3 mmol/L (females)
- Triglycerides ≥1.7 mmol/L
- Fasting glucose ≥5.6 mmol/L
- Systolic blood pressure ≥130 mmHg or diastolic blood pressure ≥85 mmHg

For the last four points, drug treatment for the measure may be an alternative indicator. Metabolic syndrome SDS is calculated based on the equations for Metabolic Syndrome Risk Z-Score developed by Gurka et al.[2]

All SF-36 scores are calculated based on the scoring instructions available here: https://www.rand.org/health-care/surveys_tools/mos/36-item-short-form/scoring.html

Self-reported physical activity is calculated based on the International Physical Activity Questionnaire[3]. Objectively measured physical activity will be extracted from wrist-worn activity trackers.

Eating behavior scores are calculated based on the TFEQ-R18[4] and appetite scores based on VAS.[5]

SDS for BMI is based on the reference values provided by Nysom et al.[6] and for FMI and waist with the use of the HOLBÆK study as reference.

1.2 Study population

Young adults (age 18–28 years) with obesity who have undergone at least one year of a non-pharmacological obesity care during childhood and a population-based reference group of young adults (age 18-28 years) with a normal weight development will be recruited from the HOLBAEK study (formerly The Danish Childhood Obesity Biobank).[7] Participants are divided into four subgroups based on their childhood treatment response and current BMI (subgroups A–C) and a reference group that has not received obesity treatment (subgroup D).

Subgroup A: young adults with obesity (BMI ≥30 kg/m^2) who completed at least one year of non-pharmacological childhood obesity treatment with a BMI-SDS reduction of less than 0.1.
Subgroup B: young adults with obesity at inclusion who completed at least one year of non-pharmacological childhood obesity treatment with a BMI-SDS reduction of more than 0.25.

Subgroups A and B will be recruited with an approximate ratio of 1:1.

Some exploratory analyses will involve an extended population including:
Subgroup C: young adults without obesity at inclusion who completed at least one year of non-pharmacological childhood obesity treatment with a BMI-SDS reduction of more than 0.5.
Subgroup D: young adults without obesity at inclusion and normal weight development.
Subgroup C and D will have baseline assessments.

1.3 Trial design

Subgroup A and B are recruited to participate in an investigator-initiated, randomized, placebo-controlled, parallel-group trial to investigate the effect of semaglutide compared with placebo as adjunct to non-pharmacological obesity care for treatment of childhood-onset obesity in young adults who underwent childhood obesity treatment with less than 0.1 (subgroup A, non-responders) or more than 0.25 (subgroup A, insufficient responders) reduction in BMI-SDS.

Participants will be randomly allocated, in a 2:1 ratio, stratified by sex (male/female), to subcutaneous semaglutide once weekly or placebo for 68 weeks. Treatment allocation will be based on a randomization list provided by Novo Nordisk and carried out by a non-blind study personnel member not otherwise related to the trial.

The trial includes a screening visit to assess the participants' eligibility followed by a baseline visit (week 0), at which participants are randomized. A period of 16 weeks of dose escalation is planned for dose escalation to 2.4 mg once weekly. Hereafter, study visits will take place approximately every 8th week until the end of treatment (week 68).

For details on trial design and participant eligibility criteria, see the published protocol paper[8] and the section *Research design and methods* in the full protocol.

1.4 Treatment strategies

Both semaglutide and placebo will be prescribed as adjuncts to hospital-based non-pharmacological obesity treatment.

1.4.1 Experimental (Semaglutide)

The treatment will be initiated at a dose of 0.24 mg once weekly for the first four weeks, with an increasing dose every fourth week to reach a maximal maintenance dose of 2.4 mg weekly by week 17 and maintained until week 68. Participants who experience unacceptable adverse effects at a given dose will receive the maximum dose at which they did not have such effects. This will be the case throughout the uptitration phase, which may be prolonged if necessary, and through the maintenance phase.

1.4.2 Control (Placebo)

Placebo will be volume-matched to semaglutide and given in identical pens. Placebo will follow the same treatment strategy as described for semaglutide.

1.5 Compliance

Compliance with trial medication will be assessed (self-reported) at weeks 1, 4, 8, 9, 12, 16, 17, 26, 35, 44, 52, 60, 68. Participants will be asked about the current dose and any missed dose, from which we will note a dose for each week in the trial. Trial medication compliance will be summarized descriptively in the full analysis set and include:

- Number and proportion of participants who discontinued medication before week 68
- Number and proportion of participants who reached maximum dose during the trial of 2.4 mg, 1.7 mg to <2.4 mg, 1.0 mg to <1.7 mg, 0.5 mg to <1.0 mg, 0.24 mg to <0.5 mg, and >0.0 mg to <0.24 mg
- Last dose for all participants
- Number and proportion with a last dose of 2.4 mg, 1.7 mg to <2.4 mg, 1.0 mg to <1.7 mg, 0.5 mg to <1.0 mg, 0.24 mg to <0.5 mg, and >0.0 mg to <0.24 mg
- Average dose after up titration

Compliance (%) will be based on number of doses and calculated as:

$$\frac{total\ number\ of\ times\ study\ medication\ were\ taken}{total\ number\ of\ times\ study\ medication\ were\ planned} \times 100$$

Per protocol is considered a compliance ≥75% without any intercurrent events (see 1.6).

Participants who permanently discontinue medication but have a BMI measurement at week 68 will be defined as retrieved (retrieved: yes/no). For retrieved participants, the last time point with an available BMI measurement while still on trial medication will be noted.

1.6 Intercurrent events

The following intercurrent events will be considered:

- Treatment discontinuation (any reasons). Treatment discontinuation will be defined as four consecutively missed doses of trial medication
- Initiation of other obesity treatment (medication or surgery)

1.7 Covariates

Only sex as a binary variable (male, female) will be considered as a covariate.

2 Statistical hypotheses

2.1 Primary analysis

The primary aim is to establish that, among young adults with obesity who received non-pharmacological childhood obesity treatment, semaglutide results in a larger mean decrease in BMI (in absolute value) from baseline to week 68 as compared with placebo, both used as adjunct to non-pharmacological obesity care, regardless of adherence to randomized treatment

or initiation of other obesity treatments (medication or surgery) (in accordance with the treatment policy strategy).

The corresponding population-level summary is therefore the mean BMI change under semaglutide minus the mean BMI change under placebo, which can be abbreviated as

$$(\overline{BMI}\,^{Semaglutide}_{week\ 68} - \overline{BMI}\,^{Semaglutide}_{week\ 0}) - (\overline{BMI}\,^{Placebo}_{week\ 68} - \overline{BMI}\,^{Placebo}_{week\ 0}) = \overline{\Delta BMI}^{Semaglutide} - \overline{\Delta BMI}^{Placebo}$$

where $\overline{BMI}\,^{g}_{t}$ denotes the mean BMI at timepoint t in group g and $\overline{\Delta BMI}^{g}$ denotes the mean change in BMI between week 68 and baseline in group g.

This population-level summary will be assessed in 3 populations sequentially:

Study population (subgroup A+B combined): Young adults with obesity who previously received non-pharmacological childhood obesity treatment with no or insufficient BMI SDS response. This aims at rejecting the null hypothesis in favor of the alternative hypotheses:

$$\boldsymbol{H_{1.0}}: \overline{\Delta BMI}^{Semaglutide} < \overline{\Delta BMI}^{Placebo}$$

(a smaller change is to be understood as more negative change, i.e., a larger decrease in BMI).

Subgroup A: young adults with obesity who completed at least one year of non-pharmacological childhood obesity treatment with a BMI-SDS reduction of less than 0.1. The corresponding alternative hypothesis $\boldsymbol{H_{1.A}}$ is $\boldsymbol{H_{1.0}}$ except that the mean BMI is evaluated in subgroup A instead of the whole study population.

Subgroup B: young adults with obesity at inclusion who completed at least one year of non-pharmacological childhood obesity treatment with a BMI-SDS reduction of more than 0.25. The corresponding alternative hypothesis $\boldsymbol{H_{1.B}}$ is $\boldsymbol{H_{1.0}}$ except that the mean BMI is evaluated in subgroup B instead of the whole study population.

2.2 Key secondary endpoints

The alternative hypotheses corresponding to the key secondary objectives of the study are as follows:

- **$H_{2.0}$; $H_{2.A}$;** and **$H_{2.B}$**: Similar approach as the primary endpoint (**$H_{1.0}$; $H_{1.A}$;** and **$H_{1.B}$**) but for change in fat mass from baseline to week 68.
- **$H_{3.0}$; $H_{3.A}$;** and **$H_{3.B}$**: Similar approach as the primary endpoint but for change in fat percentage from baseline to week 68.
- **$H_{4.0}$; $H_{4.A}$;** and **$H_{4.B}$**: Similar approach as the primary endpoint but for change in waist-to-height ratio from baseline to week 68.
- **$H_{5.0}$; $H_{5.A}$;** and **$H_{5.B}$**: Similar approach as the primary endpoint but for change in metabolic syndrome z-score from baseline to week 68.
- **$H_{6.0}$; $H_{6.A}$;** and **$H_{6.B}$**: Similar approach as the primary endpoint but for change in visceral fat from baseline to week 68.
- **$H_{7.0}$; $H_{7.A}$;** and **$H_{7.B}$**: Similar approach as the primary endpoint but for change in liver fat from baseline to week 68.

2.3 Multiplicity adjustment

The study hypotheses stated in section 2.1 and 2.2 will be tested hierarchically to preserve the family-wise error rate α at 0.05 with two sided tests. The hypothesis test strategy is shown in **Figure 1**. Specifically, the hypotheses are ordered and tested sequentially. If superiority of semaglutide 2.4 mg to placebo is confirmed, the α will be transferred to the next hypothesis, which will be tested. If superiority is not confirmed for a given endpoint, the null hypothesis testing will stop. Subsequent endpoints will be reported with estimated treatment contrast and two-sided 95% confidence intervals (CI) without P values.

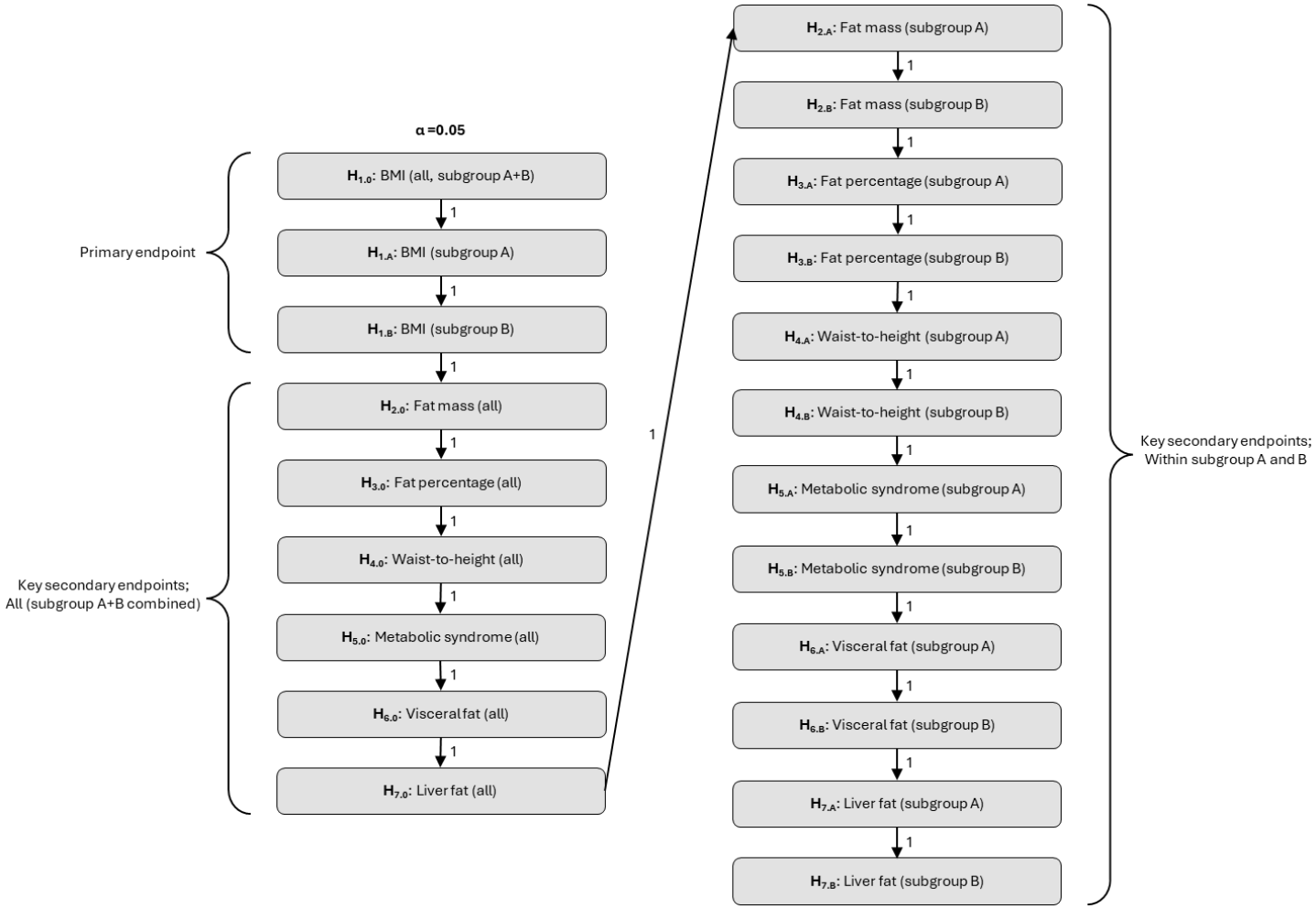

Figure 1. Graphical test strategy for primary and key secondary endpoints to control type 1 error.

3 Analysis sets

For all primary analyses of the primary and secondary endpoints and safety, the full analysis set will be used, defined as: All participants who were randomized and received at least one dose of trial medication (semaglutide or placebo), regardless of treatments adherence, discontinuation in the trial, and initiation of other obesity treatments (medication or surgery). Data will be used until week 68.

For exploratory analyses involving comparisons between subgroups, an extended analysis set will be used, defined as participants with an available measurement in childhood (HOLBAEK study) and attended baseline assessments as either subgroup A, B, C, or D.

4 Statistical analyses

4.1 General considerations

The primary analyses of efficacy and safety endpoints will be analyzed using the full analysis set (according to the intention-to-treat principle).

Statistical analyses to compare semaglutide 2.4 mg to placebo will be 2-sided, and the significance level is 5%. Results from the statistical tests for the primary and key secondary endpoints will be reported with estimated treatment contrast and two-sided 95% CI and P values (according to the hierarchical testing, i.e. P values are shown if prior tests showed superiority, Figure 1). If prior tests do not all show superiority, the estimated summary statistic will only be reported with its confidence interval but not P value. Additional secondary and exploratory endpoints will be reported with estimated treatment contrast and 95% CI unadjusted for multiple testing.

The objective of all statistical analyses of semaglutide 2.4 mg compared with placebo is to confirm superiority for semaglutide; therefore, the framework is superiority hypothesis testing. Superiority will be claimed if P values are less than 5% and the estimated treatment contrast favors semaglutide 2.4 mg.

Summary statistics for continuous variables may include number of observations, mean, standard deviation, median, and interquartile range. Summary statistics for categorical variables may include number of observations, frequency, and percentages.

4.1.1 Handling of missing data

Missing data will be handled under the hypothetical strategy, i.e., what would have happened had none of the participants included in the analysis set had missing values. These may arise due to participant dropout, technical issues preventing a measurement, or measurement not satisfying quality control. Participant dropout is defined as missing endpoint at week 68. The time of dropout is the first visit at which the endpoint is missing for this visit and all following visits.

For the primary and the key secondary analyses only missing outcome values are expected as the other variables considered are sex, randomization group, participant ID, and visit index. The primary and secondary analyses will assume that:

(1) participants with missing outcome data follow their randomization group trajectory.

Two planned sensitivity analyses will be performed assuming either:

(2) participants who dropout from the semaglutide group follow, from the time of dropout, the same trajectory as the placebo group.
(3) participants who dropout follow, from the time of dropout, the trajectory of participants in the same treatment group who permanently discontinue study medication.

Linear mixed models will be used to model the participant trajectories: by randomization group and possibly by time at which study medication is permanently discontinued. The estimated parameters from the mixed model will be used to quantify the group difference under (1). (2) and (3) will use multiple imputations, based on the mixed model estimates, to complete the dataset and assess the group difference using an ANCOVA.

4.1.2 Handling of intercurrent events

For all primary analyses of the primary and secondary endpoints, all participants outcome measurements will be included regardless of the occurrence of intercurrent events, according to the treatment policy strategy.

4.1.3 Statistical software

R version 3.6.0 or newer (The R Foundation for Statistical Computing, www.R-project.org) and SAS version 9.4 or newer (SAS Institute, Cary, NC, USA).

4.2 Primary endpoint analysis

4.2.1 Primary analytical approach

The primary analysis of the primary endpoint will be based on all available BMI data in the full analysis set.

For the primary endpoint, a mixed model for repeated measurements will be used to model the mean BMI over time, for each treatment arm and subgroup, adjusted for sex. BMI measurement at timepoint 0, 17, 35, 52, and 68 will be used and the mixed model will be parametrized as follows: at baseline, the mixed model will use two parameters to model the mean BMI, one for each subgroup (A, B). Indeed, due to randomization, there should be no baseline difference, in average, between participants from the same subgroup but allocated to a different treatment. At each follow-up, the mixed model will use four parameters to model the mean BMI, one for each combination of subgroup (A, B) and treatment arm (semaglutide, placebo). The residual variance will be modeled using an unstructured covariance matrix stratified on treatment arm and subgroup: a separate variance (resp. correlation) parameter will be used for combination of treatment arm, subgroup, and timepoint (resp. pair of timepoints). The parametrization of the mean corresponds to a three-way interaction between time, subgroup, and treatment where treatment is set to placebo for all participants at baseline and to treatment arm (placebo or semaglutide) at each follow-up.

To fit this model, the `lmm()` function from the LMMstar package[9] will be used. In the case where the optimizer would return warnings indicating optimization issues, alternative optimizers will

be used instead (e.g., `mmrm()` from the mmrm package). If this is unsuccessful, a simpler covariance pattern will be considered (unstructured covariance stratified on only treatment arm, unstructured covariance without stratification, Toeplitz stratified on only treatment arm).

The estimated mean difference in BMI change (week 68 vs. baseline) between the semaglutide arm and placebo arm will be averaged between the two subgroups to provide an estimate for $\overline{\Delta BMI}^{Semaglutide} - \overline{\Delta BMI}^{Placebo}$. This corresponds to averaging two interactions parameters. A P value and a 95% confidence interval will be based on a Wald test, assumed to follow a Student's t-distribution with a standard error derived from the observed information matrix and degree of freedom evaluated using Satterthwaite approximation (default in LMMstar, a similar strategy would be used with `mmrm()` but based on the package default: expected information and Kenward-Roger degrees of freedom).

A P value below 0.05 with a negative estimate for $\overline{\Delta BMI}^{Sema} - \overline{\Delta BMI}^{Placebo}$ would demonstrate superiority, i.e., would lead to accept $\boldsymbol{H_{1.0}}$. We would then proceed to decide upon $\boldsymbol{H_{1.A}}$ by extracting the estimated mean difference in BMI change (week 68 vs. baseline) between subgroup A in the semaglutide arm and subgroup A in the placebo arm, corresponding to one of the two interaction parameters from the same mixed model. No averaging is needed, and statistical inference would be carried out as before to conclude about $\boldsymbol{H_{1.A}}$. If rejected, we would consider the other subgroup, i.e., the other interaction term from the same mixed model and perform statistical inference similarly to conclude about $\boldsymbol{H_{1.B}}$.

4.2.2 Sensitivity analyses

Two sensitivity analyses for the primary endpoint will be performed to test the robustness of the primary analysis with different assumptions on BMI development after loss to follow-up. Both will be carried out using multiple imputations with 1000 imputed datasets. The random number generator state (seed) will be set to 1 when starting the imputation of the first dataset, 2 for the second, and so on up to 1000.

4.2.2.1 Jump to reference multiple imputations

In this approach, the BMI changes after dropout in the semaglutide group will resemble the development in the placebo group. Thus, when dropping out of the semaglutide group (i.e., no more observed endpoint) it is assumed that participants immediately lose any potential effect of semaglutide treatment and follow the same BMI trajectory as participants allocated to placebo plus hospital-based non-pharmacological obesity care. Participants who drop out of the placebo group will still follow the trajectory of their group.
This scenario will be implemented as follows:

- The imputation model is estimated as the mixed model from the primary analysis except that the covariance pattern is not stratified on treatment arm (otherwise imputation would require an estimate of the correlation when changing treatment arm which is not accessible from the observed data).
- The original dataset will be duplicated and the arm variable will be assigned to 'placebo' for each subject after the last timepoint at which an outcome variable is observed for this individual.

- This dataset will be used as an input to the `predict()` function with the argument `type = "impute"` to evaluate the mean and variance of the missing outcomes conditional on observed outcomes and covariates (including time, treatment arm, subgroup) and impute values by sampling from the corresponding normal distribution.
- In each imputed dataset, the arm variable will be re-assigned to its original value (as per randomization) and the linear mixed model defined in the primary analysis will be fit on the resulting dataset. In practice, to save computation time we will leverage that, with complete data, the mixed model estimator of $\overline{\Delta BMI}^{Semaglutide} - \overline{\Delta BMI}^{Placebo}$ reduces to an ANCOVA approach where the outcome at week 68 is regressed against baseline (interacting with subgroup) and treatment arm (interacting with subgroup).
- The estimate of $\overline{\Delta BMI}^{Semaglutide} - \overline{\Delta BMI}^{Placebo}$ will be extracted for each imputed dataset and pooled using Rubin's rule.

4.2.2.2 Retrieved participants multiple imputations

In this approach, the BMI changes after dropout will resemble the development for participants in the same treatment group who permanently discontinue study medication but have available week 68 endpoint (retrieved participants). The exact implementation of this scenario will depend on the missing data pattern and corresponding retrieved participants.
Nevertheless, it is anticipated to be implemented as follows:- The full analysis set is restricted to all participants with available BMI endpoint at week 68. Last timepoint on treatment will be noted for retrieved participants and set to 68 for non-retrieved participants. This dataset is referred to as the retrieved participant dataset (even though it contains patients who fully complied with the study medication protocol).
- the imputation model is estimated with a mixed model similar to the primary analysis with the last timepoint on treatment variable interacting with time and treatment arm variable in the mean structure for times beyond the last timepoint on treatment. This implicitly assumes that the effect of non-adherence is the same in sub-population A and B. In case of convergence issues, change of optimizer or simplification in the mean and variance structure will be considered.
- The full analysis set is then restricted to patients with missing week 68 endpoint. This dataset is referred to as the dropout dataset. The variable 'last timepoint on treatment' is set to the dropout time (i.e. last timepoint at which BMI is observed)
- If the interaction between 'last timepoint on treatment' with time and treatment arm has levels in the dropout dataset that are not present in the retrieved participant dataset, these extra levels are set to the next available last timepoint on treatment level from the same treatment arm and the corresponding extra BMI measurements set to missing (example: no retrieved participant for participants dropping out at week 17 in the treatment arm but retrieved participants for week 35).
- The dropout dataset, restricted to timepoints with observed endpoint and week 68, will be used as an input to the `predict()` function with the argument `type = "impute"` to evaluate the mean and variance of the missing outcomes conditional on observed outcomes and covariates (including time, last timepoint on treatment, treatment arm, subgroup) and impute values by sampling from the corresponding normal distribution.
- The subset of the original dataset composed of participants with observed week 68 endpoint is combined with each imputed dataset. An ANCOVA approach where the outcome at week 68 is

regressed against baseline (interacting with sub-group) and treatment arm (interacting with subgroup) to estimate $\overline{\Delta BMI}^{Sema} - \overline{\Delta BMI}^{Placebo}$. This is an equivalent approach to the mixed model with complete data that only requires to impute outcome at week 68.
- The estimate of $\overline{\Delta BMI}^{Sema} - \overline{\Delta BMI}^{Placebo}$ will be extracted for each imputed dataset pooled using Rubin's rule.

4.2.3 Supplementary analysis: Hypothetical strategy for handling intercurrent events

A supplementary analysis for the primary endpoint will be performed using a hypothetical strategy for handling intercurrent events. This analysis assesses the treatment effect if all participants had continued to receive randomized treatment without initiating other obesity treatments (medication or surgery). This analysis will use the same constrained mixed model for repeated measurements as the primary analysis, but all endpoint values after an intercurrent event will be set to missing. Intercurrent events are defined as: treatment discontinuation or initiation of other obesity medication or obesity surgery. Missing data will be assumed to be missing at random (related to observed endpoints, sex) and handled implicitly in the constrained mixed model for repeated measurements by maximum likelihood estimation.

4.3 Key secondary endpoint analysis

A schematic overview of key secondary endpoints is shown in Table 1. Key secondary endpoints will be controlled for multiple testing as shown in Figure 1. The analytical approach will be a constrained mixed model for repeated measurements similar to the primary analytical approach described for the primary endpoint, and include comparisons between semaglutide and placebo for all participants and for subgroups A and B. All key secondary endpoints are assessed at week 0 and week 68, and waist-to-height ratio is additionally assessed at weeks 17, 35, and 52. If possible, the sensitivity analyses and supplementary analysis described for the primary endpoint (section 4.2.2 and 4.2.3) will also be performed for key secondary endpoints.

4.4 Additional secondary endpoints

A schematic overview of additional secondary endpoints is shown in Table 1. Additional secondary endpoints will not be controlled for multiple testing. The analytical approach will be a constrained mixed model for repeated measurements similar to the primary analytical approach described for the primary endpoint, and include comparisons between semaglutide and placebo for all participants and for subgroups A and B. In addition, the proportion of participants with weight loss of at least 5%, 10%, 15%, and 20% by treatment group will be reported descriptively. All additional secondary endpoints are assessed at week 0 and week 68, and the following endpoints are also assessed at weeks 17, 35, and 52: body weight, systolic and diastolic blood pressure, waist circumference, and fasting plasma glucose.

4.5 Exploratory endpoints

A schematic overview of exploratory endpoints is shown in Table 1.

Metabolic syndrome prevalence will be reported descriptively for baseline and week 68. Metabolic syndrome prevalence is scored (yes/no) according to the harmonized metabolic syndrome definition.[1]

Physical activity will be measured during the study with an activity tracker and at baseline and 68 with the International Physical Activity Questionnaire.[3] The following variable will be summarized descriptively and changes over time by treatment group will also be estimated with a constrained mixed model for repeated measurements: daily steps, moderate and vigorous-intensity physical activity, and sedentary time.

Eating behavior will be measured using three factor eating questionnaire[4] and appetite will be measured with VAS[5] in fasting state and in response to standardized mixed meal at baseline and week 68. The following variables will be summarized descriptively and changes over time by intervention group will also be estimated with a constrained mixed model for repeated measurements: cognitive restraint score, emotional eating score, uncontrolled eating score and fasting and postprandial appetite ratings.

Trajectories in adiposity-related and metabolic health-related variables will include the following time points:

- First visit in childhood (subgroup A-D). For subgroup A-C, this corresponds to childhood treatment initiation. For subgroup D this corresponds to the only childhood assessment.
- Last visit after childhood obesity treatment (subgroup A-C). This is used to evaluate response to childhood obesity treatment (change from first to last childhood visit).
- Baseline (subgroup A-D). This is the first visit after recruitment to the study, where the participants are now young adults. For subgroup A and B this corresponds to week 0 in the randomized trial, i.e. the time of randomization.

Variables will be summarized descriptively at each time point. When relevant (i.e., when subgroups are not inherently different as a consequence of the subgroup inclusion criteria), subgroups will be compared in terms of absolute value of the variable and change in variable using general linear models with sex and age as covariates if relevant.
The variables analyzed over time are:

- BMI SDS
- Fat mass index SDS
- Waist circumference SDS

For other variables not obtained in childhood, we will do baseline comparisons only.

From the constrained mixed models described for the primary and secondary endpoints, we will compare participants treated with semaglutide in subgroup A versus subgroup B. Likewise, we will compare endpoints for participants treated with placebo in the two subgroups.

For variables measured for primary and secondary endpoints, values after semaglutide treatment (subgroup A and B on semaglutide at week 68) will be compared to reference groups without obesity as young adults (subgroup C and D). Comparisons will be adjusted for sex and age if relevant.

4.6 Safety endpoints

Safety will be assessed in the full analysis set. Adverse events (as defined in the protocol) with an incident of ≥ 5% and all serious adverse events will be reported by system organ class and preferred term in accordance with the Medical Dictionary for Regulatory Activities (MedDRA) version 28.1 for all participants combined and by treatment group. Events will be reported as the number and percentage of participants experiencing an adverse event. Frequencies of adverse events will be reported descriptively. Serious adverse events are defined as any adverse event leading to hospital admission, prolongation of a hospitalization or death; results in persistent or significant disability/incapacity; consists of a congenital anomaly or birth defect; or involves suspicion of transmission of infectious agents.

4.7 Statistical interim analyses and stopping guidance

No interim analysis are planned for the study, and no stopping guidelines were made. Semaglutide 2.4 mg is approved for the treatment of obesity, and the dosage will be kept within the approved maximum of 2.4 mg.

4.8 Timing of final analysis

All endpoints presented in this SAP will be analyzed collectively after the last participant's last visit and after the statistical analysis plan has been signed. The analysis of primary and key secondary endpoints will be performed blinded to treatment group. The primary endpoint will be analyzed by a statistician that has not been involved in the conduction of the trial.

5 Sample size calculations

The sample size was calculated based on the primary endpoint, change in BMI from randomization (week 0) to end of treatment (week 68). A sample size of 60 (40 in the semaglutide and 20 in the placebo group) gives 90% power to detect a difference in means of 1.8 kg/m^2 assuming that the common standard deviation (SD) for change is 2.0 kg/m^2, with a 0.05 two-sided significance level. The estimate used for the SD was based on our previous trials with GLP-1RA and placebo for treatment of obesity.[10,11] A reduction in BMI of 1.8 kg/m^2 (corresponding to approximately 5 kg) is associated with improvements in cardiometabolic risk factors (e.g., blood pressure, blood lipids, and glucose levels) and health-related quality of life,[12,13] although larger reductions yields larger benefits and may be necessary in the presence of obesity-related complications.[14-16] Semaglutide 2.4 mg has previously been shown to reduce BMI by about 4 kg/m^2.[17,18] Although our study population could experience slightly lower effect of semaglutide, a between-group difference of 1.8 kg/m^2 is considered realistic to detect and clinically relevant and is therefore defined as the minimum relevant difference (MIREDIF). To account for a dropout rate of 20%, at least 75 participants will be recruited to both subgroups A and B. We expect to have at least 60 completers in both group A and B, which will give >99% power for the primary endpoint in the groups combined (~120 completers) and 90% in group A and B, separately. If we get 75 completers in each group, the power is 95%.

We also calculated power for the key secondary endpoints included in the hypothesis test hierarchy. The power, assumed sample size, mean difference, and common SD for primary and key secondary endpoints are listed in Table 3.

For change in fat mass and fat percentage, the common SD was based on our previous studies.[10,11] Placebo-subtracted reductions in fat mass/fat percentage were 8.4 kg/3.5 %-points after semaglutide 2.4 mg for 68 weeks,[17] 3.5 kg after semaglutide 1 mg for 12 weeks,[19] and 2.6 %-points after liraglutide 3.0 mg/day for 16 weeks.[10] The MIREDIF of 3.5 kg/ 2.5%-points is therefore considered realistic to detect.

For change in waist-to-height and metabolic syndrome z-score, the mean differences and common SD were based on our previous liraglutide randomized trial.[11,20] Waist-to-height can be used as an anthropometric criterion to indirectly confirm excess adiposity,[21,22] and predicts obesity-related cardiovascular risk outcomes and all-cause mortality.[23,24] The hazard ratio for developing type 2 diabetes is 1.9 for an increase in metabolic syndrome z-score of 0.25-0.5, and 2.7 for an increase in metabolic syndrome z-score of >0.5, independent on the individual metabolic syndrome components.[25] Thus, we consider the MIREDIF in metabolic syndrome z score of 0.4 clinically relevant.

For change in visceral fat and liver fat, the common SD is based on MRI studies of liraglutide 3.0 mg and tirzepatide.[26,27] The effect size for visceral fat in these studies were 11% with liraglutide and 16-25% with tirzepatide 5-15 mg/week. In a CT study, the reduction with semaglutide was 15% with 1.7 mg and 33% with 2.4 mg. The MIREDIF for visceral fat of 10% is therefore considered realistic to detect. For liver fat, semaglutide 2.4 mg showed relative reductions of 42-50%[28] in patients with NAFLD and 33% in patients with NASH plus cirrhosis.[29] A 30% relative reduction is associated with improvement in NAFLD and fibrosis[30,31] and was therefore defined as the MIREDIF.

Table 3. Power calculations

Test order (Hypothesis)	Endpoint	Assumed available week 68 measurement (n)	Assumed mean difference	Assumed common SD	Power (%)
1 ($\mathbf{H_{1.0}}$)	Change in BMI (all)	120	1.8 kg/m^2	2.0 kg/m^2	>99
2+3 ($\mathbf{H_{1.A+1.B}}$)	Change in BMI (subgroup A+B)	60	1.8 kg/m^2	2.0 kg/m^2	90
4 ($\mathbf{H_{2.0}}$)	Change in fat mass (all)	120	3.5 kg	3.8 kg	>99
5 ($\mathbf{H_{3.0}}$)	Change in fat percentage (all)	120	2.5 %-points	3.0 %-points	99
6 ($\mathbf{H_{4.0}}$)	Change in waist-to-height ratio (all)	120	0.035	0.040	99
7 ($\mathbf{H_{5.0}}$)	Change in metabolic syndrome z-score (all)	120	0.4	0.5	>98
8 ($\mathbf{H_{6.0}}$)	Change in visceral fat (all)	100	10%	15%	87
9 ($\mathbf{H_{7.0}}$)	Change in liver fat (all)	100	30%	45%	87
10+11 ($\mathbf{H_{2.A+2.B}}$)	Change in fat mass (subgroup A+B)	60	3.5 kg	3.8 kg	90
12+13 ($\mathbf{H_{3.A+3.B}}$)	Change in fat percentage (subgroup A+B)	60	2.5 %-points	3.0 %-points	85
14+15 ($\mathbf{H_{4.A+4.B}}$)	Change in waist-to-height ratio (subgroup A+B)	60	0.035	0.040	88

16+17 ($\mathbf{H}_{5.A+5.B}$)	Change in metabolic syndrome z score (subgroup A+B)	60	0.4	0.5	82
18+19 ($\mathbf{H}_{6.A+6.B}$)	Change in visceral fat (subgroup A+B)	50	10%	15%	59
20+21 ($\mathbf{H}_{7.A+7.B}$)	Change in liver fat (subgroup A+B)	50	30%	45%	59

Power calculations were based on independent samples t-tests.

6 Participant disposition

A CONSORT flow diagram will be provided, with reported number of participants for each of the following:

For subgroup A, B, C, and D

- Invitation letters sent based on inclusion criteria
- Screened
- Screen failure with reasons

For participants initially recruited for subgroup C only

- Eligible for subgroup C; recruited for subgroup B because of current BMI ≥ 30 kg/m^2

For subgroup A and B only

- Randomized to receive semaglutide or placebo
- Received at least one dose of semaglutide or placebo
- Attended end of treatment visit (week 68)
- Discontinued study treatment with reasons
- Lost to follow-up with reasons
- Included in statistical analysis
- Excluded from analysis with reasons

7 Characteristics of the participants

Demographic and clinical characteristics of the participants will be summarized by subgroup (A-D) and by study treatment (semaglutide, placebo) for subgroup A and B. Participants' characteristics may include but are not limited to:

Demographic characteristics

- Age (years)
- Sex assigned at birth (female, male)
- Country of birth, parental country of birth, ethnicity (self-reported)
- Completed education level (ISCED)
- Father and mother education level (ISCED), employment (unemployed, part time employed, full time employed)

Clinical characteristics

- Body weight (kg), height (cm), and BMI (kg/m2)
- BMI SDS
- Fat mass (kg)

- Fat-free mass (kg)
- Fat percentage (percentage-points)
- Waist circumference (cm), waist-to-height ratio
- Liver fat content (percentage points)
- Visceral fat (liter)
- Abdominal subcutaneous fat (liter)
- Metabolic syndrome z-score
- Metabolic syndrome prevalence (n, %)
- Systolic and diastolic blood pressure (mmHg)
- Resting heart rate (bpm)
- Lipid levels (Total cholesterol (mM), LDL-c (mM), HDL-c (mM), and triglycerides (mM))
- Glycated hemoglobin (%) and fasting plasma glucose (mM)
- SF-36 component scores (0-100): Physical functioning, Role Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role Emotional, Mental Health, Physical Component Summary, Mental Component summary
- Obesity-related diseases and psychiatric disorders

Childhood clinical characteristics

- Age, first visit
- Treatment duration (for group A, B, and C)
- BMI SDS, first visit (corresponding to treatment start for group A, B, and C)
- BMI SDS, last visit (for group A, B, and C)
- BMI SDS reduction, start to last visit (for group A, B, and C)
- Similar values as BMI SDS for fat mass index
- Time from final visit to RESETTLE inclusion

8 References

1. Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009; **120**(16): 1640-5.
2. Gurka MJ, Lilly CL, Oliver MN, DeBoer MD. An examination of sex and racial/ethnic differences in the metabolic syndrome among adults: a confirmatory factor analysis and a resulting continuous severity score. *Metabolism* 2014; **63**(2): 218-25.
3. Booth M. Assessment of physical activity: an international perspective. *Res Q Exerc Sport* 2000; **71 Suppl 2**: 114-20.
4. Anglé S, Engblom J, Eriksson T, et al. Three factor eating questionnaire-R18 as a measure of cognitive restraint, uncontrolled eating and emotional eating in a sample of young Finnish females. *Int J Behav Nutr Phys Act* 2009; **6**: 41.
5. Flint A, Raben A, Blundell J, Astrup A. Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. *International journal of obesity* 2000; **24**(1): 38-48.
6. Nysom K, Mølgaard C, Hutchings B, Fleischer Michaelsen K. Body mass index of 0 to 45-y-old Danes: reference values and comparison with published European reference values. *International Journal of Obesity* 2001; **25**(2): 177-84.
7. Holm JC, Gamborg M, Bille DS, Gr Nb KH, Ward LC, Faerk J. Chronic care treatment of obese children and adolescents. *Int J Pediatr Obes* 2011; **6**(3-4): 188-96.
8. Byberg S, Holt J, Sandsdal RM, et al. Protocol for a randomised, double-blinded, controlled trial of youth with childhood-onset obesity treated with semaglutide 2.4 mg/week: the RESETTLE trial. *BMJ Open* 2024; **14**(11): e082446.
9. Ozenne B, Forman J. LMMstar: Repeated measurement models for discrete times. R package version 1.1.0. 2024.
10. Iepsen EW, Zhang J, Thomsen HS, et al. Patients with Obesity Caused by Melanocortin-4 Receptor Mutations Can Be Treated with a Glucagon-like Peptide-1 Receptor Agonist. *Cell Metab* 2018; **28**(1): 23-32.e3.
11. Lundgren JR, Janus C, Jensen SBK, et al. Healthy Weight Loss Maintenance with Exercise, Liraglutide, or Both Combined. *N Engl J Med* 2021; **384**(18): 1719-30.
12. Wing RR, Lang W, Wadden TA, et al. Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. *Diabetes Care* 2011; **34**(7): 1481-6.
13. Yumuk V, Tsigos C, Fried M, et al. European Guidelines for Obesity Management in Adults. *Obes Facts* 2015; **8**(6): 402-24.
14. Wharton S, Lau DCW, Vallis M, et al. Obesity in adults: a clinical practice guideline. *Canadian Medical Association Journal* 2020; **192**(31): E875-E91.
15. Garvey WT, Mechanick JI, Brett EM, et al. AMERICAN ASSOCIATION OF CLINICAL ENDOCRINOLOGISTS AND AMERICAN COLLEGE OF ENDOCRINOLOGY COMPREHENSIVE CLINICAL PRACTICE GUIDELINES FOR MEDICAL CARE OF PATIENTS WITH OBESITY. *Endocr Pract* 2016; **22 Suppl 3**: 1-203.
16. Logue J, Thompson L, Romanes F, Wilson DC, Thompson J, Sattar N. Management of obesity: summary of SIGN guideline. *Bmj* 2010; **340**: c154.
17. Wilding JPH, Batterham RL, Calanna S, et al. Once-Weekly Semaglutide in Adults with Overweight or Obesity. *N Engl J Med* 2021; **384**(11): 989-1002.

18. Wadden TA, Bailey TS, Billings LK, et al. Effect of Subcutaneous Semaglutide vs Placebo as an Adjunct to Intensive Behavioral Therapy on Body Weight in Adults With Overweight or Obesity: The STEP 3 Randomized Clinical Trial. *Jama* 2021; **325**(14): 1403-13.
19. Blundell J, Finlayson G, Axelsen M, et al. Effects of once-weekly semaglutide on appetite, energy intake, control of eating, food preference and body weight in subjects with obesity. *Diabetes Obes Metab* 2017; **19**(9): 1242-51.
20. Sandsdal RM, Juhl CR, Jensen SBK, et al. Combination of exercise and GLP-1 receptor agonist treatment reduces severity of metabolic syndrome, abdominal obesity, and inflammation: a randomized controlled trial. *Cardiovascular Diabetology* 2023; **22**(1): 41.
21. Rubino F, Cummings DE, Eckel RH, et al. Definition and diagnostic criteria of clinical obesity. *Lancet Diabetes Endocrinol* 2025; **13**(3): 221-62.
22. Busetto L, Dicker D, Frühbeck G, et al. A new framework for the diagnosis, staging and management of obesity in adults. *Nature Medicine* 2024; **30**(9): 2395-9.
23. Ashwell M, Gunn P, Gibson S. Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. *Obes Rev* 2012; **13**(3): 275-86.
24. Abdi Dezfouli R, Mohammadian Khonsari N, Hosseinpour A, Asadi S, Ejtahed HS, Qorbani M. Waist to height ratio as a simple tool for predicting mortality: a systematic review and meta-analysis. *Int J Obes (Lond)* 2023; **47**(12): 1286-301.
25. Gurka MJ, Golden SH, Musani SK, et al. Independent associations between a metabolic syndrome severity score and future diabetes by sex and race: the Atherosclerosis Risk In Communities Study and Jackson Heart Study. *Diabetologia* 2017; **60**(7): 1261-70.
26. Gastaldelli A, Cusi K, Fernández Landó L, Bray R, Brouwers B, Rodríguez Á. Effect of tirzepatide versus insulin degludec on liver fat content and abdominal adipose tissue in people with type 2 diabetes (SURPASS-3 MRI): a substudy of the randomised, open-label, parallel-group, phase 3 SURPASS-3 trial. *Lancet Diabetes Endocrinol* 2022; **10**(6): 393-406.
27. Neeland IJ, Marso SP, Ayers CR, et al. Effects of liraglutide on visceral and ectopic fat in adults with overweight and obesity at high cardiovascular risk: a randomised, double-blind, placebo-controlled, clinical trial. *Lancet Diabetes Endocrinol* 2021; **9**(9): 595-605.
28. Flint A, Andersen G, Hockings P, et al. Randomised clinical trial: semaglutide versus placebo reduced liver steatosis but not liver stiffness in subjects with non-alcoholic fatty liver disease assessed by magnetic resonance imaging. *Aliment Pharmacol Ther* 2021; **54**(9): 1150-61.
29. Loomba R, Abdelmalek MF, Armstrong MJ, et al. Semaglutide 2·4 mg once weekly in patients with non-alcoholic steatohepatitis-related cirrhosis: a randomised, placebo-controlled phase 2 trial. *Lancet Gastroenterol Hepatol* 2023; **8**(6): 511-22.
30. Tamaki N, Munaganuru N, Jung J, et al. Clinical utility of 30% relative decline in MRI-PDFF in predicting fibrosis regression in non-alcoholic fatty liver disease. *Gut* 2022; **71**(5): 983-90.
31. Stine JG, Munaganuru N, Barnard A, et al. Change in MRI-PDFF and Histologic Response in Patients With Nonalcoholic Steatohepatitis: A Systematic Review and Meta-Analysis. *Clin Gastroenterol Hepatol* 2021; **19**(11): 2274-83.e5.