Hydration to optimize metabolism

Detailed Statistical Analysis Plan

Principal Investigator: Olle Melander, MD, PhD, Professor of Internal Medicine Lund University and Skåne University Hospital Malmö, Sweden

Protocol identification number: H2O Metab Proposal

NCT03422848, ClinicalTrials.gov identifier 2016894

Version: Draft 0.1, 2025-05-19

Authors: Olle Melander, MD, PhD; Sofia Enhörning, MD, PhD; Mats Pihlsgård, PhD, Statistician; all at Lund University

Signatures

Olle Melander
_____

Olle Melander, Principal Investigator

_____

Sofia Enhörning, co-Investigator

Mats Pihlsgård
_____

Mats Pihlsgård, Statistician

Contents

1 ABBREVIATIONS AND DEFINITIONS 3

2 INTRODUCTION 3

3 DESCRIPTION OF THE STUDY 3

3.1 Study objectives 3

3.1.1 Primary objectives 4

3.1.2 Secondary objectives 4

3.2 Study design 5

3.3 Sample size 6

3.4 Changes in the conduct of the study compared to the protocol 6

4 ANALYSIS SETS 7

4.1 Analysis sets 7

5 PLANNED ANALYSES 8

5.1 Status of subjects, protocol deviations and population definition 8

5.1.1 Status of subjects 8

5.1.2 Protocol deviations 9

5.1.3 Populations 9

5.2 Demography 9

5.3 Baseline characteristics 10

5.4 Study conducts parameters 11

5.5 Prior / Concomitant medications 11

5.6 Analysis of primary criteria 12

5.6.1 Description of primary outcome 12

5.6.2 Analysis of primary outcome parameters 12

5.7 Analysis of secondary outcomes/endpoints 12

5.7.1 Description of secondary criteria 13

5.7.2 Analysis of secondary outcome parameters 13

5.8 Analysis of safety 14

5.8.1 Adverse events 14

5.9 Analysis of drop-outs 14

6 STATISTICAL ANALYSES 14

6.1 Hypothesis and significance level 14

6.2 Distributions and normality assessment 15

6.3 Handling of missing data 15

6.4 Handling multiple testing 15

6.5 Outlier management 15

6.6 Subgroup analyses 15

6.7 Statistical computer software 16

6.8 Reporting of results 16

7 CHANGES IN THE SAP WITH RESPECT TO THE PROTOCOL 16

8 REFERENCES 16

1 Abbreviations and definitions

Acronyms

ACTH: Adrenocorticotropic Hormone

AE: Adverse Event

OGTT: Oral Glucose Tolerance Test

RCT: Randomized Control Trial

SAE: Serious Adverse Event

VP: Vasopressin

V3: Visit 3, baseline visit

V6: Visit 6

V8: Visit 8

2 Introduction

The study is a single-centre parallel group randomized clinical trial comparing the effect of a water intake supplementation intervention vs. a control group in subjects with a high plasma level of vasopressin (VP) (measured by a stable VP marker of its precursor hormone called "copeptin") on fasting levels of glucose (primary outcome measure), risk of new-onset diabetes and other cardiometabolic risk factors (secondary outcome measures).

The intervention group will increase its daily water intake by 1,5L on top of habitual water intake. The control group will not change its daily water intake.

No interim analyses are planned. The final analysis is performed at the end of the study using the final database lock (randomized and complete subjects are currently planned).

This Statistical Analysis Plan (SAP) describes the planned statistical analyses including methods used, derivations and outputs to be provided.

3 Description of the study

3.1 Study objectives

To study if water supplementation in subjects with high plasma levels of VP (measured by copeptin) can reduce fasting levels of glucose, risk of new-onset diabetes and other cardiometabolic risk factors in comparison with subjects with high VP who do not change their water intake.

3.1.1 Primary objectives

To assess if water supplementation during one year reduces fasting plasma glucose levels.

3.1.2 Secondary objectives

To assess if the water supplementation during one year alters

- the diabetes incidence

- the level of cardiometabolic risk factors (2-hour glucose during oral glucose tolerance test (OGTT), HbA1c, waist circumference, body mass index, systolic and diastolic blood pressure, serum triglycerides, HDL- and LDL cholesterol, Apo-B, Apo-A1, cortisol, Adrenocorticotropic hormone (ACTH), insulin (fasting and 2h post OGTT), C-reactive protein, estimated GFR, Creatinine clearance) and plasma concentration of copeptin.

3.1.3 Exploratory objectives.

Exploratory objectives are not part of primary or secondary outcomes and the exact analyses methods are not detailed here. To assess if water supplementation during one year leads to alterations in

- Reported diet intake. This objective is exploratory but reported diet intake regarding alterations in e.g. sugar sweetened beverages and adherence to Nordic Diet Recommendations 2023 can have an important mediating effect on both primary and secondary endpoints. Reported diet intake was collected using a software from the Swedish National Food Agency, and they declined continuation of the license from last quarter of 2023. Thus, we have missing data in subjects who had not completed their V8 before that time point. We estimate approximately 30% of missing data for the diet analyses (i.e. do not have complete data at baseline and V8). We plan exploratory analyses to test if water vs control therapy leads to differences in change of intake of main nutrients (protein, carbohydrates and fat energy intake relative to total caloric intake) over the 12-month period. These analyses will be performed using t-test (treatment group differences between delta values).

The first step of statistical analysis will be to produce descriptive statistics for some baseline parameters on reduced population (those with baseline and at least one post-baseline value), by intervention group and compare those results with the ITT population. Variables taken into account are those potentially impacting the primary outcome (fasting glycemia) and could be the sex (Male/Female), the age at enrolment (Years), the BMI at baseline (kg/m²), the total energy intake at baseline (kcal/day) from three non-consecutive days of 24-h dietary recall, if possible, 2-h plasma glucose level from OGTT at baseline (mmol/L) and the waist circumference (cm). If we notice a clinical imbalance between study groups (control and intervention groups in the reduced population vs. the ITT population) for those variables, they could be considered as potential covariates in sensitivity models for primary and secondary analysis. Three other variables could also be considered for the sensitivity analysis: newly diagnosed with T2DM (Yes/No), management of T2DM (initiation of intensive lifestyle therapy if possible and initiation of hypoglycemic drug therapy) and drug therapy impacting the fasting glycemia (Yes/No).

The second step will be driven by both previous results (the real number of missing data and potentially imbalanced covariates detected in the previous step). The idea is to perform sensitivity analyses for primary and secondary endpoints with dietary records (minimum baseline and one post-baseline value; maximum 3 data points per subject) including covariates that tend to cause imbalance between the intervention study groups.

- Gut microbiome composition. This is a second exploratory objective. Currently, we have only saved stool samples in the biobank, but since we do not yet have funding for microbiome profiling, it is unclear which method will be used (level of resolution i.e. 16S vs metagenomic shotgun sequencing) will require completely different statistics and bioinformatics. Thus it is not possible at this stage to provide a statistical analysis plan for this potential exposure. If any impacts of water treatment on primary or secondary outcomes are observed, and if funding allows measurement, a SAP - adjusted to the method of choice - will be written at a later stage.

- Plasma metabolome: Also for this third exploratory objective, funding is not currently available and range methodological resolution and cost is wide. Thus, like described for the microbiome, if funding allows measurement, a SAP- adjusted to the method of choice - will be written at a later stage.

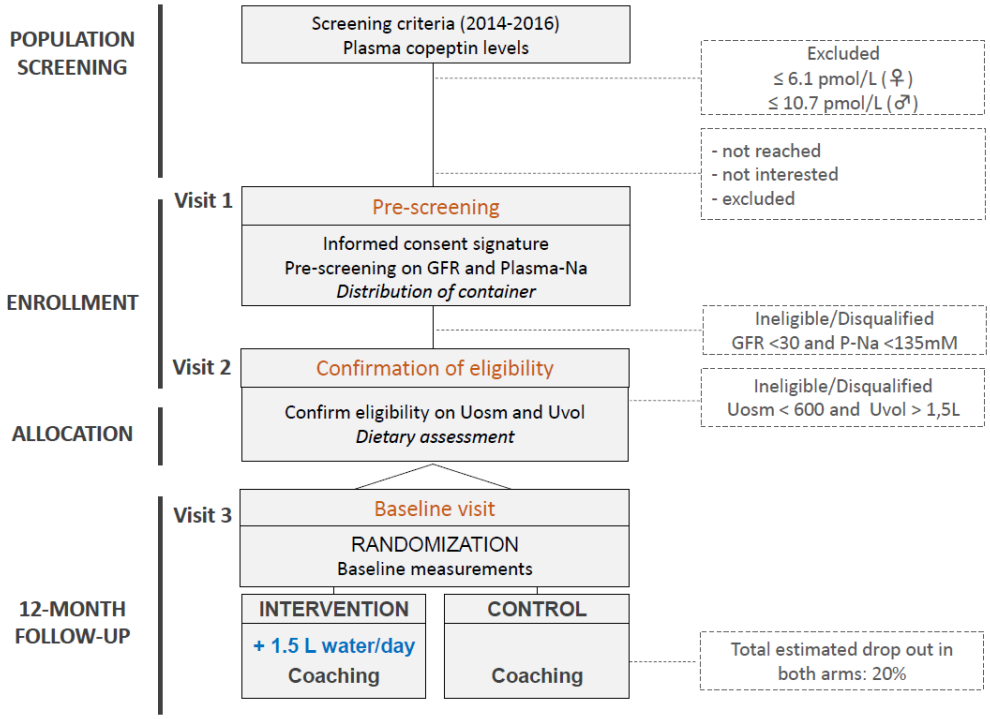
3.2 Study design

The study is a parallel-group RCT with two arms during 12 months. Subjects with the age of 20-75 years will be randomized to the water-intervention (1.5 L total on the top of habitual intake) and control group (1:1). The randomization will be stratified by gender to pursue equal distribution of intervention and control group for both male and female subjects. Both groups will receive general lifestyle advice (general oral and written advice on diet and physical activity). Reusable Bluetooth water bottles, which are volume sensitive and can be linked to an Android or iPhone application for individual monitoring purposes will be provided to subjects in the active treatment arm.

Study subjects will be recruited from 4 ongoing population studies in the Scania region encompassing altogether approximately 20 000 individuals within the current age span. Copeptin will be measured in -80 degrees frozen plasma samples from these 4 population studies. Individuals having a copeptin concentration of >6.1 pmol/L (women) or >10.7 pmol/L (men) will be invited to participate in the screening and inclusion process of this study. If fewer than expected will be recruited from these 4 studies, letters of invitation will be sent to inhabitants of the City of Malmö with surroundings, and inhabitants will also be invited through advertisements in local newspapers and social media. The same copeptin cut-off values will be used for inclusion to the study regardless of which invitation process that was used to include the participant.

All included individuals are low drinkers with high urine osmolality (Uosm ≥ 600) and low urine volume (Uvol ≤ 1,5L) as specified in the following figure.

Clinic visits are performed at 8 time points: visit 1, visit 2, baseline (visit 3, V3), 3 weeks and 3, 6 (visit 6, V6), 9 and 12 (visit 8, V8) months at which cardiometabolic risk factors and hydration parameters are measured (not all parameters at all time points). The study design is visualized in the Figure below.



3.3 Sample size

The primary outcome parameter is the difference between active (=extra water) and control treatment in the change of fasting plasma glucose between baseline and 12 months. We use prior effect estimates from the largest RCT for diabetes prevention study in Europe, i.e. the Finnish Diabetes Prevention Study (FDPS) (New Engl J Med 2001;344:1343-50), which compared individual life style counselling (active treatment) with general oral and written life-style advice (control treatment) in relation to risk of new onset type 2 diabetes and change of plasma glucose concentration and found a 58% decreased relative risk of diabetes. After 12 months, the fasting glucose in the active treatment group was reduced by 4±12 mg/dL vs a 1±12 mg/dL increase in the control group with the mean difference of 5 mg/dL of the 12-month change of fasting glucose being highly significant (P<0.001).

To obtain sufficient statistical power to detect a clinically significant effect size, we base the power calculation on an effect of water vs control that is at least 50% of what is considered an epoch changing effect of lifestyle i.e., the difference observed in the FDPS, while assuming the standard deviation for the change (s = 0.67 mmol/L) as observed in FDPS. To be able to detect ≥50% of that effect (a difference of 2.5 mg/dL = 0.14 mmol/L between treatments in 12-month change) we need 319 evaluable subjects in both the active and control treatment group at a power of 80% and a 2-tailed significance level of =0.05, a total of 638 evaluable subjects. As the glucose measurement in the study used a more recent and exact method, the standard deviation of the change of glucose was recalculated by the DSMB in October 2022 from the first completed 357 study participants (s = 0.58 mmol/L). With this smaller standard deviation, the power (with all other assumptions unchanged) was 270 evaluable subjects per group, i.e. n=540 evaluable subjects in total. When also taking into account a dysbalance (with slightly higher drop-out rate in subjects in the intervention group), a total of 544 evaluable subjects were estimated to be needed for 80% power.

3.4 Changes in the conduct of the study compared to the protocol

The key compliance/adherence measure during the study is 24-hour urine measurement of volume and osmolality. After the start of study treatment at V3, 24-hour urine is collected at V4 (3 weeks), V5 (3 months), V6 (6 months), V7 (9 months) and V8 (12 months). Apart from the main "intention to treat" analysis, we will perform "per protocol" analysis in which we exclude subjects with major protocol deviations (i.e. whose values of 24-hour urine displays major deviation from expected, given treatment assignment, with those limits defined below (see protocol deviations)).

4 Analysis sets

4.1 Analysis sets

-Included subjects

Urine collection over 24 hours is collected from participants in previous population studies and from subjects living within the greater Malmö area who respond to advertising. The analysis set is the number of subjects who fulfill all inclusion criteria (including urine volume, osmolality and plasma copeptin) and are randomized to treatment allocation after V2.

At the following "baseline visit" (V3), baseline fasting metabolic measurements are done and after the baseline visit (V3), the actual active and control treatment commence. This design is the only one practically possible as study subjects need to perform the baseline measurements in the morning in the fasted state. However, of the subjects included and randomized to treatment allocation at V2 some are expected to not come to the V3 fasting baseline examination and thus never commence active or control treatment. Furthermore, of all subjects starting intervention after V3, we will probably observe a dropout during the study.

Intention to treat

The "Intention-to-Treat" population: all randomized study subjects with at least two during study fasting plasma glucose measurements will be analyzed in intention to treat analysis (i.e. regardless of any protocol deviations). This analysis set includes the evaluation of primary and secondary outcomes as well as adverse events (AE) and serious adverse events (SAE) between intervention and control group.

Per protocol

The key protocol deviation that can influence the therapy tested (increased water intake vs control) is that the study subjects do not adhere to the assigned therapy (do not increase water intake in active group / increase water intake in control group). Generally, unless a study treatment is directly administered by study personnel (for example short-term intravenous pharmacological infusions in the inpatient setting) it is impossible to fully control the adherence to the study therapies as study subjects administer the study therapies themselves at home. This limitation is thus the same as in randomized drug trials, where "pill counting" is the most common way to assess compliance/adherence. Whereas the "intention to treat" analysis, which is the primary method, tests if the assigned therapy works regardless of adherence, the secondary "per protocol" analyses will take into account the best available measure of compliance to the current therapy, which is changes of 24-urine measures during follow-up in comparison with baseline. In these per protocol analyses, we will exclude study subjects who, based on values of their 24-hour urine, show indications of poor compliance.

The exclusion of patients let to define the Per-Protocol Set will be based on the urine volume and concentration. These parameters will be considered as the main protocol deviation criteria. Cut-offs for inclusion in the per protocol analysis are at least one of the two criteria below met for 24h UOsm OR UVol:

Delta 24h UOsm (Baseline vs. 3 months, 6 months, 9 months and 12 months):

- Control Group: Those with a decrease in delta 24h UOsm not greater than 25% from baseline.
- Intervention group: Those with a decrease in delta 24h UOsm of at least 50% from baseline.

OR Delta 24h UVol (Baseline vs. 3 months, 6 months, 9 months and 12 months):

- Control Group: Those with an increase in delta 24h UVol not greater than 500 mL from baseline.
- Intervention group: Those with an increase in delta 24h UVol of at least 1000 mL from baseline.

Furthermore, all individuals attending a clinic study visit (V3, V6 or V8) in a non-fasted state will be excluded from the per protocol analyses. Some other protocol deviations let to exclude patients are detailed in Section 5.1.2

Missing data

Missing data on outcomes (i.e. change of fasting glucose and the secondary cardiometabolic measures) will not be imputed, so subjects with missing data on outcome parameters will not be included in analysis of the respective outcome. For each of the outcomes, the maximum number of subjects will be included (e.g. if there are n=544 evaluable subjects for fasting glucose change, all of these will be included even if there are less evaluable subjects for any other cardiometabolic outcome parameter).

Given the randomized nature of the study, we will not perform any adjustments (which in turn can be affected by missing data on the covariates adjusted for) with three exceptions: (1) As detailed in "statistics" we will include adjustment for the baseline value of the outcome parameter in question (e.g. the change of glucose outcome will be adjusted for the baseline value of glucose). The motif of this adjustment is to minimize bias related to regression towards the mean. (2) As the randomization is stratified by sex, this covariate will be included in all outcome analyses. (3) If clinically significant/meaningful treatment group differences appear in baseline values of age, BMI or diabetes status in any outcome analysis, we will add a sensitivity outcome analysis in which we adjust for such baseline differences. Any missing values for these covariates will be imputed to obtain the same number of evaluable subjects in non-adjusted and adjusted analyses. The technique we propose is Multiple Imputation by Chained Equations (MICE) with subsequent application of Rubin's rule to combine results from different imputations. We will base the analyses on 20-30 imputed data sets. For this analysis we will use SAS PROC MI in combination with PROC MIXED/PROC GLM and PROC MIANALYZE.

5 Planned analyses

5.1 Status of subjects, protocol deviations and population definition

5.1.1 Status of subjects

The number of subjects will be provided for the following groups:

- Subjects who attended V2 and fulfilled low water intake criteria (based on urine volume, urine osmolality and plasma copeptin) were randomized. We will thus detail numbers of subjects randomized (total and in intervention and control groups, respectively).
- Of those who were randomized at V2, not all appeared to visit 3, where a baseline clinical examination was performed and therapy (intervention or control) commenced. Thus, number

(total and per intervention/control group) who commenced therapy will be detailed next, including the number and percentage who dropped out between randomization (V2) and baseline exam/start of study (V3) in total and per study treatment group.

- We will detail the numbers who were present at V3, V6 and V8 (totally and per study treatment group.
- We will detail the numbers who completed the study for the intention to treat analysis (i.e. subjects randomized who underwent baseline exam/commenced the study treatment and had at least one additional measurement of fasting glucose (at V6 or V8), in total and per study treatment group.
- Of randomized subjects who attended baseline exam and commenced study treatment (i.e. attended V3), we will detail the number (total and per study treatment group) who dropped out after V3 without at least one additional study fasting glucose measurement (i.e. not completed subjects).

The current numbers at early visits (V2-3) and preliminary numbers on late visits (V6) are given in the section above.

5.1.2 Protocol deviations

So far identified protocol deviations are:

Main protocol deviation:

The main protocol deviation is lack of compliance to active or control therapy, for which assessment of 24-hour urine is the best indication. We will address this in per protocol analyses, please see details in section above (part 4.1).

Other protocol deviations:

*Attendance at clinic study visit outside the predefined time windows. The time between visits summary statistics will be shown but will not result in any per protocol analyses.

* Attendance at clinic study visit V3, V6 or V8 in a non-fasted state. These participants will be excluded from the per protocol analyses.

Thus, we will detail the numbers of subjects (total and per study treatment group) who had any protocol deviation for these respective reasons, or for other reasons not yet encountered.

5.1.3 Populations

We expect a final study sample for intention to treat analysis of at least n=544 evaluable study subjects who have completed the protocol (i.e. with at least 2 fasting glucose measurements, meaning at V3 and at either V6 or V8 or at both V6 and V8). We expect a dysbalance in the final study sample (46% intervention vs 54% control therapy) due to dysbalance in the drop-out rate between intervention and control groups.

5.2 Demography

Variable name and unit	Variable source[1]	Type[2]	Unit	Set [3]
Sex	R	N	male/female	ITT, PP

Age	R	C	years	ITT, PP

1. *(R)aw or (D)erived data*
2. *(C)ontinuous or (O)rdinal or (N)ominal*
3. *Intention to treat (ITT) or Per Protocol (PP)*

5.3 Baseline characteristics

Variable name and unit	**Variable source**[1]	**Type**[2]	**Set**[3]
SURVEY			
Diabetes mellitus	R	N	ITT, PP
Hypertension	R	N	ITT, PP
Cardiovascular disease	R	N	ITT, PP
Cancer	R	N	ITT, PP
CLINICAL			
Height (cm)	R	C	ITT, PP
Weight (Kg)	R	C	ITT, PP
BMI (Kg/m2)	D	C	ITT, PP
Waist circumference (cm)	R	C	ITT, PP
Systolic blood pressure (mm Hg)	R	C	ITT, PP
Diastolic blood pressure (mm Hg)	R	C	ITT, PP
Body fat%	R	C	ITT, PP
BLOOD			
Fasting glucose (mmol/L)	R	C	ITT, PP
Fasting and 2h glucose (during OGTT) (mmol/L)	R	C	ITT, PP
Fasting and 2h insulin (during OGTT) (mIE/L)	R	C	ITT, PP
Fasting HbA1c (mmol/mol)	R	C	ITT, PP
Fasting cortisol (nmol/L)	R	C	ITT, PP
Fasting ACTH (pmol/L)	R	C	ITT, PP
Fasting copeptin (pmol/L)	R	C	ITT, PP
Fasting Apo A1 (g/L)	R	C	ITT, PP
Fasting ApoB (g/L)	R	C	ITT, PP
Fasting Triglycerides (mmol/L)	R	C	ITT, PP
Fasting LDL (mmol/L)	R	C	ITT, PP
Fasting HDL (mmol/L)	R	C	ITT, PP
C-reactive protein (mg/L)	R	C	ITT, PP
Creatinine (μmol/L)	R	C	ITT, PP
eGFR (mL/min/1,73 m^2)	D	C	ITT, PP
Sodium (mmol/L)	R	C	ITT, PP
Potassium (mmol/L)	R	C	ITT, PP
Urea (mmol/L)	R	C	ITT, PP
Osmolality (mOsm/kg)	R	C	ITT, PP
Erytrocyte Volume Fraction (%)	R	C	ITT, PP
URINE			
Volume (ml)	R	C	ITT, PP
Osmolality (mosm/kg H_2O)	R	C	ITT, PP
Creatinine (mmol/24h)	R	C	ITT, PP

Sodium (mmol/24h)	R	C	ITT, PP
Potassium (mmol/24h)	R	C	ITT, PP
Urea (mmol/24h)	R	C	ITT, PP
Albumin/creatinine ratio (g/mol)	R	C	ITT, PP

1. *(R)aw or (D)erived data*
2. *(C)ontinuous or (O)rdinal or (N)ominal*
3. *Intention to treat (ITT) or Per Protocol (PP)*

5.4 Study conducts parameters

Variable name	Variable source(1)	Type(2)	Unit	Set(3)
Compliance parameters (markers of urine concentration)				
V6 vs V3 urine osmolality difference (paired test)	D	C	mosm/kg	ITT, PP
V8 vs V3 urine osmolality difference (paired test)	D	C	mosm/kg	ITT, PP
V6 vs V3 urine volume difference (paired test)	D	C	mosm/kg	ITT, PP
V8 vs V3 urine volume difference (paired test)	D	C	mosm/kg	ITT, PP

1. *(R)aw or (D)erived data*
2. *(C)ontinuous or (O)rdinal or (N)ominal*
3. *Intention to treat (ITT) or Per Protocol (PP)*

Study conduct parameter derivations:

Study duration (days): time between the last visit (V8) and randomization visit
= (Date of V8 – Date of V3) + 1 will be detailed.

Compliance to study urine volume (see paragraph 5.4 above): This will be used to assess compliance in per protocol analysis.

Time between study visits (days): time between each study visit planned in the study.
= (Date of Vy – Date of Vx) +1 with y>x. It will be calculated between V3 – V6, V6 – V8 and V3 – V8 and presented.

5.5 Prior / Concomitant medications

Variable name	Variable source(1)	Type(2)	Unit	Set(3)
Concomitant medications of relevance at baseline (visit 3)				
Antihypertensive medications [a]	R	N	Yes/no	ITT, PP
Oral anti-diabetic medications [b]	R	N	Yes/no	ITT, PP
Lipid lowering medications [c]	R	N	Yes/no	ITT, PP
[a] Defined as use of medications included in ATC-groups C02, C03, C07, C08 or C09.				

[b] Defined as use of medications included in ATC-group A10 [c] Defined as use of medications included in ATC-group C10

1. *(R)aw or (D)erived data*
2. *(C)ontinuous or (O)rdinal or (N)ominal*
3. *Intention to treat (ITT) or Per Protocol (PP)*

5.6 Analysis of primary criteria

5.6.1 Description of primary outcome measure

The primary outcome measure is between group comparison of the change from baseline to 12-month of fasting plasma glucose.

Variable name	Variable source[(1)]	Type[(2)]	Unit	Set [(3)]
Change of fasting plasma glucose from baseline to 6-and 12-months	D	C	mmol/L	ITT, PP

2. *(R)aw or (D)erived data*
4. *(C)ontinuous or (O)rdinal or (N)ominal*
5. *Intention to treat (ITT) or Per Protocol (PP)*

5.6.2 Analysis of primary outcome parameters

The primary analysis (intention-to-treat) will compare the between study groups (intervention vs control treatment) LS-means differences on the changes from baseline to 12 months and from baseline to 6 months in fasting plasma glucose. The main analysis will be based on a linear mixed model with fasting plasma glucose change from baseline to study visits where it is measured during the study (6 and 12 months) as dependent variable. This model produces unbiased estimates under the assumption that missing data are missing at random. The model will be adjusted (through fixed effects) for sex (because of the stratified randomization), baseline value of fasting plasma glucose, treatment, study visit and interaction between treatment and study visit. We will in the mixed regression analysis model within-patient correlation to decrease exponentially with number of days between measurements or through an AR(1) model. The choice of correlation structure specifies the variance-covariance matrix in the mixed regression model. The choice of covariance structure rests on principles of parsimony, simplicity and prior beliefs, rather than post-analysis testing. Model assumptions will be assessed through visual inspection of plots of suitable regression residuals. The estimated LS-mean differences between treatment levels for the change in fasting plasma glucose from baseline to 12 months and baseline to 6 months, the corresponding 95% confidence intervals (CIs) and p-values will be provided. The same analyses will be conducted on the imputed data set (see "Missing Data" sub-section under section 4.1) and the per protocol data set. The analyses will be conducted using SAS PROC MIXED.

Sensitivity analysis as described in Section 4.1, part "Missing Data" could be performed. Conclusions regarding the inclusion of baseline characteristics in the linear mixed model will be based on clinical significances.

5.7 Analysis of secondary outcomes/endpoints

The secondary outcome measures are between group comparison of the change from baseline to 12 months of diabetes incidence and cardiometabolic risk factors.

5.7.1 Description of secondary criteria

Variable name and unit	Variable source[1]	Type[2]	Set [3]
Diabetes			
Diabetes incidence after baseline visit during 12 months	R	N	ITT, PP
Change from baseline to 6 and 12 months of cardiometabolic risk factors			
Fasting and 2h glucose (during OGTT) (mmol/L)	D	C	ITT, PP
Fasting and 2h insulin (during OGTT) (mIE/L)	D	C	ITT, PP
Fasting HbA1c (mmol/mol)	D	C	ITT, PP
Fasting cortisol (nmol/L)	D	C	ITT, PP
Fasting ACTH (pmol/L)	D	C	ITT, PP
Fasting Apo A1 (g/L)	D	C	ITT, PP
Fasting ApoB (g/L)	D	C	ITT, PP
Fasting Triglycerides (mmol/L)	D	C	ITT, PP
Fasting LDL (mmol/L)	D	C	ITT, PP
Fasting HDL (mmol/L)	D	C	ITT, PP
C-reactive protein (mg/L)	D	C	ITT, PP
Creatinine (µmol/L)	D	C	ITT, PP
eGFR (mL/min/1,73 m^2)	D	C	ITT, PP
BMI (Kg/m2)	D	C	ITT, PP
Waist circumference (cm)	D	C	ITT, PP
Systolic blood pressure (mm Hg)	D	C	ITT, PP
Diastolic blood pressure (mm Hg)	D	C	ITT, PP
Body fat%	D	C	ITT, PP
Other blood laboratory parameters			
Fasting copeptin (pmol/L)	D	C	ITT, PP
Urine laboratory parameters			
Albumin/creatinine ratio (g/mol)	D	C	ITT, PP

1. *(R)aw or (D)erived data*
2. *(C)ontinuous or (O)rdinal or (N)ominal*
3. *Intention to treat (ITT) or Per Protocol (PP)*

5.7.2 Analysis of secondary outcome parameters

New-onset diabetes is defined as 2 consecutive fasting plasma glucose values ≥7.0 mmol/L or 2-hour post OGTT value of ≥11.0 mmol/L after the baseline exam (V3). In addition, new onset diabetes is considered present if diagnosed by a physician outside of the current study, as assessed in the questionnaire (answering yes on having had a physician diagnosis of diabetes or having been prescribed

medication for diabetes) after the baseline exam (V3). All individuals with a diagnosis of diabetes at or before the baseline exam (V3) will be excluded in the analysis of incidence of diabetes.
The difference in diabetes incidence between the intervention and control group will be analyzed using logistic regression with diabetes incidence as the binary outcome and treatment arm as exposure variable and reported as the odds ratio (OR) including 95 % CI for diabetes incidence in intervention vs control. As the randomization should lead to no difference in diabetes risk factors, the main analysis will only be adjusted for sex (as the randomization is stratified by sex, this covariate will be included in all outcome analyses), but we will also add a sensitivity analysis to adjust for key diabetes risk factors at baseline (fasting glucose, diabetes status, BMI, age), if there is a clinically significant difference in these factors between treatment groups despite the randomization.

Cardiometabolic risk factors will be analyzed using the same method as for the primary outcome, but with baseline value for the actual factor as a covariate instead of baseline value of fasting plasma glucose.

In addition, to assess the possible association between change in fasting plasma glucose levels, copeptin, other cardiometabolic risk factors, other blood laboratory parameters and urine laboratory parameters, we propose to analyze correlation(s) between delta value of outcome parameters and delta value of the efficacy parameters (e.g. urine volume, urine osmolality and plasma copeptin) with Pearson or Spearman correlation analyses.

- Outcome parameters are fasting plasma glucose and all secondary continuous outcome parameters listed in section 5.7.1 above.

We emphasize that in these correlation analyses, copeptin will be regarded both as an outcome parameter and as an efficacy parameter.

Analysis of safety

5.7.3 Adverse events

See section "*Definition and registration of adverse events*" of the protocol. AE and SAE will be reported for all subjects who attended baseline exam and commenced study treatment (i.e. attended V3). AE and SAE will be reported as a diagnosis code (http://icd.internetmedicin.se/) associated with the AE/SAE, and the incidence of AE and SAE will be compared between treatment groups (intervention vs control).

5.8 Analysis of drop-outs

Of randomized subjects who attended baseline exam and commenced study treatment (i.e. attended V3), we will detail the number (total and per study treatment group) who completed vs dropped out after V3. The definition of drop-outs are subjects attending V3 but dropped out without at least one additional study fasting glucose measurement either at V6 or V8. Baseline characteristics for the intention to treat population will be provided for both intervention and control subjects and will be separated by drop-outs.

6 Statistical analyses

6.1 Hypothesis and significance level

- The primary and secondary outcomes will be tested using two-sided tests (P<0.05) with the hypothesis that there will be a significant difference in the change of the outcome parameter between intervention and control group at P<0.05 level in a 2-sided test.

6.2 Distributions and normality assessment

For normally distributed data we will use parametric tests. If a variable deviates from normal distribution, we will log transform the variable. If the variable still deviates from normal distribution, we will use a non-parametric test.

6.3 Handling of missing data

When running sensitivity analyses addressing unbalanced (despite randomization) distributions of baseline age, BMI or diabetes status for the ITT population, model based multiple imputation (MI) will be used for these covariates. The number of imputations will be 20-30 and we will apply MICE (Multiple Imputation by Chained Equations) to perform the imputations. Trace plots and distribution plots will be created to check the accuracy of the imputations.

Due to the balanced dropout between V2 and the baseline visit (V3) and the slightly dysbalanced dropout during the study (as indicated by preliminary data from September 2024 regarding completed V6), we will in a post-hoc analysis in an experimental fashion fill the gaps in data according to different scenarios for what the value of an unobserved measurement would have been had it been observed. In other words, we will create different data sets based on different more or less pessimistic assumptions regarding the values of unobserved measurements. These data sets can subsequently be analyzed to see what happens with treatment effects under the different scenarios.

6.4 Handling multiple testing

No multiplicity adjustment strategy is planned.

6.5 Outlier management

For analysis on the primary and secondary outcome parameters only: in case the data review meeting and/or database explorations reveal outliers, as identified by independent experts and/or statistical and/or Medical Reviewer outlier identification, a sensitivity analysis could be performed excluding these outliers. Reasons for outlier classification and results of these sensitivity analyses will be described separately in the study report.

6.6 Subgroup analyses

Six subgroups will be analyzed. All subgroups will be analyzed using both ITT and PP populations.

High-high

All randomized study subjects having copeptin concentration above the previously specified cut-off values, 6.1 pmol/L (women) or 10.7 pmol/L (men), at both population screening and baseline visit.

Top tertile
All randomized study subjects having copeptin concentration in the top tertile (gender specific) at baseline visit.

Diabetes mellitus
All randomized study subjects will be divided into two subgroups according to diabetes mellitus status at baseline visit.

Impaired fasting glucose
All randomized study subjects will be divided into two subgroups according to fasting glucose status (diabetes+impaired/not impaired) at baseline visit.

Gender
All randomized study subjects will be divided into two subgroups according to gender.

6.7 Statistical computer software

SAS 9.4 will be used for all statistical analysis.

6.8 Reporting of results

Results to be shown in the data report are detailed above.

7 Changes in the SAP with respect to the protocol

SAP Section	Change	Rationale
3.3	Sample size	The reason for the recalculation is that the instruments used are more accurate in the measurements than the instruments used in the Finnish Diabetes Prevention Study (FDPS). The pooled standard deviation of delta fasting plasma glucose (mmol/L) at month 12 is 0.58.

8 References

1. Verbeke G, Molenberghs G. Linear Mixed Models for Longitudinal Data. 2000.
2. Gadbury et al. Modern statistical methods for handling missing repeated measurements in obesity trial data: beyond LOCF. Obesity Reviews 2003.
3. Meinert Larsen et al. Diets with high or low protein content and glycemic index for weight-loss maintenance. NEJM. 2010.
4. Bell et al. Differential dropout and bias in randomised controlled trials: when it matters and when it may not. BMJ 2013.