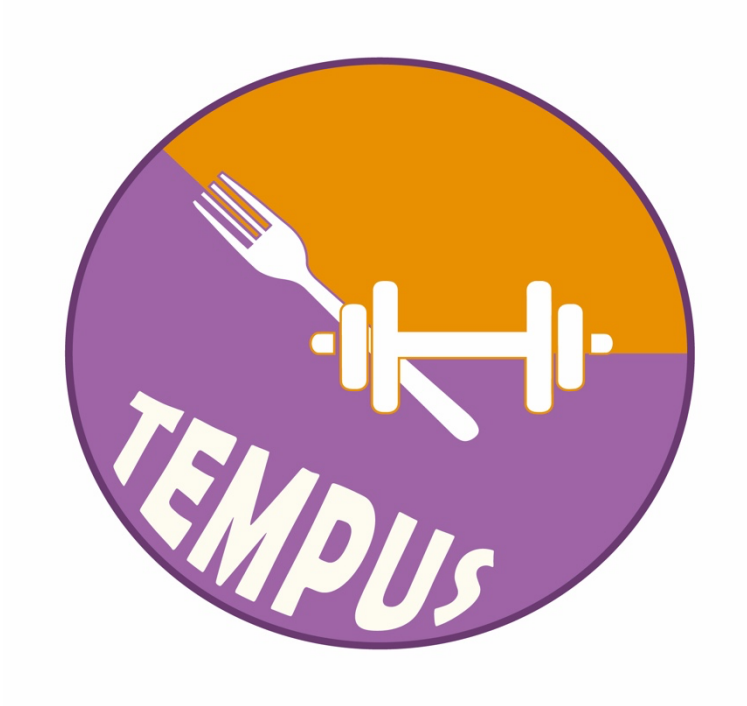
# TEMPUS




# Statistical Analysis Plan

University of Granada, Spain
Trial registration number: NCT05897073
Version: 1
Date: 01/07/2023

**Statistical analysis plan (SAP) revision history**: no revisions

**Roles and responsibility**:

Name and address of principal investigator and trial coordinators:

Prof. Dr. Jonatan R. Ruiz (principal investigator)
Dr. Alba Camacho-Cardenosa (trial coordinator)
Mr. Antonio Clavero-Jimeno (trial coordinator)
PROFITH (PROmoting FITness and Health Through Physical Activity) research group, Department of Physical Education and Sports, Sport and Health University Research Institute (iMUDS), University of Granada
Faculty of Sport Sciences, Carretera de Alfacar s/n, Granada, 18071, Spain

**Name of person writing SAP:**

E-mail principal investigator: ruizj@ugr.es
E-mails trial coordinators: acamachocardenos@ugr.es & claveroa@ugr.es
PROFITH research group, Department of Physical Education and Sports, Faculty of Sport Sciences, Sport and Health University Research Institute (iMUDS), University of Granada, Granada, Spain

**Signatures**:

| | |
|---|---|
| Prof. Dr. Jonatan R. Ruiz (principal investigator) | Date: 01/07/2023 |
| Dr. Alba Camacho-Cardenosa (trial coordinator) | Date: 01/07/2023 |
| Mr. Antonio Clavero-Jimeno (trial coordinator) | Date: 01/07/2023 |

# Table of contents

*1. General information of TEMPUS study* ..... *3*

**1.1. Background and rationale** ..... **3**

**1.2. Objective** ..... **3**

*2. Study methods* ..... *4*

**2.1. Trial design** ..... **4**

**2.2. Randomization** ..... **4**

**2.3. Sample size** ..... **4**

**2.4. Framework** ..... **5**

**2.5. Statistical interim analyses and stopping guidance** ..... **5**

**2.6. Timing of the final analyses** ..... **5**

**2.7. Timing of outcome assessment** ..... **5**

*3. Statistical principles* ..... *5*

**3.1. Confidence and P values** ..... **5**

**3.2. Adherence, attendance, and protocol deviations** ..... **6**

**3.3. Analysis populations** ..... **6**

*4. Trial population* ..... *6*

**4.1. Screening data** ..... **6**

**4.2. Eligibility** ..... **6**

**4.3. Recruitment** ..... **7**

**4.4. Withdrawal/follow-up** ..... **8**

**4.5. Baseline patient characteristics** ..... **8**

*5. Analysis* ..... *8*

**5.1. Outcome definitions** ..... **8**

**5.2. Analysis methods** ..... **13**

**5.3. Missing data** ..... **14**

**5.4. Additional analyses** ..... **14**

**5.5. Harms** ..... **14**

**5.6. Statistical software** ..... **14**

*6. References* ..... *14*

# 1. General information of TEMPUS study

## *1.1. Background and rationale*

Metabolic dysfunction-associated steatotic liver disease (MASLD) is a major public health concern due to its high prevalence and strong association with extrahepatic conditions, particularly among adults with obesity, in whom it often remains undiagnosed. Emerging approaches such us time-restricted eating (TRE)—a novel form of intermittent fasting—and supervised exercise may represent promising strategies to mitigate associated health risks. Preliminary evidence suggests that combining TRE and exercise may be an effective strategy for improving body composition and cardiometabolic health in adults with obesity. However, studies examining the combined effects of TRE and exercise in humans are scarce and have notable limitations.

## *1.2. Objective*

The overall objective is to investigate the effects of TRE combined with supervised exercise on hepatic fat content (primary outcome) and cardiometabolic health (secondary outcomes) in adults with obesity.

The primary objective is to investigate the effects of a 12-week TRE intervention combined with a supervised exercise intervention (TRE + Exercise), as compared with TRE (TRE) or supervised exercise alone (Exercise), and usual care control group (UC), on hepatic fat content (primary outcome) and cardiometabolic health (secondary outcomes) in adults with obesity.

Secondary aims are to:

i. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on other specific fat depots: abdominal visceral and subcutaneous adipose tissue, pancreatic fat and mid-thigh intermuscular and intramuscular adipose tissue, and body composition (fat mass and fat-free mass).
ii. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on cardiometabolic health markers including blood pressure, liver function enzymes, glucose homeostasis, blood lipid profile, and circulating inflammatory biomarkers.
iii. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on fecal microbiota.
iv. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on sleep and physical activity patterns.
v. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on psychological outcomes (i.e., depression, stress and anxiety) and quality of life.
vi. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on physical fitness (i.e., cardiorespiratory fitness and muscular strength).
vii. Examine the effect of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on dietary habits and eating behaviour.
viii. To quantify the persistence of the effects of TRE + Exercise intervention, TRE alone, and

supervised exercise alone on hepatic fat content and cardiometabolic health in the long-term (12-month follow-up).

## 2. Study methods

### *2.1. Trial design*

The TEMPUS project is an investigator initiated, parallel group, randomized controlled trial (RCT). The RCT will include four-arms containing a UC control group and three intervention groups that will receive three different 12-week TRE and/or supervised exercise programs (**Figure 1**): i) TRE alone, ii) Exercise alone and iii) TRE + Exercise. The study will be conducted in ~184 patients (50% women) with obesity who meet the eligibility criteria. Measurements will be performed at baseline, at midpoint (6 weeks), post intervention (12 weeks) and 12-month follow-up (12 months post intervention). More detailed information is described in the study protocol.[1]

The TEMPUS research team will perform the data quality control, data processing, writing analyses scripts/programs and statistical analysis.

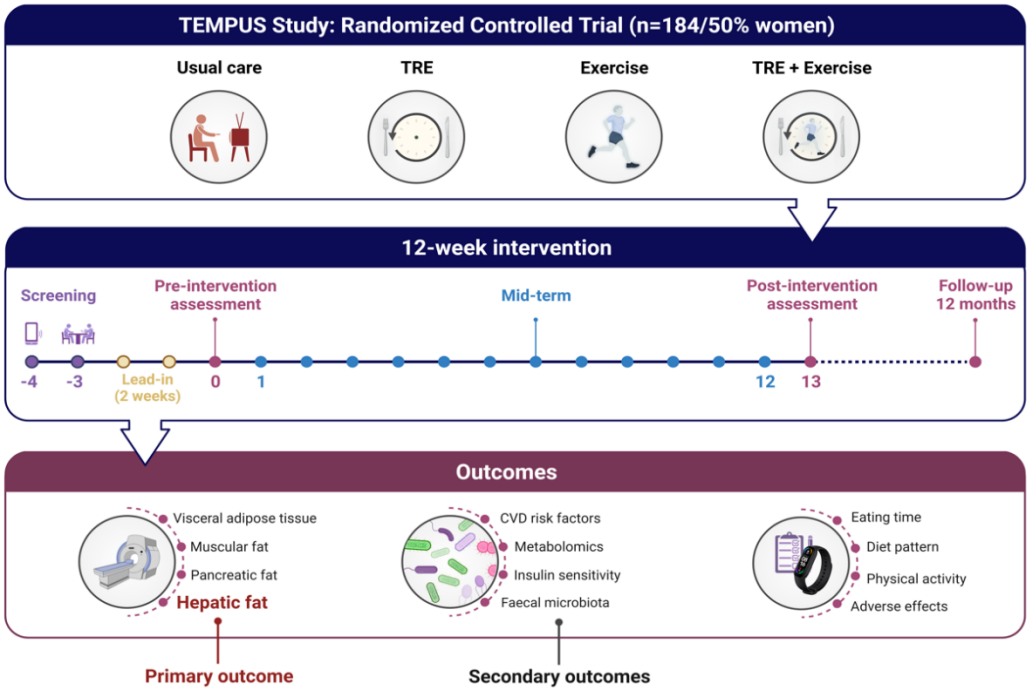



**Figure 1.** Overview of the TEMPUS Study. TRE, time-restricted eating.

### *2.2. Randomization*

Participants will be randomized using a parallel design with a 1:1:1:1 allocation ratio and stratified for sex.

### *2.3. Sample size*

Based on previous findings from a recent trial on the combination of alternate day fasting and exercise on hepatic fat content,[2] we anticipate approximately 5.0% reduction in hepatic fat content in the TRE + Exercise group, 2.5% in the TRE group, 2.5% in the exercise group, and

no significant change in this outcome in the usual-care group. Assuming a pre-post correlation of 0.8 and a standard deviation of 6 points in the main outcome, we estimate a medium effect size of 0.45. To detect this effect size as statistically significant in a one-way ANOVA with α = 0.05 and a power of 0.8, a minimum of 19 patients per group is required. Accounting for subgroup analyses based on sex and a maximum dropout rate of 20%, we will aim to recruit ~46 participants for each trial group, resulting in a total sample size of ~184 participants, with ~92 women included.

Our sample size calculations have taken into account subgroup analyses by sex. We have conservatively estimated a maximum dropout rate of approximately 20%. By considering this dropout rate, we have ensured that our study is adequately powered to detect the specified effect size even if there are differential dropout rates between men and women. For example, if men have a dropout rate of 5% and women have a dropout rate of 15%, our study will still have sufficient power to analyze the data separately in men and women.

### *2.4. Framework*

A superiority hypothesis testing framework will be used. In the main analyses, we will compare whether TRE plus exercise interventions (TRE + Exercise group) is superior to TRE or Exercise alone or UC (Mediterranean education program only).

### *2.5. Statistical interim analyses and stopping guidance*

No pre-specified interim analyses are performed, and therefore, a stopping guidance will not be applicable.

### *2.6. Timing of the final analyses*

The final analyses will be performed after the finalization of the data collection and processing of the primary and secondary outcomes.

### *2.7. Timing of outcome assessment*

The primary (i.e., hepatic fat content) and secondary outcomes (i.e., other specific fat depots measured by magnetic resonance imaging [MRI], and cardiometabolic health markers) will be assessed at baseline, post intervention (12 weeks), and at a 12-month follow-up (12 months post intervention). The time frame of each outcome is explained in section '*5.1 Outcome definitions*'

## 3. Statistical principles

### *3.1. Confidence and P values*

All statistical tests will be two-tailed. *P* for significance will be set at 0.05 and 95% confidence intervals will be estimated. No adjustment for multiplicity will be made as multiplicity adjustments may be of lesser importance in the case of distinct treatment arms.[3] Moreover, only one primary outcome will be defined, and other outcomes (i.e., secondary outcomes and exploratory analyses) will not require adjustment for multiple testing. However, we will include a hierarchical analytic approach for the interventions tested and the primary outcome.

*3.2. Adherence, attendance, and protocol deviations*

During the 12-week intervention period, participants will be required to record their eating times (i.e., exact times of the beginning of the first meal and of the end of the last meal) in a mobile phone app specifically designed for the study. These data will be revised from 2 to 3 times every week, asking the participant for missing records, and will provide insights into their adherence to the prescribed eating window. For participants in the TRE + Exercise group and TRE alone group, each day will be labeled as adherent if the participant meets their self-selected 8-hour (±30 minutes) eating window. Adherence will then be defined as the percentage of days on which the eating window was met. Finally, we will assess the long-term adherence to the intervention (at the 12-month follow-up). This will allow us to evaluate the sustainability of participants' adherence over an extended period.

Attendance will be defined as the percentage of exercise sessions attended by the participant (TRE + Exercise and Exercise alone groups), recorded by the trainers, divided by the actual number of exercise sessions offered (12 weeks x 2 sessions / week = 24).

All protocol deviations made to the protocol (e.g., change in pre-defined inclusion/exclusion criteria, baseline and post assessments, data cleaning/processing) will be reported and described.

*3.3. Analysis populations*

We plan to use an intention-to-treat approach for our primary analyses, which will include all randomized participants. With this approach, participants are analysed in the groups to which they were originally assigned, regardless of whether they completed the intervention. As a result, some participants will have valid data at both time points, while others may have missing data at baseline, post-intervention, or the 12-month follow-up assessment.

## 4. Trial population

*4.1. Screening data*

Screening data will be based on patients who are defined as eligible by the clinical and research team. Screening is performed in three phases:

1) Screening based on a phone call by the research team.
2) Screening based on the eligibility criteria (see '*4.2 Eligibility'*) by the clinicians of the research team.
3) Screening during the 2-week lead-in period prior to pre-intervention assessments and group allocation.

More details on the eligibility criteria are provided in '*4.2 Eligibility*'. The number of patients within each phase will be reported.

*4.2. Eligibility*

Eligible patients will be defined by the inclusion and exclusion criteria. In general, adults with obesity between 25 and 65 years old (both included) who are physically inactive will be included. Specific inclusion and exclusion criteria are:

Inclusion criteria:

1. Aged 25-65 years.
2. Body mass index (BMI) between 30 and <40 kg/m$^2$.
3. Weight stability (within 3% of screening weight) for >2 months prior to study entry.
4. Habitual eating window ≥ 11 hours.

Exclusion criteria:

1. History of a major adverse cardiovascular event (e.g., acute myocardial infarction, ischemic or hemorrhagic stroke, or peripheral arterial ischemia, among others), kidney failure, chronic liver disease, or human immunodeficiency virus (HIV) / acquired immunodeficiency syndrome (AIDS).
2. Active endocrinological disease, innate errors of metabolism, myopathies or epilepsy.
3. Patients who have undergone bariatric surgery or other surgical techniques used for the treatment of other pathologies (e.g., "Roux-en-Y").
4. Rheumatoid arthritis, Parkinson's disease, active cancer treatment in the past year or another medical condition in which fasting is contraindicated.
5. Use of medications that may affect the results of the study such as drugs for glycemic control (e.g., antidiabetic, steroids, beta-blockers, antibiotics, prebiotics, probiotics and symbiotics).
6. Diagnosis of major sleep or eating disorders.
7. Caregiver for a dependent requiring frequent nocturnal care/sleep interruption or shift workers with nocturnal hours.
8. Metal or electrical prosthesis.
9. Foreign bodies in the eyes.
10. Fear of needles and claustrophobia to MRI.
11. Active tobacco or illicit drug use or a history of alcohol abuse treatment (i.e., moderate or severe alcoholism).
12. Participating in a weight loss or a supervised exercise program (i.e., > 30 min in 3 times per week, or > 45 min in 2 or more times per week at moderate-to-vigorous intensity).
13. Pregnancy and lactation or planned pregnancy (within the study period).
14. Frequent travel over time zones during the study period.
15. Being unable to understand and accept the instructions or the study objectives and protocol.
16. Not having or being able to use a smartphone with Apple iOS or Android OS.
17. Are deemed unsuitable by the investigator for any other reason.

*4.3. Recruitment*

Information necessary for the CONSORT flow diagram will be collected[4]. For the enrolment

phase, we will note the number of patients that are assessed for eligibility by the research team, the number of excluded patients (plus reason for exclusion), and the number of randomized participants. For the allocation, the number of participants allocated to the intervention, and the number of participants who received or did not receive the intervention (plus reasons) will be noted. For follow-up, the number of participants lost to follow-up and the number who discontinued the intervention (plus reasons) will be counted. Finally, the number of participants included in the analyses using the intention-to-treat principle will be described.

### *4.4. Withdrawal/follow-up*

The number, timing, and reasons (e.g., adverse events, lost motivation to continue with the study, withdrew consent, could not or did not want to attend the post-intervention evaluations) for withdrawal will be noted and described.

### *4.5. Baseline patient characteristics*

A baseline table will be created to describe the characteristics of the study population. The characteristics include general sociodemographic outcomes (e.g., age, sex), anthropometry and body composition (e.g., body weight, height, BMI, waist circumference, fat-free mass, fat mass), cardiometabolic health (e.g., blood pressure, glucose homeostasis, blood lipid profile, liver health markers), dietary intake (e.g., energy intake, percentage of macronutrients) and physical fitness (e.g., cardiorespiratory fitness, muscular strength). The characteristics of the total study population and each study arm will be summarized using mean (SD) or median (interquartile range) for normally and not normally distributed continuous variables, respectively, and as number (percentage) for categorical variables.

## 5. Analysis

### *5.1. Outcome definitions*

Primary outcome

1. Change in hepatic fat content (baseline and 12 weeks). The quantification of hepatic fat content will be performed using MRI with a Siemens 3T Magnetom Vida scanner.

Secondary outcomes

The secondary outcomes are:

SPECIFIC FAT DEPOTS:

2. Change in hepatic fat content (baseline and 12-month follow-up). The quantification of hepatic fat content will be performed using MRI with a Siemens 3T Magnetom Vida scanner.
3. Change in hepatic elasticity (baseline, 12 weeks, and 12-month follow-up). The quantification of viscosity and fibrosis severity will be assessed using attenuation imaging, shear wave elastography and shear wave dispersion with a Canon Aplio i800.
4. Change in abdominal subcutaneous adipose tissue (baseline, 12 weeks, and 12-month follow-up). The quantification of abdominal subcutaneous adipose tissue will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.

5. Change in abdominal visceral adipose tissue (baseline, 12 weeks, and 12-month follow-up). The quantification of visceral adipose tissue will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.
6. Change in abdominal intramuscular fat content (baseline, 12 weeks, and 12-month follow-up). The quantification of intramuscular fat content will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.
7. Change in pancreatic fat content (baseline, 12 weeks, and 12-month follow-up). The quantification of pancreatic fat content will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.

ANTHROPOMETRY AND BODY COMPOSITION:

8. Change in body weight (baseline, 12 weeks, and 12-month follow-up). Body weight will be measured by a digital scale.
9. Change in neck circumference (baseline, 12 weeks, and 12-month follow-up). Neck circumference will be assessed by measuring tape following the procedures outlined by the International Society for the Advancement of Kinanthropometry.
10. Change in waist circumference (baseline, 12 weeks, and 12-month follow-up). Waist circumference will be assessed by measuring tape following the procedures outlined by the International Society for the Advancement of Kinanthropometry.
11. Change in hip circumference (baseline, 12 weeks, and 12-month follow-up). Hip circumference will be assessed by measuring tape following the procedures outlined by the International Society for the Advancement of Kinanthropometry.
12. Change in calf girth (baseline, 12 weeks, and 12-month follow-up). Calf girth will be assessed by measuring tape following the procedures outlined by the International Society for the Advancement of Kinanthropometry.
13. Change in fat mass (baseline, 12 weeks, and 12-month follow-up). Fat mass will be assessed by bioelectrical impedance analysis (BIA).
14. Change in fat-free mass (baseline, 12 weeks, and 12-month follow-up). Fat-free mass will be assessed by BIA.
15. Change in bone mineral density and content (baseline, 12 weeks, and 12-month follow-up). Bone mineral density and content will be assessed by dual-energy X-ray absorptiometry (DXA).

BLOOD PRESSURE:

16. Change in systolic blood pressure (baseline, 12 weeks, and 12-month follow-up). Systolic blood pressure will be assessed by a blood pressure monitor.
17. Change in diastolic blood pressure (baseline, 12 weeks, and 12-month follow-up). Diastolic blood pressure will be assessed by a blood pressure monitor.

LIVER FUNCTION ENZYMES:

18. Change in alkaline phosphatase (baseline and 12 weeks). Fasting blood samples will be used to assess alkaline phosphatase.
19. Change in alanine transaminase (baseline and 12 weeks). Fasting blood samples will be used to assess alanine transaminase.
20. Change in gamma-glutamyl transferase (baseline and 12 weeks). Fasting blood samples will be used to assess gamma-glutamyl transferase.

GLUCOSE HOMEOSTASIS:

21. Change in fasting glucose (baseline and 12 weeks). Fasting blood samples will be used to assess fasting glucose.
22. Change in HbA1c (baseline and 12 weeks). Fasting blood samples will be used to assess HbA1c.
23. Change in HOMA-IR (baseline and 12 weeks). Fasting blood samples will be used to assess glucose and insulin, and HOMA-IR will be computed.
24. Change in mean glucose (baseline and the last 2 weeks of intervention). Mean glucose over 14 days will be assessed by continuous glucose monitoring during 2 weeks.
25. Change in 2-hour plasma glucose (baseline and 12 weeks). 2-hour plasma glucose will be assessed by an oral glucose tolerance test.

BLOOD LIPID PROFILE:

26. Change in fasting triglycerides (baseline and 12 weeks). Fasting blood samples will be used to assess levels of triglycerides.
27. Change in fasting high-density lipoprotein cholesterol (baseline and 12 weeks). Fasting blood samples will be used to assess levels of high-density lipoprotein cholesterol.
28. Change in fasting low-density lipoprotein cholesterol (baseline and 12 weeks). Fasting blood samples will be used to assess levels of low-density lipoprotein cholesterol.
29. Change in fasting total cholesterol (baseline and 12 weeks). Fasting blood samples will be used to assess levels of total cholesterol.
30. Change in apolipoprotein A1 (baseline and 12 weeks). Fasting blood samples will be used to assess levels of apolipoprotein A1.
31. Change in apolipoprotein B (baseline and 12 weeks). Fasting blood samples will be used to assess levels of apolipoprotein B.

CIRCULATING INFLAMMATORY BIOMARKERS:

32. Change in C-reactive protein (baseline and 12 weeks). Fasting blood samples will be used to assess levels of C-reactive protein.
33. Change in interleukin 6 (baseline and 12 weeks). Fasting blood samples will be used to assess levels of interleukin 6.

DIETARY HABITS AND EATING BEHAVIOUR:

34. Change in energy intake (baseline and 12 weeks). Energy intake (kcal/day) will be assessed by 24h recalls.
35. Change in carbohydrate intake (baseline and 12 weeks). Carbohydrate intake (g/day and percentage of energy intake) will be assessed by 24h recalls.
36. Change in fat intake (baseline and 12 weeks). Fat intake (g/day and percentage of energy intake) will be assessed by 24h recalls.
37. Change in protein intake (baseline and 12 weeks). Protein intake (g/day and percentage of energy intake) will be assessed by 24h recalls.
38. Change in dietary habits (baseline and 12 weeks). Dietary habits will be assessed by a food frequency questionnaire (FFQ). Minimum value is 1 (never) and maximum value is 9 (more than 6 times per day). Higher values mean a higher frequency of a certain food consumption.
39. Change in appetite traits (baseline and 12 weeks). Appetite traits will be assessed by the Adult Eating Behavior Questionnaire (AEBQ). Minimum value is 1 (completely disagree) and maximum value is 5 (completely agree). Higher values mean a worse outcome.

SLEEP AND PHYSICAL ACTIVITY PATTERNS:

40. Change in subjective sleep quality (baseline and 12 weeks). Subjective sleep quality will be assessed by the Pittsburgh Sleep Quality Index (PSQI). Minimum value is 0 (never) and maximum value is 3 (3 or more times per week). Higher values mean a worse outcome.
41. Change in objective sleep quality (baseline and the last 2 weeks of intervention). Objective sleep quality will be assessed by accelerometry.
42. Change in chronotype (baseline and 12 weeks). Chronotype will be assessed by the Munich Chronotype Questionnaire (MCTQ).
43. Change in Morning-Evening type (baseline and 12 weeks). Morning-Evening type will be assessed by the Morningness-Eveningness Questionnaire Self-Assessment Version. It defines if a person is more morningness or eveningness based on daily time preferences.
44. Change in moderate to vigorous physical activity levels (baseline and the last 2 weeks of intervention). Physical activity levels will be assessed by accelerometry.

PSYCHOLOGICAL OUTCOMES AND QUALITY OF LIFE:

45. Change in depression traits (baseline and 12 weeks). Depression aspects will be assessed by the Beck Depression Inventory Fast Screen (BDI-FS). Values ranged from 0 to 63. Higher values mean a worse outcome.
46. Change in stress traits (baseline and 12 weeks). Stress aspects will be assessed by the Perceived Stress Scale (PSS). Values ranged from 0 to 40. Higher values mean a worse outcome.
47. Change in anxiety traits (baseline and 12 weeks). Anxiety aspects will be assessed by the State-Trait Anxiety Inventory (STAI). Values ranged from 0 to 60. Higher values mean a worse

outcome.

48. Change in general health (baseline and 12 weeks). General health will be assessed by the EuroQol 5 dimensions 5 levels (EQ-5D-5L). Values ranged from 0 to 100. Higher values mean a better outcome.
49. Change in quality of life (baseline and 12 weeks). Quality of life will be assessed by the Rand Short Form 36 (SF-36). Values ranged from 0 to 100. Higher values mean a better outcome.

FECAL MICROBIOTA:

50. Change in fecal microbiota composition (baseline and 12 weeks). DNA sequencing to determine fecal microbiota composition (e.g., phylum and genera).
51. Change in fecal microbiota diversity (baseline and 12 weeks). DNA sequencing to determine fecal microbiota diversity (e.g., beta and alpha).

PHYSICAL FITNESS:

52. Change in cardiorespiratory fitness (baseline and 12 weeks). Cardiorespiratory fitness will be measured by a maximal treadmill test.
53. Change in lower body muscular strength (baseline and 12 weeks). Lower body muscular strength will be measured by the chair stand test.
54. Change in upper body muscular strength (baseline and 12 weeks). Upper body muscular strength will be measured by the handgrip strength test.
55. Change in walking speed (baseline and 12 weeks). Walking speed will be measured by the gait speed test. Higher values mean worse performance.

ADHERENCE and ATTENDANCE:

56. Adherence to the eating window (during the 12 weeks). Adherence will be assessed by eating records through the mobile phone app.
57. Attendance to the exercise intervention (during the 12 weeks). Attendance will be assessed by number of completed exercise sessions.

OTHERS:

58. Change in levels of vitamin D (baseline and 12 weeks). Fasting blood samples will be used to assess vitamin D.
59. Change in levels of calcium (baseline and 12 weeks). Fasting blood samples will be used to assess calcium.
60. Change in levels of bilirubin (baseline and 12 weeks -up). Fasting blood samples will be used to assess bilirubin.
61. Change in levels of glomerular filtration rate (baseline and 12 weeks). Fasting blood samples will be used to assess glomerular filtration rate.

62. Change in levels of estradiol (baseline and 12 weeks). Fasting blood samples will be used to assess estradiol.
63. Change in levels of progesterone (baseline and 12 weeks). Fasting blood samples will be used to assess progesterone.
64. Change in levels of testosterone (baseline and 12 weeks). Fasting blood samples will be used to assess testosterone.
65. Change in levels of follicle-stimulating hormone (baseline and 12 weeks). Fasting blood samples will be used to assess follicle-stimulating hormone.
66. Change in levels of luteinizing hormone (baseline and 12 weeks). Fasting blood samples will be used to assess luteinizing hormone.
67. Change in levels of thyrotropin (baseline and 12 weeks). Fasting blood samples will be used to assess thyrotropin.
68. Change in levels of thyroxine (baseline and 12 weeks). Fasting blood samples will be used to assess thyroxine.
69. Change in levels of triiodothyronine (baseline and 12 weeks). Fasting blood samples will be used to assess triiodothyronine.
70. Change in levels of creatinine (baseline and 12 weeks). Fasting blood samples will be used to assess levels of creatinine.
71. Change in levels of creatine kinase (baseline and 12 weeks). Fasting blood samples will be used to assess levels of creatine kinase.
72. Change in levels of iron (baseline and 12 weeks). Fasting blood samples will be used to assess levels of iron.
73. Change in levels of ferritin (baseline and 12 weeks). Fasting blood samples will be used to assess levels of ferritin.
74. Change in levels of folic acid (baseline and 12 weeks). Fasting blood samples will be used to assess levels of folic acid.

### *5.2. Analysis methods*

The effects on primary and secondary outcomes in response to the present 12-week intervention will be assessed based on repeated-measures linear mixed-effects multilevel models. Individual measures of change will be therefore modeled as a function of the randomly assigned group, assessment time, and their interaction terms. All the analyses will also be conducted separately for men and women. Model-based estimations will be performed with an intention-to-treat approach (primary analyses) using the restricted maximum-likelihood method, the model assuming that missing values are missing-at-random. Analyses and estimations will also be performed with a per-protocol approach, and an attrition propensity will be calculated using a logistic model predicting attrition with baseline values of allocation group, age, sex, and BMI. Additional models will be conducted including energy intake, physical activity, or reproductive status in women. In addition to the conventional approach of assessing intervention effects

based on statistical and practical significance, it is important to highlight that this study will employ a practical benefit approach. This approach emphasizes the reporting of unadjusted values that are intuitive to human judgment and easily replicable, considering the design and methodology of the study.

The main statistical analyses will be performed by an independent researcher (Dr. Almudena Carneiro-Barrera), who is not involved in the recruitment, evaluations, and interventions, and will be performed blinded to the treatment allocation by coding the intervention arms (e.g., A, B, C, D).

### *5.3. Missing data*

The number of missing data will be reported, and patterns of missing data will be explored. Based on previous experience, we expect that missing data will be assumed as missing at random. Therefore, the linear mixed model analyses will handle our missing data. However, once the data processing is finalized, we will reconsider this assumption. In case we believe this assumption does not hold, we will take appropriate measures for the data analyses by using, for example, other multiple imputation techniques.

### *5.4. Additional analyses*

Changes in other outcomes will be analysed using a similar protocol as described by '*5.2 Analysis methods*' unless other analyses would be more appropriate depending on the outcome (e.g., ordinal / dichotomous outcomes). Furthermore, cross-sectional analyses will be performed using the baseline data of the RCT. For example, cross-sectional associations will be explored by bivariate and partial correlations. Then, further analyses will be performed using linear and logistic regression depending on the nature of the study variables and research questions. ANOVAs will also be used to test differences in outcomes among groups.

### *5.5. Harms*

The number and reasons of adverse events (e.g., headache, dizziness, falls, injuries, musculoskeletal problems, cardiovascular disease events, and any other events potentially related to the implementation of the trial protocol) at each time point will be collected, reported, and described separately for each study arm. No formal statistical testing will be undertaken.

### *5.6. Statistical software*

The analyses will be performed using R. For the main analyses we will use the 'lme4' or 'LMMstar' package. The use of packages will be reported in the manuscript.

## 6. References

1. Camacho-Cardenosa A, Clavero-Jimeno A, Martin-Olmedo JJ, et al. Time-restricted eating and supervised exercise for improving hepatic steatosis and cardiometabolic health in adults with obesity: protocol for the TEMPUS randomised controlled trial. *BMJ Open*. 2024;14(1):e078472. doi:10.1136/bmjopen-2023-078472

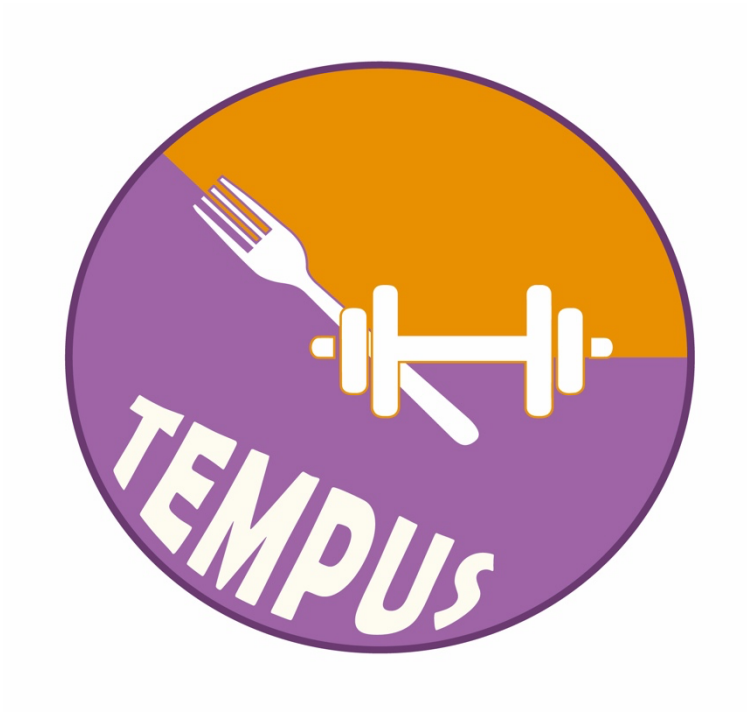
2. Ezpeleta M, Gabel K, Cienfuegos S, et al. Effect of alternate day fasting combined with

aerobic exercise on non-alcoholic fatty liver disease: A randomized controlled trial. *Cell Metab*. 2023;35(1):56-70.e3. doi:10.1016/j.cmet.2022.12.001

3. Li G, Taljaard M, Van den Heuvel ER, et al. An introduction to multiplicity issues in clinical trials: the what, why, when and how. *Int J Epidemiol*. 2017;46(2):746-755. doi:10.1093/ije/dyw320

4. Junqueira DR, Zorzela L, Golder S, et al. CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials. *BMJ*. 2023;381:e073725. doi:10.1136/bmj-2022-073725

# TEMPUS




# Statistical Analysis Plan

University of Granada, Spain
Trial registration number: NCT05897073
Version: 2
Date: 01/07/2024

**Statistical analysis plan (SAP) revision history**: revision made on 1st, July, 2024

**Roles and responsibility**:

Name and address of principal investigator and trial coordinators: Prof. Dr. Jonatan R. Ruiz (principal investigator)
Dr. Alba Camacho-Cardenosa (trial coordinator)
Mr. Antonio Clavero-Jimeno (trial coordinator)
PROFITH (PROmoting FITness and Health Through Physical Activity) research group, Department of Physical Education and Sports, Sport and Health University Research Institute (iMUDS), University of Granada
Faculty of Sport Sciences, Carretera de Alfacar s/n, Granada, 18071, Spain

**Name of person writing SAP:** E-mail principal investigator: ruizj@ugr.es
E-mails trial coordinators: acamachocardenos@ugr.es & claveroa@ugr.es
PROFITH research group, Department of Physical Education and Sports, Faculty of Sport Sciences, Sport and Health University Research Institute (iMUDS), University of Granada, Granada, Spain

**Signatures**:

| | |
|---|---|
| Prof. Dr. Jonatan R. Ruiz (principal investigator) | Date: 01/07/2024 |
| Dr. Alba Camacho-Cardenosa (trial coordinator) | Date: 01/07/2024 |
| Mr. Antonio Clavero-Jimeno (trial coordinator) | Date: 01/07/2024 |

# Table of contents

*1. General information of TEMPUS study ..... 17*

**1.1. Background and rationale ..... 17**

**1.2. Objective ..... 17**

*2. Study methods ..... 18*

**2.1. Trial design ..... 18**

**2.2. Randomization ..... 19**

**2.3. Sample size ..... 19**

**2.4. Framework ..... 19**

**2.5. Statistical interim analyses and stopping guidance ..... 19**

**2.6. Timing of the final analyses ..... 19**

**2.7. Timing of outcome assessment ..... 19**

*3. Statistical principles ..... 19*

**3.1. Confidence and P values ..... 20**

**3.2. Adherence, attendance, and protocol deviations ..... 20**

**3.3. Analysis populations ..... 20**

*4. Trial population ..... 20*

**4.1. Screening data ..... 20**

**4.2. Eligibility ..... 21**

**4.3. Recruitment ..... 22**

**4.4. Withdrawal/follow-up ..... 22**

**4.5. Baseline patient characteristics ..... 22**

*5. Analysis ..... 22*

**5.1. Outcome definitions ..... 22**

**5.2. Analysis methods ..... 27**

**5.3. Missing data ..... 28**

**5.4. Additional analyses ..... 28**

**5.5. Harms ..... 28**

**5.6. Statistical software ..... 28**

*6. References ..... 28*

*7. Summary of changes to Statistical Analysis Plan ..... 30*

# 1. General information of TEMPUS study

## *1.1. Background and rationale*

Metabolic dysfunction-associated steatotic liver disease (MASLD) is a major public health concern due to its high prevalence and strong association with extrahepatic conditions, particularly among adults with obesity, in whom it often remains undiagnosed. Emerging approaches such us time-restricted eating (TRE)—a novel form of intermittent fasting—and supervised exercise may represent promising strategies to mitigate associated health risks. Preliminary evidence suggests that combining TRE and exercise may be an effective strategy for improving body composition and cardiometabolic health in adults with obesity. However, studies examining the combined effects of TRE and exercise in humans are scarce and have notable limitations.

## *1.2. Objective*

The overall objective is to investigate the effects of TRE combined with supervised exercise on hepatic fat content (primary outcome) and cardiometabolic health (secondary outcomes) in adults with obesity.

The primary objective is to investigate the effects of a 12-week TRE intervention combined with a supervised exercise intervention (TRE + Exercise), as compared with TRE (TRE) or supervised exercise alone (Exercise), and usual care control group (UC), on hepatic fat content (primary outcome) and cardiometabolic health (secondary outcomes) in adults with obesity.

Secondary aims are to:

i. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on other specific fat depots: abdominal visceral and subcutaneous adipose tissue, pancreatic fat and mid-thigh intermuscular and intramuscular adipose tissue, and body composition (fat mass and fat-free mass).
ii. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on cardiometabolic health markers including blood pressure, liver enzymes, glucose homeostasis, blood lipid profile, and circulating inflammatory biomarkers.
iii. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on fecal microbiota.
iv. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on sleep and physical activity patterns.
v. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on psychological outcomes (i.e., depression, stress and anxiety) and quality of life.
vi. Investigate the effects of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on physical fitness (i.e., cardiorespiratory fitness and muscular strength).
vii. Examine the effect of a 12-week TRE + Exercise intervention, as compared with TRE or supervised exercise alone, and UC, on dietary habits and eating behaviour.
viii. To quantify the persistence of the effects of TRE + Exercise intervention, TRE alone,

and supervised exercise alone on hepatic fat content and cardiometabolic health in the long-term (12-month follow-up).

## 2. Study methods

### *2.1. Trial design*

The TEMPUS project is an investigator initiated, parallel group, randomized controlled trial (RCT). The RCT will include four-arms containing a UC control group and three intervention groups that will receive three different 12-week TRE and/or supervised exercise programs (**Figure 1**): i) TRE alone, ii) Exercise alone and iii) TRE + Exercise. The study will be conducted in ~184 patients (50% women) with obesity who meet the eligibility criteria. Measurements will be performed at baseline, at midpoint (6 weeks), post intervention (12 weeks) and 12-month follow-up (12 months post intervention). More detailed information is described in the study protocol.[1]

The TEMPUS research team will perform the data quality control, data processing, writing analyses scripts/programs and statistical analysis.

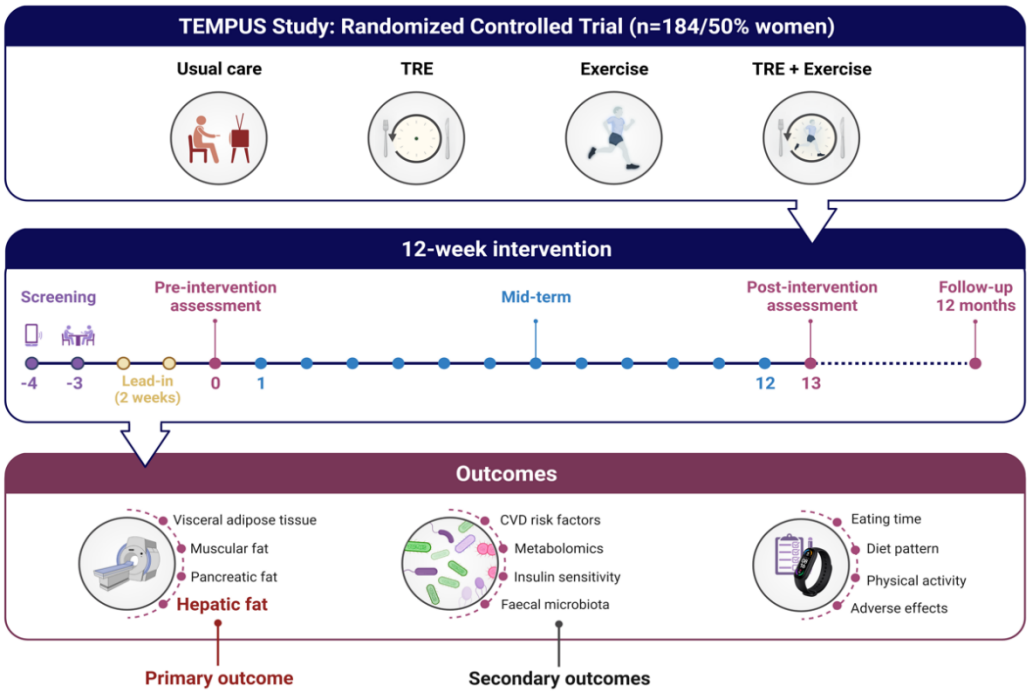



**Figure 1.** Overview of the TEMPUS Study. TRE, time-restricted eating.

### *2.2. Randomization*

Participants will be randomized using a parallel design with a 1:1:1:1 allocation ratio and stratified for sex.

### *2.3. Sample size*

Based on previous findings from a recent trial on the combination of alternate day fasting and exercise on hepatic fat content,[2] we anticipate approximately 5.0% reduction in hepatic fat content in the TRE + Exercise group, 2.5% in the TRE group, 2.5% in the exercise group, and no significant change in this outcome in the usual-care group. Assuming a pre-post correlation of 0.8 and a standard deviation of 6 points in the main outcome, we estimate a medium effect size of 0.45. To detect this effect size as statistically significant in a one-way ANOVA with $\alpha = 0.05$ and a power of 0.8, a minimum of 19 patients per group is required. Accounting for subgroup analyses based on sex and a maximum dropout rate of 20%, we will aim to recruit ~46 participants for each trial group, resulting in a total sample size of ~184 participants, with ~92 women included.

Our sample size calculations have taken into account subgroup analyses by sex. We have conservatively estimated a maximum dropout rate of approximately 20%. By considering this dropout rate, we have ensured that our study is adequately powered to detect the specified effect size even if there are differential dropout rates between men and women. For example, if men have a dropout rate of 5% and women have a dropout rate of 15%, our study will still have sufficient power to analyse the data separately in men and women.

### *2.4. Framework*

A superiority hypothesis testing framework will be used. In the main analyses, we will compare whether TRE plus exercise interventions (TRE + Exercise group) is superior to TRE or Exercise alone or the UC (Mediterranean education program only).

### *2.5. Statistical interim analyses and stopping guidance*

No pre-specified interim analyses will be performed, and therefore, a stopping guidance will not be applicable.

### *2.6. Timing of the final analyses*

The final analyses will be performed after the finalization of the data collection and processing of the primary and secondary outcomes.

### *2.7. Timing of outcome assessment*

The primary (i.e., hepatic fat content) and secondary outcomes (i.e., other specific fat depots measured by magnetic resonance imaging [MRI], and cardiometabolic health markers) will be assessed at baseline, post intervention (12 weeks), and at a 12-month follow-up (12 months post intervention). The time frame of each outcome is explained in section '*5.1 Outcome definitions*'

## 3. Statistical principles

### *3.1. Confidence and P values*

All statistical tests will be two-tailed. *P* for significance will be set at 0.05 and 95% confidence intervals will be estimated. No adjustment for multiplicity will be made as multiplicity adjustments may be of lesser importance in the case of distinct treatment arms.[3] Moreover, only one primary outcome will be defined and other outcomes (i.e., secondary outcomes, and exploratory analyses) will not require adjustment for multiple testing. However, we will include a hierarchical analytic approach for the interventions tested and the primary outcome.

### *3.2. Adherence, attendance, and protocol deviations*

During the 12-week intervention period, participants will be required to record their eating times (i.e., exact times of the beginning of the first meal and of the end of the last meal) in a mobile phone app specifically designed for the study. These data will be revised from 2 to 3 times every week, asking the participant for missing records, and will provide insights into their adherence to the prescribed eating window. For participants in the TRE + Exercise group and TRE alone group, each day will be labeled as adherent if the participant meets their self-selected 8-hour (±30 minutes) eating window. Adherence will then be defined as the percentage of days on which the eating window was met. Finally, we will assess the long-term adherence to the intervention (at the 12-month follow-up). This will allow us to evaluate the sustainability of participants' adherence over an extended period.

Attendance will be defined as percentage of exercise sessions attended by the participant (TRE + Exercise and Exercise alone groups), recorded by the trainers, divided by the actual number of exercise sessions offered (12 weeks x 2 sessions / week = 24).

All protocol deviations made to the protocol (e.g., change in pre-defined inclusion/exclusion criteria, baseline and post assessments, data cleaning/processing) will be reported and described.

### *3.3. Analysis populations*

We plan to use an intention-to-treat approach for our primary analyses, which will include all. randomized participants. With this approach, participants are analysed in the groups to which they were originally assigned, regardless of whether they completed the intervention. As a result, some participants will have valid data at both time points, while others may have missing data at baseline, post-intervention, or the 12-month follow-up assessment.

## 4. Trial population

### *4.1. Screening data*

Screening data will be based on patients who are defined as eligible by the clinical and research team. Screening is performed in three phases:

1) Screening based on phone call by the research team.
2) Screening based on the eligibility criteria (see '*4.2 Eligibility'*) by the clinicians of the research team.
3) Screening during the 2-week lead-in period prior to pre-intervention assessments and group allocation.

More details on the eligibility criteria are provided in '*4.2 Eligibility*'. The number of patients within each phase will be reported.

*4.2. Eligibility*

Eligible patients were defined by the inclusion and exclusion criteria. In general, adults with obesity between 25 and 65 years old (both included) who are physically inactive will be included. Specific inclusion and exclusion criteria are:

Inclusion criteria:

1. Aged 25-65 years.
2. Body mass index (BMI) between 30 and <40 kg/m$^2$.
3. Weight stability (within 3% of screening weight) for >2 months prior to study entry.
4. Habitual eating window ≥ 11 hours.

Exclusion criteria:

1. History of a major adverse cardiovascular event (e.g., acute myocardial infarction, ischemic or hemorrhagic stroke, or peripheral arterial ischemia, among others), kidney failure, chronic liver disease, or human immunodeficiency virus (HIV) / acquired immunodeficiency syndrome (AIDS).
2. Active endocrinological disease, innate errors of metabolism, myopathies or epilepsy.
3. Patients who have undergone bariatric surgery or other surgical techniques or used for the treatment of other pathologies (e.g., "Roux-en-Y").
4. Rheumatoid arthritis, Parkinson's disease, active cancer treatment in the past year or another medical condition in which fasting is contraindicated.
5. Use of medications that may affect the results of the study such as drugs for glycemic control (e.g., antidiabetic, steroids, beta-blockers, antibiotics, prebiotics, probiotics and symbiotics).
6. Diagnosis of major sleep or eating disorders.
7. Caregiver for a dependent requiring frequent nocturnal care/sleep interruption or shift workers with nocturnal hours.
8. Metal or electrical prosthesis.
9. Foreign bodies in the eyes.
10. Fear of needles and claustrophobia to MRI.
11. Active tobacco or illicit drug use or a history of alcohol abuse treatment (i.e., moderate or severe alcoholism).
12. Participating in a weight loss or a supervised exercise program (i.e., > 30 min in 3 times per week, or > 45 min in 2 or more times per week at moderate-to-vigorous intensity).
13. Pregnancy and lactation or planned pregnancy (within the study period).
14. Frequent travel over time zones during the study period.

15. Being unable to understand and accept the instructions or the study objectives and protocol.
16. Not having or being able to use a smartphone with Apple iOS or Android OS.
17. Are deemed unsuitable by the investigator for any other reason.

### *4.3. Recruitment*

Information necessary for the CONSORT flow diagram will be collected[4]. For the enrolment phase, we will note the number of patients that are assessed for eligibility by the research team, the number of excluded patients (plus reason for exclusion), and the number of randomized participants. For the allocation, the number of participants allocated to the intervention, and the number of participants who received or did not receive the intervention (plus reasons) will be noted. For follow-up, the number of participants lost to follow-up and the number who discontinued the intervention (plus reasons) will be counted. Finally, the number of participants included in the analyses using the intention-to-treat will be described.

### *4.4. Withdrawal/follow-up*

The number, timing, and reasons (e.g., adverse events, lost motivation to continue with the study, withdrew consent, could not or did not want to attend the post-intervention evaluations) for withdrawal will be noted and described.

### *4.5. Baseline patient characteristics*

A baseline table will be created to describe the characteristics of the study population. The characteristics include general sociodemographic outcomes (e.g., age or sex), anthropometry and body composition (e.g., body weight, height, BMI, waist circumference, fat-free mass, fat mass), cardiometabolic health (e.g., blood pressure, glucose homeostasis, blood lipid profile, liver health markers), dietary intake (e.g., energy intake, percentage of macronutrients) and physical fitness (e.g., cardiorespiratory fitness, muscular strength). The characteristics of the total study population and each study arm will be summarized using mean (SD) or median (interquartile range) for normally and not normally distributed continuous variables, respectively, and as number (percentage) for categorical variables.

## 5. Analysis

### *5.1. Outcome definitions*

<u>Primary outcome</u>

1. Change in hepatic fat content (baseline and 12 weeks). The quantification of hepatic fat content will be performed using MRI with a Siemens 3T Magnetom Vida scanner.

<u>Secondary outcomes</u>

The secondary outcomes are:

<u>SPECIFIC FAT DEPOTS</u>:

2. Change in hepatic fat content (baseline and 12-month follow-up). The quantification of hepatic fat content will be performed using MRI with a Siemens 3T Magnetom Vida

scanner.

3. Change in abdominal visceral adipose tissue (baseline, 12 weeks, and 12-month follow-up). The quantification of visceral adipose tissue will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.
4. Change in abdominal subcutaneous adipose tissue (baseline, 12 weeks, and 12-month follow-up). The quantification of abdominal subcutaneous adipose tissue will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.
5. Change in abdominal intermuscular fat content (baseline, 12 weeks, and 12-month follow-up). The quantification of intermuscular fat content will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.
6. Change in pancreatic fat content (baseline, 12 weeks, and 12-month follow-up). The quantification of pancreatic fat content will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.
7. Change in abdominal skeletal muscle tissue (baseline, 12 weeks, and 12-month follow-up). The quantification of skeletal muscle tissue will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.
8. Change in mid-thigh subcutaneous adipose tissue (baseline, 12 weeks, and 12-month follow-up). The quantification of abdominal subcutaneous adipose tissue will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.
9. Change in mid-thigh intermuscular fat content (baseline, 12 weeks, and 12-month follow-up). The quantification of intermuscular fat content will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.
10. Change in mid-thigh intramuscular fat content (baseline, 12 weeks, and 12-month follow-up). The quantification of intramuscular fat content will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.
11. Change in mid-thigh skeletal muscle tissue (baseline, 12 weeks, and 12-month follow-up). The quantification of skeletal muscle tissue will be assessed using MRI with a Siemens 3T Magnetom Vida scanner.

ELASTOGRAPHY:

12. Change in liver stiffness (baseline, 12 weeks, and 12-month follow-up). The quantification of liver stiffness will be assessed using shear wave elastography with a Canon Aplio i800.
13. Change in liver steatosis (baseline, 12 weeks, and 12-month follow-up). The quantification of liver steatosis will be assessed using attenuation imaging coefficient with a Canon Aplio i800.
14. Change in liver viscosity (baseline, 12 weeks, and 12-month follow-up). The quantification of liver viscosity will be assessed using shear wave dispersion with a Canon Aplio i800.

ANTHROPOMETRY AND BODY COMPOSITION:

15. Change in body weight (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Body weight will be measured by a digital scale.
16. Change in waist circumference (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Waist circumference will be assessed by measuring tape following the procedures outlined by the

International Society for the Advancement of Kinanthropometry.

17. Change in hip circumference (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Hip circumference will be assessed by measuring tape following the procedures outlined by the International Society for the Advancement of Kinanthropometry.
18. Change in neck circumference (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Neck circumference will be assessed by measuring tape following the procedures outlined by the International Society for the Advancement of Kinanthropometry.
19. Change in calf girth (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Calf girth will be assessed by measuring tape following the procedures outlined by the International Society for the Advancement of Kinanthropometry.
20. Change in fat mass (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Fat mass will be assessed by bioelectrical impedance analysis (BIA).
21. Change in fat-free mass (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Fat-free mass will be assessed by BIA.
22. Change in bone mineral density and content (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Bone mineral density and content will be assessed by dual-energy X-ray Absorptiometry (DXA).

BLOOD PRESSURE:

23. Change in systolic blood pressure (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Systolic blood pressure will be assessed by blood pressure monitor.
24. Change in diastolic blood pressure (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Diastolic blood pressure will be assessed by blood pressure monitor.

LIVER ENZYMES:

25. Change in levels of alkaline phosphatase (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to asses alkaline phosphatase.
26. Change in levels of alanine transaminase (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to asses alanine transaminase.
27. Change in levels of aspartate aminotransferase (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to asses aspartate aminotransferase.
28. Change in levels of gamma-glutamyl transferase (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess gamma-glutamyl transferase.

GLUCOSE HOMEOSTASIS:

29. Change in levels of fasting glucose (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess fasting glucose.
30. Change in levels of HbA1c (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess HbA1c.
31. Change in levels of fasting insulin (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess insulin.

32. Change in levels of HOMA-IR (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess glucose and insulin and HOMA-IR will be computed.
33. Change in 2-hour plasma glucose (baseline and 12 weeks). 2-hour plasma glucose will be assessed by an oral glucose tolerance test.
34. Change in levels of mean glucose (baseline and the last 2 weeks of intervention). 24-hour, diurnal and nocturnal mean glucose over 14 days will be assessed by continuous glucose monitoring during 2 weeks.

BLOOD LIPID PROFILE:

35. Change in levels of fasting triglycerides (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess levels of triglycerides.
36. Change in levels of fasting high-density lipoprotein cholesterol (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess levels of high-density lipoprotein cholesterol.
37. Change in levels of fasting low-density lipoprotein cholesterol (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess levels of low-density lipoprotein cholesterol.
38. Change in levels of fasting total cholesterol (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess levels of total cholesterol.
39. Change in levels of apolipoprotein A1 (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess levels of apolipoprotein A1.
40. Change in levels of apolipoprotein B (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess levels of apolipoprotein B.

CIRCULATING INFLAMMATORY BIOMARKERS:

41. Change in levels of C-reactive protein (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess levels of C-reactive protein.
42. Change in levels of interleukin 6 (baseline, 12 weeks, and 12-month follow-up). Fasting blood samples will be used to assess levels of interleukin 6.

AMINO ACIDS:

43. Change in levels of essential amino acids (baseline and 12 weeks). Fasting urine samples will be used to assess levels of essential amino acids.
44. Change in levels of non-essential amino acids (baseline and 12 weeks). Fasting urine samples will be used to assess levels of non-essential amino acids.

DIETARY HABITS AND EATING BEHAVIOUR:

45. Change in energy intake (baseline, 6 weeks and 12 weeks). Energy intake (kcal/day) will be assessed by 24h recalls.
46. Change in carbohydrate intake (baseline, 6 weeks and 12 weeks). Carbohydrates intake (g/day and percentage of energy intake) will be assessed by 24h recalls.

47. Change in fat intake (baseline, 6 weeks and 12 weeks). Fat intake (g/day and percentage of energy intake) will be assessed by 24h recalls.
48. Change in protein intake (baseline, 6 weeks and 12 weeks). Protein intake (g/day and percentage of energy intake) will be assessed by 24h recalls.
49. Change in fiber intake (baseline, 6 weeks and 12 weeks). Fiber intake (g/day and percentage of energy intake) will be assessed by 24h recalls.
50. Change in dietary habits (baseline and 12 weeks). Dietary habits will be assessed by food frequency questionnaire (FFQ). Minimum value is 1 (never) and maximum value is 9 (more than 6 times per day). Higher values mean a higher frequency of a certain food consumption.
51. Change in appetite traits (baseline and 12 weeks). Appetite traits will be assessed by the Adult Eating Behavior Questionnaire (AEBQ). Minimum value is 1 (completely disagree) and maximum value is 5 (completely agree). Higher values mean a worse outcome.
52. Change in adherence to Mediterranean diet (baseline, 6 weeks, 12 weeks, and 12-month follow-up). Adherence to the Mediterranean dietary pattern will be assessed using validated questionnaires such as the PREDIMED questionnaire. The minimum value is 0 and the maximum value is 14. Higher values indicate better adherence to the Mediterranean diet and represent a better outcome.

SLEEP AND PHYSICAL ACTIVITY PATTERNS:

53. Change in Subjective sleep quality (baseline and 12 weeks). Subjective sleep quality will be assessed by the Pittsburgh Sleep Quality Index (PSQI). Minimum value is 0 (never) and maximum value is 3 (3 or more times per week). Higher values mean a worse outcome.
54. Change in objectively sleep quality (baseline and the last 2 weeks of intervention). Objectively sleep quality will be assessed by accelerometry.
55. Change in chronotype (baseline and 12 weeks). Chronotype will be assessed by the Munich Chronotype Questionnaire (MCTQ).
56. Change in morning-evening type (baseline and 12 weeks). Morning-Evening type will be assessed by the Morningness-Eveningness Questionnaire Self-Assessment Version. It defines if a person is more morningness or eveningness based on daily times preferences.
57. Change in moderate to vigorous physical activity levels (baseline, the last 2 weeks of intervention, and 12-month follow-up). Physical activity levels will be assessed by accelerometry.
58. Change in rest-activity rhythms (baseline and the last 2 weeks of intervention). Rest-activity rhythms will be assessed by accelerometry.
59. Change in step count (baseline and during the 12-week intervention). Step count will be assessed by activity band.

PSHYCHOLOGICAL OUTCOMES AND QUALITY OF LIFE:

60. Change in depression traits (baseline and 12 weeks). Depression aspects will be assessed by the Beck Depression Inventory Fast Screen (BDI-FS). Values ranged from 0 to 63. Higher values mean worse outcome.
61. Change in stress traits (baseline and 12 weeks). Stress aspects will be assessed by the Perceived

Stress Scale (PSS). Values ranged from 0 to 40. Higher values mean worse outcome.

62. Change in anxiety traits (baseline and 12 weeks). Anxiety aspects will be assessed by the State-Trait Anxiety Inventory (STAI). Values ranged from 0 to 60. Higher values mean worse outcome.
63. Change in general health (baseline and 12 weeks). General health will be assessed by the EuroQol 5 dimensions 5 levels (EQ-5D-5L). Values ranged from 0 to 100. Higher values mean better outcome.
64. Change in quality of life (baseline and 12 weeks). Quality of life will be assessed by the Rand Short Form 36 (SF-36). Values ranged from 0 to 100. Higher values mean better outcome.

FECAL MICROBIOTA:

65. Change in fecal microbiota composition (baseline and 12 weeks). Shotgun metagenomic sequencing of DNA extracted from stool samples to determine taxonomic profiling (e.g., phylum, genera, species).
66. Change in fecal microbiota diversity (baseline and 12 weeks,). Shotgun metagenomic sequencing of DNA extracted from stool samples to determine fecal microbiota diversity (e.g., beta and alpha diversity metrics).
67. Change in fecal microbiota functionality (baseline and 12 weeks,). Shotgun metagenomic sequencing of DNA extracted from stool samples to determine microbial functional capacity through the analysis of metabolic pathways.

PHYSICAL FITNESS:

68. Change in cardiorespiratory fitness (baseline and 12 weeks). Cardiorespiratory fitness will be measured by a maximal treadmill test.
69. Change in lower muscular strength (baseline and 12 weeks). Lower body muscular strength will be measured by the chair stand test.
70. Change in upper muscular strength (baseline and 12 weeks). Upper body muscular strength will be measured by the handgrip strength test.
71. Change in walking speed (baseline and 12 weeks). Walking speed will be measured by the gait speed test. Higher values mean worse performance.

ADHERENCE AND ATTENDANCE:

72. Adherence to the eating window (during the 12 weeks). Adherence will be assessed by eating records through the mobile phone app.
73. Attendance to the exercise intervention (during the 12 weeks). Attendance will be assessed by number of completed exercise sessions.

### *5.2. Analysis methods*

The main analyses will consist of the intention-to-treat analyses for the primary and secondary outcomes using a constrained baseline (meaning baseline-adjusted) linear mixed model, which accounts for baseline differences among the study groups. The model will include fixed effects

for time (three levels) and treatment (four levels) as well as the unique participant identifier as a random effect, and sex will be included as a covariate. Model assumptions will be checked. In case we detect model violations, we will take appropriate measures by e.g., conducting data transformations. If the global test of significance indicates between-group differences, pairwise comparisons will be explored. Although no adjustments for multiplicity will be performed, family-wise type 1 error rate on the primary outcome will be retained by using a hierarchical analytic approach. The prespecified hierarchical hypotheses will be tested using the prespecified sequence: TRE + Exercise versus UC, TRE versus UC, Exercise versus UC, TRE + Exercise versus TRE, TRE + Exercise versus Exercise, and TRE versus Exercise. The main statistical analyses will be performed by an independent researcher (Dr. Almudena Carneiro-Barrera), who is not involved in the recruitment, evaluations, and interventions, and will be performed blinded to the treatment allocation by coding the intervention arms (e.g., A, B, C, D).

### *5.3. Missing data*

The number of missing data will be reported, and patterns of missing data will be explored. Based on previous experience, we expect that missing data will be assumed as missing at random. Therefore, the linear mixed model analyses will handle our missing data. However, once the data processing is finalized, we will reconsider this assumption. In case we believe this assumption does not hold, we will take appropriate measures for the data analyses by using, for example, other multiple imputation techniques.

### *5.4. Additional analyses*

Changes in other outcomes will be analyzed using a similar protocol as described by '*5.2 Analysis methods*' unless other analyses would be more appropriate depending on the outcome (e.g., ordinal / dichotomous outcomes). Furthermore, cross-sectional analyses will be performed using the baseline data of the RCT. For example, cross-sectional associations will be explored by bivariate and partial correlations. Then, further analyses will be performed using linear and logistic regression depending on the nature of the study variables and research questions. ANOVAs will also be used to test differences in outcomes among groups.

### *5.5. Harms*

The number and reasons of adverse events (e.g., headache, dizziness, falls, injuries, musculoskeletal problems, cardiovascular disease events, and any other events potentially related to the implementation of the trial protocol) at each time point will be collected, reported, and described separately for each study arm. No formal statistical testing will be undertaken.

### *5.6. Statistical software*

The analyses will be performed using R. For the main analyses we will use the 'lme4' or 'LMMstar' package. The use of packages will be reported in the manuscript.

## 6. References

1. Camacho-Cardenosa A, Clavero-Jimeno A, Martin-Olmedo JJ, et al. Time-restricted

eating and supervised exercise for improving hepatic steatosis and cardiometabolic health in adults with obesity: protocol for the TEMPUS randomised controlled trial. BMJ Open. 2024;14(1):e078472. doi:10.1136/bmjopen-2023-078472

2. Ezpeleta M, Gabel K, Cienfuegos S, et al. Effect of alternate day fasting combined with aerobic exercise on non-alcoholic fatty liver disease: A randomized controlled trial. Cell Metab. 2023;35(1):56-70.e3. doi:10.1016/j.cmet.2022.12.001

3. Li G, Taljaard M, Van den Heuvel ER, et al. An introduction to multiplicity issues in clinical trials: the what, why, when and how. Int J Epidemiol. 2017;46(2):746-755. doi:10.1093/ije/dyw320

4. Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomized trials. Nat Med. 2025;31(6):1776-1783. doi:10.1038/s41591-025-03635-5

## 7. Summary of changes to Statistical Analysis Plan

| Date | Amendment |
|---|---|
| 01/07/2024 | Review and improvement of the statistical analysis methods, including constrained baseline (meaning baseline-adjusted) linear mixed model. |
| 01/07/2024 | Update of secondary outcomes including new outcomes derived from magnetic resonance imaging, elastography, fasting urine, and fasting blood samples. |