TirolGESUND Study
Statistical Analysis Plan

Chiara Herzog, Ph.D.

November 2022

Contents

1 Introduction

This document follows outlines the proposed presentation and detailed analysis for the main papers reporting results from the single-centre TirolGESUND study to investigate the impact of a lifestyle intervention on a variety of health parameters, importantly epigenetics. This and the original study protocol, available in German, follow the SPIRIT guidelines[1] for study protocols.

The results reported in these papers will follow the strategy set out here. Any subsequent analyses of a more exploratory nature will not necessarily be bound by this strategy and will be detailed in separate analysis plans.

Suggestions for subsequent analyses by journal editors and referees will be considered carefully in line with the principles of this analysis plan. Any deviations from the statistical analysis plan will be described and justified. The analyses will be carried out by identified appropriately qualified and experienced statisticians or researchers who will ensure integrity of the data during their processing.

2 Background information

2.1 Rationale

A detailed outline of the background and rationale, including age and disease risk and epigenetic risk stratification, are detailed in the study protocol, available in German. Briefly, a majority of malignancies may be caused modifiable risk factors[2] and could therefore be prevented. Smoking[3] and diet[4,5] are known risk factors for cancer but also other disorders such as cardiovascular or metabolic disorders and neurodegeneration, and may promote premature cellular ageing. TirolGESUND investigates the effect of two lifestyle interventions, namely smoking cessation or intermittent fasting (both with additional adjuvant exercise), over 6 months for the promotion of health and reduction of disease risk.

As investigation of disease incidence and mortality following intervention require long-term follow-up, here we explore the use of primarily epigenetic signatures indicative of disease risk as so-called "surrogate endpoints" (surrogate to follow-up of long-term follow-up). Our primary aim is to identify the malleability of the epigenetic cancer risk- and age-associated signature scores in this cohort of healthy-albeit-at-increased-risk women (cervical, blood, and buccal samples). Women's cancer risk identification (WID) indices have been shown to have high sensitivity and specificity in identifying current cancer patients from healthy controls, less is known about their suitability for risk prediction and continuous monitoring during lifestyle or pharmacological changes. There is preliminary evidence that WID indices may reflect systemic epigenetic field defects and enable prediction. For example, the indices are increased in women with a *BRCA1* or *BRCA2* mutation who have an increased risk for breast cancer.[6] There are currently no prospective validation data defining thresholds for WID index positivity. Prospective validation is not the aim of TirolGESUND, instead we are assessing the impact of lifestyle intervention of such cancer risk signature scores as surrogate endpoints.

As smoking and diet constistute risk factors for a variety of other diseases, including cardiovascular, neurodegenerative, pulmonary, and intestinal, we are additionally evaluating the effect of lifestyle changes on these organ systems via additional surrogate markers or clinical measures.

The study is primarily a explorative study to evaluate longitudinal impacts of lifestyle change on health markers, and may not be sufficiently powered for each measure outlined below.

2.2 Study objectives

2.2.1 Primary objective

To assess whether a reduction and/or cessation in risk exposure (smoking, continuous non-time restricted intake of calories) via lifestyle modification over six months results in significant absolute changes in DNA methylation-based Women's cancer risk identification (WID) signature scores compared to baseline in cervical samples. WID signatures act as intermediate surrogate markers for disease risk (e.g. breast[6] or ovarian cancer[7]) and cellular ageing; high scores of these signatures indicate a higher risk for certain cancers, and a reduction thereof would indicate a reduced risk of later development of this disease. Successful completion relates to a significant reduction in two ore more WID risk scores with no significant in increase in any of the scores (see 5.4.1). Individuals that did not adhere to the intervention but continued to provide samples may be used as a nested control, if at least 10 individuals are present. Testing will be conducted using paired t tests or appropriate non-parametric tests. Linear or non-linear mixed-effects models may be employed.

2.2.2 Secondary objectives

The TirolGESUND study is novel type of participatory, lifestyle intervention study and may represent a considerable effort for participants. Few studies have so far so comprehensively evaluated health parameters longitudinally. Therefore, we will describe and report general study characteristics relating to participation and dropout, and evaluate changes, as well as dynamics, in several biological markers, including molecular (epigenetic, genetic, microbial, metabolic), clinical, and psychological (see 5.4.2).

We are also interested in investigating the spatial and temporal dynamics of DNA methylation changes in response to lifestyle intervention and will hence be investigating changes over time in different surrogate samples (cervical, buccal, blood) using mixed-effects models, and moreover investigate whether surrogate tissue DNA methylation changes occur in a synchronous or discordant manner.

Lastly, we will explore dynamics of other biomarkers and indicators of health and disease, including microbiome, metabolome, inflammation, body composition and smoking status, vascular health, pulmonary health, physical activity, and health-related quality of life.

2.2.3 Subsidiary outcomes or objectives

Subsidiary endpoints include analysis of longer-term changes at month 12 and 18 for participants who provided longer-term samples.

In a subgroup of participants (n=10), we will apply single cell multiomic analysis for RNA and ATAC-seq to investigate cellular heterogeneity at an early time point (two months

of intervention).

A subgroup of participants will be selected to investigate of which epigenetic sites are (individually) altered with cellular Yamanaka factor-induced reprogramming in samples collected at baseline (M0). In these participants we will investigated if the same or similar sites are significantly altered by the intervention at M6 (within-individual); if consistent regions are found across individuals, these sites may also applied to those individuals that were not included in reprogramming analysis.

Lastly, we may explore potential DNAme biomarkers that change in response to the intervention and are associated with clinically meaningful features (e.g. microbiome, lung function).

3 Study methods

3.1 Study design

TirolGESUND is a single-centre, open-label baseline-controlled study that encompasses two parallel arms with 6 months of intervention. One study is aimed at investigating the effect of smoking cessation. The other study is a randomised trial that evaluates the effect of intermittent fasting with or without ketogenic supplement. The two sub-studies will be evaluated separately. The study does not include a control group but instead investigates changes compared to baseline. Individuals that do not comply with the intervention but continue to provide samples may be utilised as a nested control group. The study is predominantly of exploratory nature, although comparisons to baseline measurements will be made as outlined in the primary, secondary, and subsidiary objectives. We aim to include a total of one-hundred and eighty participants as follows:

1. **Smoking cessation (n = 60)**
   Sixty women aged 30-60 without a significant prior medical history who have been smoking above or equal 10 cigarettes a day for at least five years will be recruited and instructed to stop smoking via a professional smoking cessation programme (provided by Suchthilfe Tirol) and 1:1 personal coaching.

2. **Dietary intervention (n = 120)**
   One-hundred and twenty women without significant prior medical history with a body mass index (BMI) between 25 and 35 will be recruited and aided to conduct 16:8 intermittent fasting (16 h fasting, 8 h window for food consumption each day). The window during which food can be consumed can be freely chosen by each participant Participants will be randomly allocated one of the following groups using age- and BMI-stratified block randomisation:

   - *Intermittent fasting (n = 60)*
     Intermittent fasting has been shown to convey a series of metabolic and health benefits, promoting longevity and reducing tumour incidence in animal models.[8] Recent human studies suggest similar benefits in humans.[9] We aim to assess whether intermittent fasting can improve a variety of health parameters.
   - *Intermittent fasting with ketogenic supplement (n = 60)*
     Intermittent fasting may mediate its benefits via timing of feeding/circadian

rhythms, caloric restriction,[10] and/or, in part, by promoting nutritional ketosis via the absence of carbohydrates;[11] promotion of ketosis has been shown to induce longevity in mice.[12] Ketogenic supplements are already commonly used in dietary therapy for pediatric epilepsy, but have not been routinely used in other settings. We aim to assess whether ketogenic supplements may boost ketosis and further enhance the health benefits of intermittent fasting.

3.2 Study setting

The study will be conducted at the EUTOPS (European Translational Oncology Prevention and Screening Institute) research clinic located at the Landeskrankenhaus Hall. Vascular examinations and optional skin biopsies are taken at the Krankenhaus Innsbruck. Sports examinations are conducted at the Institut für Sport-, Alpinmedizin und Gesundheitstourismus. Smoking cessation courses may be conducted in local seminar rooms.

3.3 Interventions

Participants will be allocated to the smoking cessation or dietary intervention sub-study based on inclusion criteria (described in detail in section 4.1). In the dietary intervention arm, participants will be further randomised between two subgroups. All participants will receive 1:1 personal coaching once a week in an attempt to minimise the attrition rate. Regardless of intervention group, participants will also receive an adjuvant targeted exercise program (detailed in section 3.3.4), aimed at improving overall health.

3.3.1 Smoking cessation

- Three smoking cessation group therapy sessions (6-12 participants per session) ideally within the first month of the study. Smoking cessation occurs in the second session. Thus, for each participant, a clearly defined (theoretical) smoking cessation date will be available
- Two individual telephone appointments with a professional addiction counsellor
- Adjuvant exercise programme and dietary guidance to promote a healthy diet at start of the study and 2 and 4 months after study initiation. These aspects were added to prevent weight gain after smoking cessation

3.3.2 Intermittent fasting

- Induction (1 month) and maintenance (5 months) of a 16:8 intermittent fasting regime (time restricted feeding). During the introductory month, women will only maintain a 14:10 intermittent fasting regime
- Detailed analysis of dietary behaviours and motivation and guidance for induction or maintenance of the intervention at the start of the study and months 2 and months 4 after study initiation
- Dietary guidance at the end of the study to either discontinue intermittent fasting or continue longer-term if desired by participant
- Adjuvant exercise programme

3.3.3 Intermittent fasting with ketogenic supplement

- Induction (1 month) and maintenance (5 months) of a 16:8 intermittent fasting regime (time restricted feeding)
- Detailed analysis of dietary behaviours and motivation and guidance for induction or maintenance of the intervention at the start of the study and months 2 and months 4 after study initiation
- Ketogenic supplement (Kanso MCTfiber, medium-chain triglyceride fiber), provided powderised in a sachet (daily dose) which can be added to food or diluted in water
- Dietary guidance at the end of the study to either discontinue intermittent fasting or continue longer-term
- Adjuvant exercise programme

3.3.4 Exercise programme

The exercise programme is not a key aim of analysis. It is provided to all participants for its general health-promoting effects and includes:

- Analysis of fitness levels at the start of the study (including exercise electrocardiogram)
- Provision of an exercise programme
- Three targeted exercise sessions (resistance, endurance/cardio, flexibility) for motivation and guidance

3.4 Quantification of compliance

Participants will be semi-continuously monitored for adherence to the intervention using the following measures:

- Smoking cessation:
  - Reported cigarettes smoked in the last 7 days via personal coaches
  - Urine cotinine levels at month 2, 4, and 6
- Intermittent fasting:
  - Reported compliance with the intermittent fasting regime over the last 7 days via coaches
  - Capillary β-hydroxybutyrate levels throughout the intervention as measured three times a week using FORA6 blood monitoring at home - available only for a subset of participants

Adherence to the compliance will also be reported by study doctors at each visit (compliance higher or lower than 75%).

A **compliance score** will be derived, taking into account the days compliant and the days between the visitations. If a subject reported compliant on all days between two visitations, the compliance score is 100%. Compliance will be cross-checked for concordance

using objective measurements (urine cotinine levels for smoking cessation, and capillary β-hydroxybutyrate levels for the diet intervention). High compliance with smoking cessation should result in no measurable cotinine levels in urine, although the metabolite is only short-lived. High compliance with diet intervention should result in elevated ketone body measurements in capillary blood. As an additional quantitative score of compliance, it may also be measured in number of cigarettes since last visit and cumulative ketone body measures in capillary blood.

3.5 Randomisation

Eligible participants in the dietary intervention group will be randomised to one of the two interventions using a secure Excel sheet developed by the statistician. The randomisation strategy is menopause- and BMI-stratified block randomisation in blocks of 4. Participants within each block will be allocated to a treatment 1:1. Users (clinicians) will provide menopausal status and BMI group and treatment allocation is then provided. The allocated treatment arm is recorded in the eCRF (Askimed).

No randomisation will be conducted in the smoking cessation sub-study.

3.6 Sample size

The sample size of n = 60 per group is based on an initial estimation of the potential clinically relevant effect and variance as part of the preceding FORECEE project: a total of 60 participants are required to detect a difference of 0.2 in DNA methylation indices (e.g. WID-BC index), assuming a standard deviation of 0.44, at a statistical power of 0.94 and a significance level of 0.05. There are no preliminary data for longitudinal changes in the WID scores, thereby TirolGESUND acts as a pilot study to define effect sizes and variance for larger prospective studies in the future.

3.7 Hypothesis framework

For each of the primary, secondary, or subsidiary outcomes, the null hypothesis is that there is no true difference in effect between the baseline measurement and the measurement taken at 6 months.

3.8 Blinding

This is an open-label study. While the study is in progress, access to results of key study outcomes will not be available to the research team, trial statisticians, or other research collaborators.

3.9 Timing of interim analyses

No interim analysis is planned for the TirolGESUND study.

3.10 Timing of analysis

The analysis of primary endpoints will be conducted collectively per sub-study once all samples and data become available for month 6 (main endpoint).

Secondary endpoints may be evaluated separately if data become available at a later timepoint, after initial investigation of the primary endpoints.

3.11 Timing of participant visits

Outcomes will be assessed using samples collected at the following timepoints:

- Baseline, including sports medicine examination and vascular analysis. All baseline measurements should be taken within 10 days of inclusion in the study, where feasible.
- Month 2 (7-10 weeks after inclusion)
- Month 4 (16-19 weeks after inclusion)
- Month 6 (25-28 weeks after inclusion)
- Month 12 (optional; 50-55 weeks after inclusion)
- Month 18 (optional; 75-79 weeks after inclusion)

Given practical limitations including annual leave, participant or staff illness, or other unforeseen circumstances, sample and data collection may be conducted outside these windows at the nearest possible date.

Baseline and regular outcome information will be collected in a trial-specific electronic case report form (eCRF), Askimed, by a member of the hospital or research staff at 2, 4, and 6 months of the study regardless of whether the participant is compliant with the intervention or not, although compliance will be monitored via regular check-ups by the coaches, biological measurements, and during study visitations. There are two optional follow-ups at 12 and 18 months. These will not be analysed with regards to primary endpoint but may inform stability of epigenetic changes for follow-up studies.

3.12 Study reporting

The study will be reported according to the principles of the CONSORT Statement.[13]

3.13 Recording of adverse or serious adverse events

Adverse events, or serious adverse events, are all (serious) adverse changes from the baseline of the participant which are deemed clinically relevant by the respective clinical doctor, including:

- Occurence of a new clinically significant disorder
- Worsening or increased frequency and/or intensity of pre-existing medical conditions
- clinically significant laboratory results outside of reference standards

These changes may not be necessarily directly related to the study intervention but will be recorded in the eCRF by the respective study doctor. Adverse or serious adverse events do not include:

- Planned interventions or hospitalisations
- Planned medical or surgical treatment, including surgeries, endoscopic examinations, tooth extractions, or blood transfusions

- Underlying medical conditions (except for those mentioned in the exclusion criteria) which do not worsen throughout the course of the study

No adverse or serious adverse events are expected in TirolGESUND as the interventions are based solely on health-promoting lifestyle changes.

4 Study population

4.1 Eligibility

4.1.1 Inclusion criteria

1. Women aged 30 to 60
2. Motivated to change their lifestyle
3. Smoking cessation: $\geq$ 10 cigarettes per day for at least the last five years
4. Dietary intervention: BMI between 25 and 35

* Should 3 and 4 apply, the participant will be allocated to the smoking cessation group.

4.1.2 Exclusion criteria

1. Relevant underlying conditions
   (a) Current or previous malignant tumour or cancer
   (b) Current or previous significant cardiovascular disorder
   Women with elevated blood pressure are allowed to participate as long as it is well controlled under their current medication
   (c) Current or previous metabolic disorder (e.g., diabetes type I or II)
   In the dietary intervention arm, participants with current hypothyroidism/Morbus Hashimoto will be excluded as the switch to intermittent fasting may require a adjustment of their medication
   (d) Current or previous psychiatric disorder (e.g., eating disorder, depression)
2. Current pregnancy or lactation period
3. Total hysterectomy
4. Known current or previous premalignant lesion of the cervix uteri (CIN2/3)
5. Concurrent participation in another interventional trial

Participants for which inclusion criteria 1b, c, d, 2, 3, and 5 occur after inclusion in the study will be excluded from the study.

4.2 Recruitment

Participants will be recruited via Tirol Kliniken, University of Innsbruck, other larger employers in Tirol (e.g. Bäcker Ruetz), via staff emails/intranet, online videos, posters, and employee magazines, or via personal recommendation from participating study investigators.

5 Statistical principles

5.1 Significance levels and adjustment of p-values for multiplicity

Significance levels throughout are set at $p < 0.05$. Formal adjustment for primary and secondary outcomes may be conducted using Benjamini-Hochberg or Bonferroni correction as appropriate, although formal adjustments may not be strictly necessary for outcomes of more exploratory or descriptive nature, as previously argued.[14] For subsidiary outcomes, no formal correction will be applied. 95% confidence intervals will be presented for estimates of effects throughout.

5.2 Analysis populations

The primary analysis will be an intention to treat (ITT) analysis in all participants, separately per study arm. A per protocol (PP) population will be all study subjects completing the study period with a median compliance above or equal to 80% (if sufficient n, e.g. n=30) or alternatively stratifying group by compliance above and beyond the compliance median. Additionally, the effect of compliance to the intervention will be evaluated using linear mixed-effects models with the compliance score as covariate. Timing may be coded as days since baseline rather than categorical variable (M2, M4, M6).

5.3 Control group

A control group of women with similar characteristics would have been desirable, yet for practical reasons would have proved challenging. The study constitutes a considerable effort for the participating group of healthy volunteers. Participants in a control group may not have seen a sufficient benefit for their participation and would have likely dropped out at high rates. Moreover, it would have not been possible to include a group of smokers and instruct them to not stop smoking for at least another 6 months. It is furthermore possible that participants in the control arm might have changed their health behaviours as they are increasingly confronted with health as part of the study.

Instead we have opted to utilise a longitudinal baseline-controlled design. Importantly, participants who are non-compliant are not excluded from the study. If they able to continue to provide samples, we include these samples as a type of "nested control" (i.e. what would happen without intervention?). Descriptive statistics of dynamics over 6 months without an intervention will be reported and if sufficient individuals are present per study arm ($n \geq 10$) that did not adhere, these may be used as a nested control group, although this number is small.

5.4 Definition of primary, secondary, and subsidiary endpoints

5.4.1 Primary objective

To assess whether a reduction/cessation in risk exposure (smoking, continuous non-time-restricted calorie intake resulting in being overweight/obese) via lifestyle modification over six months results in significant absolute changes in DNA methylation-based disease risk and ageing scores compared to baseline in cervical samples. Scores of the following age signatures will be evaluated:

- signatures that have previously been published, including
  - WID-BC
  - WID-OC
  - WID-EC
  - WID-CIN
  - WID-REA
  - WID-RIA
  - pcgtAge
- signatures that are not yet published but have been developed by the study group
  - Group of markers indicative of (stem) cellular age, e.g. WID-PCGT (general/immune/epithelial), WID-SEN (general/immune/epithelial), WID-PRO (general/immune/epithelial), etc. - thereafter referred to collectively as "WID-SOLA"
- signatures indicative of exposure, e.g.
  - WID-SMK smoking index

Successful completion relates to a significant reduction in two ore more risk scores with no significant in increase in any of the scores.

Analysis will be conducted both ITT, intention to treat, and PP, per protocol, comparing baseline and month 6. ITT analysis may be conducted as a mixed-effects models, using compliance score as a covariate. A per protocol analysis may be conducted either by median compliance rate above or equal to 80%, or alternative split by mean or median compliance rate (see Figure 1).

Subgroup analyses will also be conducted as follows:

1. smoking cessation study arm: investigation of the impact of the use of electronic cigarettes on smoking signature scores
2. intermittent fasting study arm: investigation of difference between individuals who received ketogenic supplement and those that did not at 6 months
3. intermittent fasting study arm: subgroup analysis by fasting pattern (breakfast or dinner skipping)

5.4.2 Secondary objectives

We will describe and report general study characteristics:

General study characteristics

- *dropout rate*
  Is the dropout rate consistent with the literature (approximately 15-20% for lifestyle

interventions)? The analysis will be reported separately for the two sub-studies and randomised trial arms.

- *compliance*
  What is the compliance across the two sub-studies?

- *sign-up rate*
  Out of participants that originally registered interest, how many wanted to finally participate and how many were included after meeting the criteria?

Moreover, the study will describe longitudinal alterations in a variety of health parameters (see also Fig. 2):

Molecular endpoints

- *epigenetic markers*
  DNA methylation-based disease risk, ageing, and exposure signatures in cervical, buccal, and blood samples.

- *microbiome*
  - changes in overall diversity and proportions of certain beneficial or harmful microbial species in fecal microbiome depending on the intervention
  - distinction of health and lean individuals using microbiome
  - distinction of smokers and non-smokers using microbiome
  - shift in dominant types and association with weight loss
  - potential changes in functional prediction

- *metabolites*
  alterations in metabolic profile of urine and saliva

- *blood profile*
  - lipids and HbA1c levels
  - cellular markers of inflammation
  - guided by cellular alterations, targeted investigation of levels of inflammatory cytokines and markers in blood plasma

- ageing markers in the skin and skin barrier function

Clinical and epidemiological endpoints

- body mass index (BMI), *diet arm of the study*
- body composition (muscle, fat, water), *diet arm of the study*
- smoking status, *smoking arm of the study*, including
  - current smoking (yes/no)

  - cigarettes in the last 7 days
  - use of nicotine replacement products

- *vascular health*
  pulse wave velocity, intima-media thickness, and plaque score before and after intervention
- abdominal fat composition, *diet arm of the study*
  visceral and subcutaneous fat before and after 6 months of intervention
- *physical activity and health*
  measured using the international physical activity questionnaire (IPAQ), sports examination (VO2max), and fitness tracker (daily steps, active minutes, floors climbed); resting heart rate
- *pulmonary health*

Psychological endpoints

- health-related quality of life (EuroQoL) before and after 6 months of intervention
- psychological self-efficacy before and after 6 months of intervention

5.4.3 Subsidiary outcomes or objectives

Optional samples are collected at month 12 and month 18 after inclusion into the study. Samples may be collected to evaluate long-term stability of changes in those participants that showed epigenetic and other biological changes in response to the intervention:

- ongoing "compliance" to smoking cessation or intermittent fasting: participants are not specifically instructed to continue with any intervention but it will be evaluated how many participants continue to be smoke-free or undergo intermittent fasting. It is likely that individuals that were highly compliant to begin with may be more likely to return for optional follow ups, which will need to be considered for this analysis.

- variability in epigenetic and other health markers evaluated in primary and secondary endpoints at month 6 and month 12 and/or month 18

Additionally, it is a subsidiary objective to identify changes in the following parameters in response to the intervention in subgroups of participants:

- blood (immune) cell identity and composition as assessed via multiomic single cell RNA- and ATAC-seq in a subgroup of patients (pilot study, baseline and 2 months): reduced numbers of inflammatory cell populations, increased chromatin accessibility at PCGT sites.
- epigenetic reprogramming signatures
- characterisation of epigenetic alterations following intervention in "compliers", and their relation to ageing and exposure signature (scores). Many studies have previously identified signatures that that capture epigenetic ageing yet the biological meaning of them remains unclear. It may be beneficial to identify CpG markers that are changed in response to a putative health-promoting intervention specifically and correlate with clinically meaningful changes (e.g. gut microbiome, lung function) in

a well-characterised longitudinal dataset. We may chose to undertake this subsidiary analysis pending the outcome of initial results.

6 Analysis

Analyses will be conducted and reported separately per sub-study (i.e. smoking cessation and intermittent fasting) after all samples for the month 6 timepoint have been collected. Key analyses will be performed on the ITT population. A per-protocol analysis will be conducted if sufficient individuals are available with above 80% compliance (n=30), or alternatively and additionally a subgroup analyses split by median compliance score (above and beyond, respectively). Descriptive statistics of dynamics over 6 months without an intervention will be reported and if sufficient individuals are present per study arm ($n \geq 10$) that did not adhere, these may be used as a nested control group; although this number is small, it could be informative for dynamic changes without intervention.

The longitudinal manner of data collection enables baseline-controlled pairwise comparison (paired t or Wilcoxon test) for each participant at baseline and after 6 months; linear mixed effects-models may also be employed. Longitudinal dynamics and repeated measurements per participant, evaluating data from baseline, 2, 4, and 6 months, will be assessed using linear (or non-linear) mixed-effects models to account for inter-individual heterogeneity and differences in response (random intercept and slopes, if appropriate). In the intermittent fasting group, cross-sectional analyses may be used to compare effects for use with and without ketogenic supplement. Change from baseline may be reported in a descriptive manner.

Methylation data will be adjusted for sample cell type composition which is known to influence methylation levels. Many of the assessed factors are influenced by age. Dependence on other, non-intervention related covariates will be investigated during the exploratory data visualisation and assessment, and factors may be associated as covariates in mixed-effects models if necessary.

If a number of participants does not adhere to the intervention but continues to provide samples and the assumption of common support holds, we may be able to explore the average treatment effect (ATE) and causal average treatment effect (CATE) using causal inference (g methods) via covariate adjustment for selected secondary outcomes (in particular longitudinal DNA methylation analyses).[15]

6.1 Statistical software

The R programming environment will be primarily used for analysis.

6.2 Assumption checking

Data will be assessed for normality via q-q plots, histogram (visually) and Shapiro-Wilk tests. For normal and non-normal data, appropriate tests will be applied (e.g. Wilcoxon or Student's t test).

6.3 General study characteristics

Summary statistics on general study data, including dropout rate, compliance, and sign-up rate, will be summarised and can be presented in the manuscript text or participant flowchart.

6.4 Epigenetic data

6.4.1 Quality control

Samples from the same participant and sample type will be processed as much as possible on the same Illumina Methylation EPIC beadchip (n=8 samples per beadchip), but time-point position will be randomised on the array to minimise batch effects. DNA methylation data quality control and preprocessing will be performed using a standardised process in the Widschwendter group. The R packages minfi, ChAMP, and ENmix will be used for preprocessing. Briefly, any samples not below a median unmethylated or methylated intensity threshold will be excluded from analysis. Probes that fail the detection p threshold ($p < 0.01$) will be excluded from the dataset. Samples that have more than 10% failed probes will be excluded from the dataset. Beta value distributions will be visualised and outliers identified (for each sample type). Dye bias correction and BMIQ normalisation will be conducted. In addition, control probes will be visualised to check for efficiency of bisulfite conversion, for instance. An amended version of the preprocessing pipeline may be utilised if ongoing experiments demonstrated improvements in reproducibility and noise.

Potential sample mix-ups will be checked by extracting information on control single nucleotide polymorphism (SNP) probes present on the EPIC array. Obvious discrepancies will be further investigated and samples may need to be excluded.

As part of quality control, sample principal components will be checked for association with age, immune cell composition, processing date, beadchip, and other potential biological and technical effects to evaluate batch effects. These will be visualised by correlation heatmaps for the first 10 principal components and principal component and uniform manifold approximation and projection (UMAP) plots. If significant non-biological batch effects are found (e.g. processing day, beadchip), these may be corrected using the ComBat function within the ChAMP R package. As part of quality control, we will also investigate correlation of WID general age with chronological age.

6.4.2 Exploratory analysis

Initial data exploration will assess the association of cancer risk- and age-associated signature scores with immune cell composition. DNAme signature scores of interest highlighted above (e.g. but not limited to WID-BC, WID-OC, WID-EC, WID-CIN, WID-REA, WID-SOLA) at month 0 will be visualised as follows:

- chronological age
- immune cell composition (ic)
- BMI
- smoking status

for each tissue, amongst other molecular and cellular analyses. It is assumed that epigenetic indices should be corrected for immune cell composition (for blood: myeloid fraction, consisting of monocyte, eosinophil and neutrophil composition) using baseline values for all participants in a given study arm.

To investigate longitudinal effects, spaghetti plots will be generated for all timepoints (baseline, M2, M4, M6; adjusted for age (across groups) and immune cell composition (ic)), grouping or colouring by the following variables:

- compliance score (colour gradient), separate for each intervention (smoking cessation, intermittent fasting with ketogenic supplement, intermittent fasting without ketogenic supplement)
- smoking status at the end of the intervention (smoking cessation only)
- delta BMI month 6 and month 0 (colour gradient) (intermittent fasting only)
- average β-hydroxybutyrate value (colour gradient) (subgroup of intermittent fasting only)

The suitability of random intercepts and slopes will also be explored during exploratory analysis. Individual slopes for scores will be plotted against covariates.

In addition to visualisation, summary tables (median and standard deviation) will be generated per group per timepoint, assessing cross-sectional (between groups) and longitudinal differences (within subjects), stratifying median compliance score.

To evaluate potential associations between changes in epigenetic scores and clinical or other factors, visualisation plots may be generated and covariate adjustment may be conducted for primary and secondary analyses, where appropriate.

6.4.3 Primary endpoint reporting

For primary endpoint analysis, paired Wilcoxon tests will be performed on adjusted WID scores associated with disease risk, cellular ageing, or exposure, in cervical samples at baseline and month 6 of the intervention (adjusted for immune cell composition). For visualisation, points and lines may be coloured by compliance score, delta BMI, or smoking status at the end of the intervention. Data will be visualised using Spaghetti plots and may additionally be visualised using violin/box plots with overlayed jitter plots.

Analyses will be presented separately per sub-study and a comparison between the two randomised arms in the diet intervention study may be conducted. The analysis is predominantly conducted ITT, but separate subgroup analyses may be conducted as follows:

- per protocol, if at least 30 individuals have a cumulative compliance score of more than 80%
- diet intervention: timing of fasting, breakfast versus dinner cancellation

Mixed-effects models could provide a complimentary method of reporting individual effects over time.

6.4.4 Secondary endpoint reporting

To investigate the magnitude and temporal dynamics of changes in methylation indices for each surrogate tissue, namely cervical, buccal, and blood, linear mixed-effects models will be estimated, combining tissues where applicable. The use of random slopes and intercepts will be explored (R package lme4 or nlme, lmer function). Compliance score, sample cell composition, BMI, smoking status, or other factors identified during exploratory data analysis may be used as covariates. Additionally, in a descriptive manner, the change from baseline will be reported within each tissue and paired comparisons will be carried out. Examples of linear models (covariates may be altered or extended) are provided below:

$$model_{intermittent\ fasting} = lmer(score \sim compliance + sample\ type + BMI + ic + age_{baseline} + (time|id), data = data) \quad (1)$$

$$model_{smoking\ cessation} = glmer(score \sim compliance + sample\ type + smoking\ status + ic + age_{baseline} + (time|id), data = data) \quad (2)$$

Separate models per sample type may be even more informative, this will be explored during exploratory data analysis.

Using models based on equations in equations 1 and 2, the data will be characterised as follows:

1. Estimation of average changes in all participants
2. Estimation of changes for individual participants
3. Characterisation of degree of heterogeneity across participants (variance across participants, quantified by random effects)
4. Identify factors (fixed effects) that predict changes (e.g. compliance score)

Beta estimation will be conducted using maximum likelihood estimation. Following model fitting, residuals will be visualised and qq plots will be generated. If p values are needed for reporting, these will be obtained via Kenward-Roger estimation in the R package lmerTest. A sensitivity analysis will be conducted using either complete case analysis or multiple imputed values

Comparison between sample types may be done via incorporation of data into mixed effects models with an additional covariate (sample type). The sample type with the biggest changes in association with the compliance score may be the most suitable surrogate type for monitoring longitudinal health effects.

Data will be visualised using spaghetti plots, colouring by compliance score and applying facet wraps by sample type.

The abovementioned aspects cover the dynamics of epigenetic indices. Additionally, it will be investigated whether epigenetic indices in certain tissues are associated with other clinical or molecular health-related biomarkers via correlation analysis. No causality can be established for these factors.

6.5 Clinical and epidemiological data

6.5.1 Quality control

Clinical data is recorded using an electronic case report form (eCRF) or databases (e.g. fitness tracker). The eCRF data entry includes a plausibility check. Data will be harmonised by the data analyst prior to exploration and missing data, and patterns thereof, will be explored.

6.5.2 Exploratory analysis

Data normality will be thoroughly visualised and summarised prior to analysis. The data distribution, including distribution of missing data, will be evaluated.

6.5.3 Secondary endpoint reporting

Paired analysis (ITT) will be conducted for baseline and month 6 measurements. As with the primary endpoint, PP or above/below median compliance score subgroup analysis may be additionally constructed. Linear-mixed effects models will be constructed to explore heterogeneity in responses among participants and leverage the strength of our longitudinal data. In addition, descriptive statistics relating to change to baseline can be reported. The following outcomes will be assessed:

- body mass index (kg/m^2), *diet arm of the study*
- fat-free body mass (%), *diet arm of the study*
- current smoking ("yes" or "no", binary response, can be analysed using a generalised estimation equation or generalised linear mixed model with a log link function)
- vascular health parameters
  - pulse wave velocity (mm/s)
  - intima-media thickness (mm)
  - plaque score
- physical activity
  - international physical activity questionnaire (IPAQ) score
  - average daily steps per week
  - active minutes per week
- abdominal fat composition, *diet arm of the study*
  - visceral fat (mm)
  - subcutaneous fat (mm)
- heart rate variability (ms)
- resting heart rate (beats per minute)
- sleep quality score (fitness tracker, average score per week)
- subgroup in diet arm of the study: β-hydroxybutyrate levels

- pulmonary health, including FEV1, FVC, and FEV1/FVC
- ergonometry functional values

Data will be reported using spaghetti plots (visualising individual values) and/or tables of mean and standard deviations. In the diet arm, changes in body mass index, fat-free body mass, and abdominal fat composition may also be compared between groups (with and without ketogenic supplement) using Wilcoxon testing of delta change, or via inclusion of ketogenic supplement covariate in linear mixed models.

6.6 Microbiome data

The following questions will be addressed:

1. How do intermittent fasting or smoking cessation influence the gut microbiome? Is there a shift in dominant types? Is there an increase in species richness? Is there an increase in the relative abundance of genera associate with a healthy phenotype, or loss of species associated with a disease phenotype?
   A comparison will be conducted between month 0 and month 6. Data will be analysed in association with the compliance score, as well as epigenetic markers and BMI. Beneficial species include probiotic and short-chain fatty acid producing microbes:

   - *Lactobacillus reuteri*, which produces reuterin, a protein oxidant that helps to protect against colorectal cancer.[16]
   - *Parabacteroides distasonis*, which can increase levels of unconjugated bile acids that activate GLP1 and UCP1 pathways and improve the glucose and energy metabolism.[17]
   - *Akkermansia muciniphilia*, which induces mucin production necessary for intestinal integrity.[18] This can also be associated with blood markers of LBP, an indicator of gut integrity.

   Conversely, harmful microbiobes include species such as *Clostridium difficile*.

2. Can the microbiome be used to distinguish lean from overweight or obese individuals, using fecal microbial taxonomic composition?
   Taxonomic composition of individuals with BMI below and above 25 will be analysed at month 0. The potential species contributing to difference may be analysed using PCA and correlation analysis.

3. Can the microbiome be used to distinguish smokers from non-smokers?
   Taxonomic composition of smokers and non-smokers will be analysed, with the caveat that non-smokers are all of BMI above 25 which may have an impact on microbiome (see above).

4. Are the changes in the microbiome associated with predicted functional alterations?

6.6.1 Quality control

Barcoded amplicons of the variable V3-V4 region of the 16S rRNA gene will be sequenced on an Illumina MiSeq platform. Paired end fastq files will be demultiplexed and checked for

remaining adaptors using FastQC. Further quality control on the raw reads and subsequent generation of Exact Sequence Variants (ESVs) can be done using the DADA2 package in R.[19] In short, for each sequence run separately, reads are trimmed and filtered, sequence error rates are estimated, read abundance is calculated after dereplication and read pairs are merged. Next, sequence sets for the different runs are merged, and chimeric sequences and singletons removed. Optionally, spurious ESVs (e.g., relative abundance in no sample $\geq 0.01\%$) can be removed from the final dataset. Taxonomic identification of the ESVs (reliable down to the genus-level) can be performed using the SILVA taxonomic database (version 138) with the RDP naïve Bayesian classifier method.[20]

6.6.2 Exploratory analysis

In addition to exploratory analysis embedded in QC, further exploratory analysis may be performed in R, using packages such as tidyverse, vegan and ampvis2. Principal component analysis (PCA) and other ordination analyses can be used to study the correlation of ESV abundance and different interventions. ESVs that explain most of the variation in the PCA can be identified using species. Multivariate statistics, such multivariate analysis of variance (PERMANOVA), on different distance measures can be used to validate the significance of correlations between interventions and the microbial community composition of the samples.

6.6.3 Secondary outcome reporting

To report on the distinction of lean and obese/overweight individuals and smokers and non-smokers, principal component plots will be generated and species plots will be provided.

To evaluate species diversity before and after intervention, richness indices such as the Shannon Diversity or inverse Simpson index will be reported before and after intervention. The relative gain or loss of beneficial and harmful individuals will be reported using ANOVA or Wilcoxon tests and visualised with barplots, and association of relative change could be associated with changes in epigenetic markers, BMI, or e.g. markers of gut leakiness (LBP).

A shift in dominant types and the association with stronger weight loss or epigenetic markers will be evaluated using heatmaps (e.g. the top 30 ASVs/genera/families). Subgroups of individuals where this shift does not occur will also be evaluated.

Functional pathway prediction will be done in R using the Tax4Fun2 package, evaluating pathways such as carbohydrate degradation, lipid metabolism, inflammatory properties, and hormone metabolism. A comparison may be conducted between individuals with and without ketogenic supplement in the diet intervention group.

6.7 Blood data

From blood samples the following data will be generated:

- Lipid levels
- HbA1c
- c-reactive protein
- General immunological profile (cell subsets)

- T cell stimulation assay
- Cytokine and markers of inflammation

6.7.1 Quality control and exploratory analysis

Data will be generated using standardised protocols. Missingness of data will be visualised.

6.7.2 Secondary endpoint reporting

Paired baseline-controlled analyses are used to compare month 0 and month 6 in each sub-study. Linear mixed-effects models (potentially random slopes/intercepts) will be utilised to leverage longitudinal data and include further covariates. Change from baseline may be reported in a descriptive manner.

6.8 Metabolomic data

NMR profiling will be conducted which provides information on over 100 metabolites, including amino acids and organic acids.

6.8.1 Quality control and exploratory analysis

Data will be generated using standardised protocols. Missingness of data will be visualised.

6.8.2 Secondary endpoint reporting

Data will be analysed in a longitudinal manner and changes associated with clinical parameters including markers of ketosis, where available, and fasting pattern (dinner or breakfast cancellation).

6.9 Skin data

Evaluation of ageing markers and barrier function may be conducted in a subgroup of participants for which skin data is available.

6.9.1 Quality control and exploratory analysis

Data will be generated using standardised protocols. Missingness of data will be visualised.

6.9.2 Secondary endpoint reporting

Paired baseline-controlled analyses are used to compare month 0 and month 6 separate for each sub-study.

6.10 Psychological data

6.10.1 Quality control and exploratory analyses

Data will be collected using standardised protocols. Missingness of data will be visualised.

6.10.2 Secondary endpoint reporting

Paired baseline-controlled analyses are used to compare month 0 and month 6 separate for each sub-study.

6.11 Subsidiary analysis: continuous compliance

Continuous compliance to intervention following the initial 6 months will be evaluate at month 6 and month 12 and reported in a summary table.

6.12 Subsidiary analysis: single cell sequencing

Single cell sequencing will be conducted to assess cell identity and population heterogeneity before and after two months of intermittent fasting, as well as chromatin accessibility and markers of senescence in cellular subgroups.

6.12.1 Quality control

Single cell analysis will follow current guidelines. A pilot experiment will be carried out to optimise sequencing depth. Sequencing data will be processed using the 10X Genomic cellranger or other available and validated tools at the time of analysis.

6.12.2 Exploratory analysis

Exploratory analysis may be conducted using Seurat for R, visualising cell heterogeneity and expression of genes across subgroups. Clustering of samples before and after intervention will be visualised.

6.12.3 Subsidiary endpoint reporting

The proportion of senescence-expressing immune cells and cells with lower-than-average PCGT chromatin accessibility will be reported. Visualisation of single cell populations before and after intervention (averaged over participants or per individual) will be carried out via two-dimensional projections such as UMAP. Changes may be taken into context with compliance of the individual.

6.13 Subsidiary analysis: epigenetic signatures of reprogramming

We will investigate whether DNAme regions or individual CpGs that change after Yamanaka factor reprogramming (OSKM) in blood cells (PBMCs) at month 0 also change during the intervention, in both smoking cessation and intermittent fasting.

6.13.1 Quality control

Cell death will be monitored in samples "before" and "after" reprogramming, and quality control will be in line with other epigenetic data.

6.13.2 Subsidiary endpoint reporting

We will investigate epigenetic changes before and after reprogramming in samples collected at baseline. The same regions identified here will be investigated in the same individuals at month 6, e.g. paired t-test mean or median methylation in significant CpGs before and after reprogramming and before and after intervention). For regions that change across participants, these can be investigated after intervention even in those individuals for whom no reprogramming analysis was conducted

6.14 Subsidiary analysis: meaningful biomarkers of intervention

Many studies have previously identified signatures that that capture epigenetic ageing yet the biological meaning of them remains unclear. It may be beneficial to identify CpG markers that are changed in response to a putative health-promoting intervention specifically and correlate with clinically meaningful changes (e.g. gut microbiome, lung function) in a well-characterised longitudinal dataset. We may chose to undertake this subsidiary analysis pending the outcome of initial results. This analysis is largely explorative and may depend on the data. DNA methylation features associated with the intervention and changes in other biologically meaningful changes (e.g. microbiome, lung function) may be identified. Scores can be investigated in a hold-out test set and ongoing follow-on projects at our institute (SUN and LIFE Tirol), which investigate smoking cessation in n=200 participants and randomised intermittent fasting in approximately n=800 participants.

6.15 Additional exploratory analyses

Post-hoc analyses requested by journal editors or referees, as well as exploratory analyses, will be labelled explicitly as such. Any further analyses not specified in the analysis protocol will be exploratory in nature and may be documented in a separate statistical analysis plan(s).

6.16 Tabular summary of analyses

Outcome	Value	Type of analysis	Study populations	Adjustments, covariates	Tissue	Notes
Baseline	Epidemiological data	Descriptive statistics	All included			
Primary	Epigenetic scores WID	Paired before-after	ITT	ic	cervix	Comparison randomised arms between and subgroups for different fasting patterns
	Epigenetic scores WID	Paired before-after	Per protocol (≥80% compliance) or above/below median compliance score	ic	cervix	Comparison randomised arms between and subgroups for different fasting patterns
	Epigenetic scores WID	Longitudinal (linear/non-linear mixed-effects models)	ITT	ic, age, e.g. BMI, compliance score	cervix	
Secondary	Epigenetic markers WID	Longitudinal (linear/non-linear mixed-effects models)	ITT	ic, BMI, age, other markers associated	cervix, buccal, blood - separate and combined	
	Epigenetic markers WID	Longitudinal (linear/non-linear mixed-effects models)	Per protocol (≥80% compliance) or above/below median compliance score	ic, BMI, age, other markers associated	cervix, buccal, blood - separate and combined	
	Clinical and epidemiological markers	Paired before-after	ITT	compliance, covariates identified during exploration	n/a	Association with epigenetic and genetic factors
	Clinical and epidemiological markers	Paired before-after	Per protocol (≥80% compliance) or above/below median compliance score	compliance, covariates identified during exploration	n/a	Association with epigenetic and genetic factors
	Clinical and epidemiological markers	Longitudinal (linear/non-linear mixed-effects models)	ITT	compliance, covariates identified during exploration	n/a	Association with epigenetic and genetic factors
	Clinical and epidemiological markers	Longitudinal (linear/non-linear mixed-effects models)	Per protocol (≥80% compliance) or above/below median compliance score	compliance, covariates identified during exploration	n/a	Association with epigenetic and genetic factors
	Mutation data	Paired before-after	ITT	compliance, covariates identified during exploration	blood	Association with epigenetic factors
	Mutation data	Paired before-after	Per protocol (≥80% compliance) or above/below median compliance score	compliance, covariates identified during exploration	blood	Associateion with epigenetic factors
	Microbiome data	Baseline comparison	Lean versus obese, smokers versus non-smokers		fecal	

Microbiome data	Paired after	before-	ITT	compliance, weight loss, covariates identified during exploration	fecal	
Microbiome data	Paired after	before-	Per protocol (≥80% pliance) or above/below median compliance score	compliance, weight loss, covariates identified during exploration	fecal	
Blood data	Paired after	before-	ITT	compliance, weight loss, covariates identified during exploration	blood	Association with epigenetic and clinical factors
Blood data	Paired after	before-	Per protocol (≥80% pliance) or above/below median compliance score	compliance, weight loss, covariates identified during exploration	blood	Association with epigenetic and clinical factors
Blood data	Longitudinal (linear/non-linear mixed-effects models)		ITT	compliance, weight loss, covariates identified during exploration	blood	Association with epigenetic and clinical factors
Blood data	Longitudinal (linear/non-linear mixed-effects models)		Per protocol (≥80% pliance) or above/below median compliance score	compliance, weight loss, covariates identified during exploration	blood	Association with epigenetic and clinical factors
Metabolomic data	exploratory, before-after				urine, saliva	Association with ketone levels, epigenetic markers
Skin data	Paired after	before-	ITT	compliance, weight loss, covariates identified during exploration	skin biopsy	Association with epigenetic factors
Skin data	Paired after	before-	Per protocol (≥80% pliance) or above/below median compliance score	compliance, weight loss, covariates identified during exploration	skin biopsy	Association with epigenetic factors
Psychological data	Paired after	before-	ITT		n/a	
Psychological data	Paired after	before-	Per protocol (≥80% pliance) or above/below median compliance score		n/a	
Subsidiary						
Continuous compliance	Descriptive statistics					
Single-cell sequencing	Paired after	before-	ITT (subgroup of all participants included for single-cell sequencing)	compliance, ketone bodies, factors identified during exploration	blood	Association with epigenetic factors
Reprogramming	Paired after	before-	ITT	compliance, ketone bodies, factors identified during exploration	blood	Association with epigenetic factors
Biomarkers of intervention	Exploratory, longitudinal		all		all	

7 Systems medicine approach

In addition to the above, we also intend to assess which of the WID scores that dynamically change as a function of protocol adherence are correlating with the change of any of the non-methylation parameters, for example association of microbiome markers in with epigenetic age (WID-SOLA markers).

Briefly, we aim to conduct three analyses:

1. Evaluation of dynamics of prespecified epigenetic markers (WID scores) within the three surrogate tissues cervix, blood, and buccal, in order to assess whether changes are concordant or discordant, both directional and temporal (part of secondary outcome analyses)

2. Investigation of prespecified epigenetic markers and their correlation to non-methylation parameters, for instance whether microbiome or lung function markers are associated with any cellular/immune ageing markers (WID-RIA/WID-SOLA)

3. Agnostic approach, where non-methylation parameters are the outcome and predictors are defined based on DNAme to identify whether CpG sites which predict the non-methylation features (part of subsidiary exploration).

8 Publication outline

For each sub-study (smoking cessation and intermittent fasting), one main paper will be prepared describing the overall effects on a variety of organ systems. Key aspects to epigenetic and clinical effects will be presented in the main data, while additional outcomes may be presented in supplementary data, although this will need to be confirmed.

For each study, a participant flowchart diagram as well as overview for compliance will be shown. The following characteristics will be described for all participants at the beginning (and at the end) of the study, separately for each intervention:

- age
- ethnicity
- highest education
- menopausal status
- parity
- use of menopausal hormone replacement therapy (postmenopausal only)
- use of oral contraceptives
- body mass index (kg/m^2)
- body composition from bioelectric impedance analysis (% mass), *diet group only*
- abdominal fat (visceral and subcutaneous, both in mm)
- resting heart rate
- smoking status (current/ex/never smoker)
- smoking pack year

Data will be reported in a table, describing the parameters using mean or median and standard deviation or quantiles, respectively, or proportions (%) depending on data type and distribution.

9 Appendix

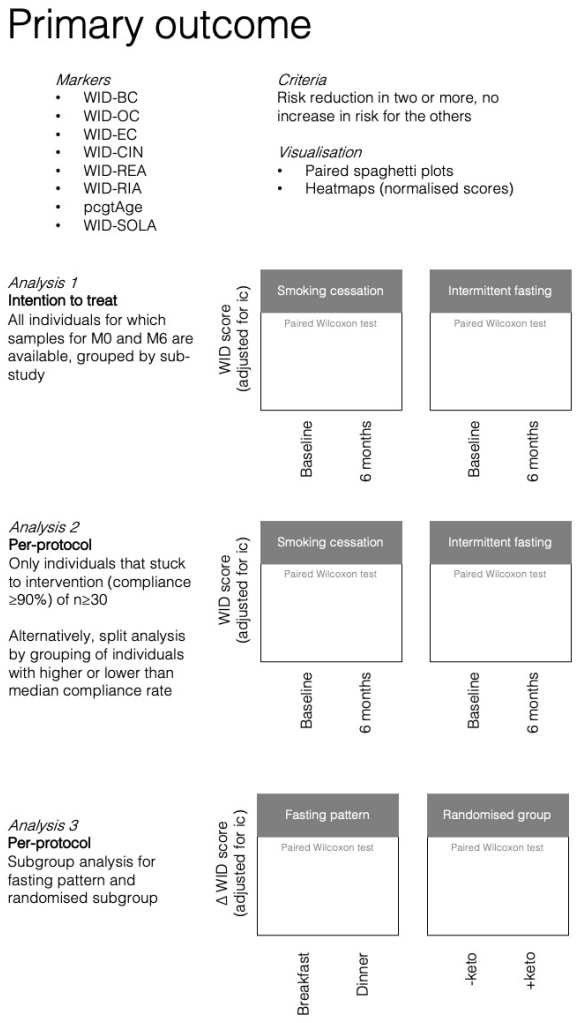

Figure 1: Graphical schematic of primary outcome evaluation.

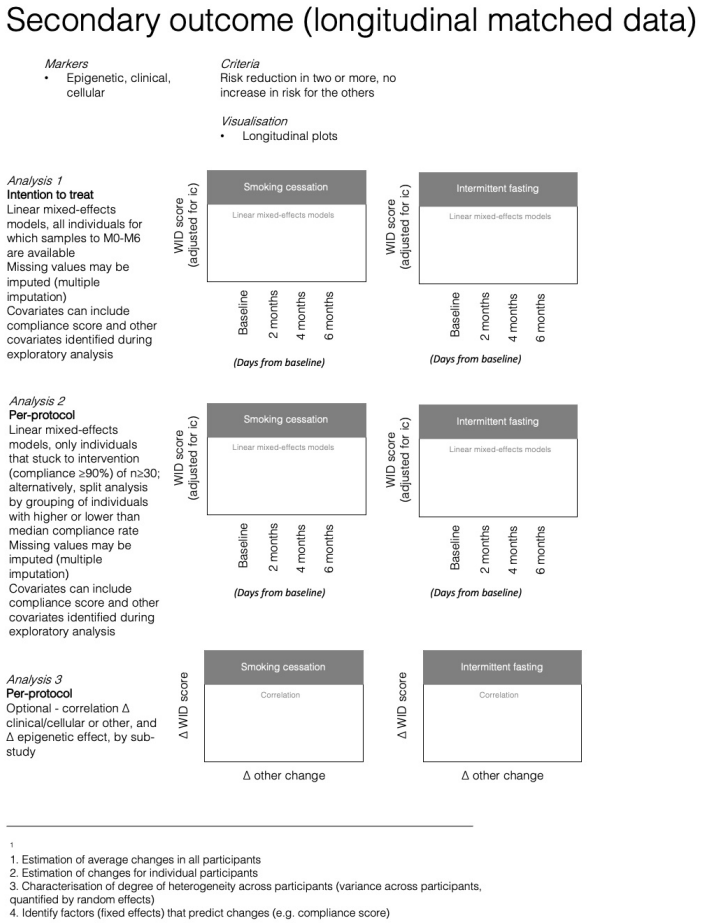

Figure 2: Graphical schematic of secondary outcome evaluation.

References

[1] Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. Annals of Internal Medicine. 2013;158(3):200.

[2] Collaborators GCRF. The global burden of cancer attributable to risk factors, 2010–19: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2022;400(10352):563–591.

[3] Jacob L, Freyn M, Kalder M, Dinas K, Kostev K. Impact of tobacco smoking on the risk of developing 25 different cancers in the UK: a retrospective study of 422,010 patients followed for up to 30 years. Oncotarget. 2018;9(25):17420–17429.

[4] Steck SE, Murphy EA. Dietary patterns and cancer risk. Nature Reviews Cancer. 2020;20(2):125–138.

[5] Key TJ, Allen NE, Spencer EA, Travis RC. The effect of diet on risk of cancer. The Lancet. 2002;360(9336):861–868.

[6] Barrett JE, Herzog C, Jones A, Leavy OC, Evans I, Knapp S, et al. The WID-BC-index identifies women with primary poor prognostic breast cancer based on DNA methylation in cervical samples. Nat Commun. Feb 1;13(1):449.

[7] Barrett JE, Jones A, Evans I, Reisel D, Herzog C, Chindera K, et al. The DNA methylome of cervical cells can predict the presence of ovarian cancer. Nat Commun. Feb 1;13(1):448.

[8] Longo VD, Tano MD, Mattson MP, Guidi N. Intermittent and periodic fasting, longevity and disease. Nature Aging. 2021;1(1):47–59.

[9] Jamshed H, Beyl RA, Manna DLD, Yang ES, Ravussin E, Peterson CM. Early Time-Restricted Feeding Improves 24-Hour Glucose Levels and Affects Markers of the Circadian Clock, Aging, and Autophagy in Humans. Nutrients. 2019;11(6):1234.

[10] Acosta-Rodríguez V, Rijo-Ferreira F, Izumo M, Xu P, Wight-Carter M, Green CB, et al. Circadian alignment of early onset caloric restriction promotes longevity in male C57BL/6J mice. Science. 2022;376(6598):1192–1202.

[11] Gershuni VM, Yan SL, Medici V. Nutritional Ketosis for Weight Management and Reversal of Metabolic Syndrome. Current Nutrition Reports. 2018;7(3):97–106.

[12] Roberts MN, Wallace MA, Tomilov AA, Zhou Z, Marcotte GR, Tran D, et al. A Ketogenic Diet Extends Longevity and Healthspan in Adult Mice. Cell Metab. Sep 5;26(3):539–546 e5.

[13] Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomized Trials. Annals of Internal Medicine. 2010;152(11):726.

[14] Althouse AD. Adjust for Multiple Comparisons? It's Not That Simple. The Annals of Thoracic Surgery. 2016;101(5):1644–1645.

[15] Naimi AI, Cole SR, Kennedy EH. An introduction to g methods. International Journal of Epidemiology. 2016;46(2):756–762.

[16] Bell HN, Rebernick RJ, Goyert J, Singhal R, Kuljanin M, Kerk SA, et al. Reuterin in the healthy gut microbiome suppresses colorectal cancer growth through altering redox balance. Cancer Cell. 2022;40(2):185–200.e6.

[17] Li M, Wang S, Li Y, Zhao M, Kuang J, Liang D, et al. Gut microbiota-bile acid crosstalk contributes to the rebound weight gain after calorie restriction in mice. Nature Communications. 2022;13(1):2060.

[18] Bodogai M, O'Connell J, Kim K, Kim Y, Moritoh K, Chen C, et al. Commensal bacteria contribute to insulin resistance in aging by activating innate B1a cells. Science Translational Medicine. 2018;10(467).

[19] Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet C, Al-Ghalith GA, et al. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. PeerJ Preprints; 2018.

[20] Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology. 2007;73(16):5261–5267.