STATISTICAL ANALYSIS PLAN (SAP)

Increase in colonic PRopionate as a method of prEVENTing weight gain in adults aged 20-40 years (**iPREVENT**)

SAP Version No. 2.0

Chief Investigator:

Professor Gary Frost (CI)

SAP Working Group:

Joana Vasconcelos and Toby Prevost

Table of Contents

page

ABBREVIATIONS

AE	Adverse Event
ANCOVA	Analysis of Covariance
AUC	Area Under the Curve
BMI	Body Mass Index
BP	Blood Pressure
^{13}C	Carbon-13
CACE	Complier Average Causal Effects
CCK	Cholecystokinin
CONSORT	Consolidated Standards of Reporting Trials
95% CI	95% Confidence Interval
CTU	Clinical Trials Unit
CVD	Cardiovascular Disease
DBP	Diastolic Blood Pressure
DMEC	Data Monitoring and Ethics Committee
eCRF	Electronic Case Report Form
ECG	Electrocardiogram
EDC	Electronic Data Capture
FM	Fat Mass
FMI	Fat Mass Index
FFM	Fat Free Mass
GCP	Good Clinical Practice
GLP-1	Glucagon-like Peptide 1
HLGT	High Level Group Term
ICH	International Conference on Harmonisation
ICTU	Imperial Clinical Trials Unit
IPAQ	International Physical Activity Questionnaire
IQR	Interquartile range
IPE	Inulin-Propionate Ester
ITT	Intention To Treat

LDL	Low Density Lipoprotein
LLT	Lowest Level Term
LME	Linear Mixed Effects
MAR	Missing At Random
MedDRA	Medical Dictionary for Regulatory Activities
MET	Metabolic Equivalent of Task
PT	Preferred Term
PYY	Peptide YY
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SBP	Systolic Blood Pressure
SCFA	Short Chain Fatty Acid
SD	Standard Deviation
SE	Standard Error
SOC	System Organ Class
SOP	Standard Operating Procedure
TSC	Trial Steering Committee
UK	United Kingdom
VAS	Visual Analogue Scales

1. INTRODUCTION

1.1. Purpose and scope of the statistical analysis plan

The purpose of this Statistical Analysis Plan is to set out the study objectives and hypotheses, and the analytical approaches and procedures necessary to address these for the main trial paper and to provide guidance for further research reported in other papers, promoting consistent approaches and methods.

As there can typically be more than one analytical approach to address a hypothesis, there is the potential for different results to be produced from using alternative approaches, alternative methods, alternative outcome definitions and the alternative data involved in analyses. These differences can be influential, for example, when results are of borderline statistical significance or change clinical interpretation. The presence of an approved SAP enables choices to be more fully discussed and justified in advance of the data analysis.

Therefore, the approved Version 1.0 of this SAP records those decisions that can be made about study hypotheses, outcome definitions and statistical procedures, along with their basis and the appropriateness of the assumptions required for their use, in advance of the main trial analysis, while any access to follow-up data by trial arm is prevented. Changes made within any subsequent versions of the Plan prior to analysis will be listed and dated, with the basis for the changes reasoned, and recorded within the plan for re-approval (e.g. from a change to the protocol, or identification of a more appropriate analysis method).

According to the MHRA 'Grey Guide' and ICH E9, blinded reviews of trial data can be used for the development of the SAP (1, 2). These are generally between final data lock and the beginning of the final analysis. Analysis decisions can be made at an earlier point, based on viewing the observed distribution of the data. In this trial, the trial statistician will be unblinded to produce open and closed DMEC reports and the senior statistician will remain blinded to closed reports and participant arm. The analysis decisions at early review points around DMEC meetings will be made without access to the primary timepoint of analyses by arm. Other decisions may inevitably be made later, involving the blinded statistician, or potentially after unblinding. However, these will be documented by trial stage in the SAP and the final clinical report, and supported by reasoning and justification as mentioned above. It is not intended that the strategy set out in the plan should prohibit sensible practices. However, the principles established in the plan will be followed as closely as possible when analysing and reporting the trial.

1.2. Derivation of the statistical analysis plan

The present statistical analysis plan was derived from the trial protocol version no. 4.1, by the trial statistician, Joana Vasconcelos. The trial statistician is responsible for the development of the SAP as well as for carrying out the statistical analysis for interim and final statistical reporting of the trial. The senior trial statistician will suggest revisions to the SAP and ensure an overall verification of the analysis throughout the study, in keeping with the Standardised Operating Procedures (SOPs) of the Imperial Clinical Trials Unit, including the SOP for developing the Statistical Analysis Plan.

The formation of this Plan has drawn on statistical guidance from: the ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trial E9 and E3(2, 3), the CONSORT statement for reporting trials(4), the Committee for Medicinal Products for Human Use (formerly known as the Committee for Proprietary Medicinal Products) (5).

The first version of the plan was numbered draft version 0.1. Draft Version 0.2 incorporated comments from the Operations manager, Trial Manager, Chief Investigator and DMEC, and both statisticians having discussed points to be agreed as decisions or as queries for investigators. The plan was then sent to the TSC for comments and saved as draft version 0.3. When this was agreed and approved by sign-off it became SAP Version 1.0. Any further amendments will require a minimum of TSC approval, and involve opening up a draft version 1.1 (and if needed 1.2 and so on), and which on agreement will generate re-approved signed-off SAP Version 2.0 (and if further amended then to approved SAP Version 3.0 and so on). The signed-off SAP versions (from 1.0) will be saved in the Statistical Master File.

2. OVERVIEW OF THE CONDITION AND TREATMENT

2.1. Overview of the condition and its importance/scale

This trial aims to evaluate an intervention to prevent weight gain in adults.

Weight gain is related to many negative health outcomes. Overweight and obesity affects over 60% of the UK population and drives the prevalence of a number of common co-morbidities, including type 2 diabetes, cardiovascular disease and cancer. It is estimated that the cost of obesity to the UK economy was £27million in 2015 and this expenditure is set to rise, as 700,000 new cancer diagnoses directly caused by being overweight or obese are predicted by 2035, whilst the number of people living with diabetes in the UK has topped 4 million for the first time(6).

Weight gain occurs commonly throughout adulthood and it is adults who are at the greatest risk of substantial gains in body weight. The study intervenes in the population of adults (aged 20 to 40) who are overweight and close to obesity (having a body mass index 24-27kg/m^2 if South Asian, or 25-30kg/m^2 otherwise) who are at greatest risk of substantial increases in body weight (through recent body weight gain and/or related lifestyle behaviours that are self-reported to be low [see inclusion criteria]).

Although there has been a great focus on treatments for obesity, there has been comparatively very little on prevention. ***A greater emphasis on the prevention of obesity at the population level is urgently required***, as recent findings indicate that current non-surgical obesity treatments are ineffective, as once an individual becomes obese the probability of returning to a normal body weight is extremely low (1 in 210 for men and 1 in 124 for women)(7).

A recent systematic review indicates that any treatment or strategy that can prevent the incremental upward trajectory in body weight currently observed in young adults will have significant benefits to long-term health in this population(8).

2.2. Description of the standard treatment (or placebo or current care)

The comparator intervention in this trial is a placebo supplement called inulin given to participants in 10g sachets to take one sachet per day, at any time with their normal diet for 12 months. Inulin is a naturally occurring non-digestible carbohydrate that is readily fermented by the gut microbiota in the colon. Different terms can be used interchangeably for the placebo arm, such as control, comparator, standard, inulin control, inulin placebo.

2.3. Description of the investigational treatment

The investigational treatment (or intervention) in this trial is the consumption of supplemented IPE (Inulin-Propionate Ester) given to participants in 10g sachets to take one sachet per day, at any time with their normal diet for 12 months. Different terms can be used interchangeably for the active arm, such as IPE, intervention, investigative, treatment, inulin propionate.

2.4. Description of the motivation for the study / mechanisms supporting the need to investigate the new treatment

IPE is a proposed alternative supplement to the consumption of dietary fibre. Dietary fibre, when consumed, can have a considerable beneficial impact on increasing the concentrations of SCFA (short chain fatty acids) produced in the large bowel, but with adverse Gastrointestinal side effects. The impact is beneficial because increased SCFA promote satiation (through the release of Gastrin directly, and indirectly through releasing Cholecystokinin). The previous trial (9), which was preliminary to this one, indicated a low rate of gastrointestinal adverse events associated with IPE. Other previous research indicated a beneficial effect also on increased energy expenditure associated with IPE, which was thought to arise via enhanced hepatic lipid oxidation. This is a hypothesised mechanism (from rodent models) that the IPE will modulate hepatic lipid metabolism to augment lipid oxidation and supress De Novo Lipogenesis.

2.5. Overview of the mechanistic sub-study

A sub-sample of study participants (volunteers) will form the mechanistic sub-study. This will enable us to understand the mechanistic network behind the primary objective of the study, and the effects of inulin-propionate ester supplementation (IPE) on small changes in energy balance (colonic metabolism, appetite regulation and energy expenditure).

3. STUDY OBJECTIVES / HYPOTHESES TESTING

3.1. Principal Trial objective

The principal objective of the trial is to investigate the effect of increasing colonic propionate production through randomised supplementation of IPE on preventing weight gain, in a population of adults who are at the greatest risk of substantial increases in body weight.

3.2. Principal Research Question

The principal research question is as follows: Is weight gain after 12 months different following IPE supplementation compared to placebo inulin?

3.3. Hypotheses

The hypotheses refer to the underlying populations of relevant patients rather than the actual study subjects.

The *Working hypothesis*: The so-called "working hypothesis" is the hypothesis which motivates the trial, which the trial results may or may not support. It is that mean body weight gain is lower at 12 months follow-up in the IPE population compared to the inulin population.

The Statistical *Null Hypothesis*: There is no difference in mean body weight gain at 12 months in a population randomised to receive IPE compared to that randomised to receive inulin.

Statistical *Alternative hypothesis*: There is a difference in mean body weight gain at 12 months in a population randomised to receive IPE compared to that randomised to receive inulin.

3.4. Primary Objective

To measure weight gain following a 12 month intervention of IPE versus inulin control to allow between-group comparison.

3.5. Secondary Objectives

1. To determine the safety profile of IPE

2. To determine effects on glucose homeostasis as a surrogate marker of type-2 diabetes risk

3. To determine effects on blood lipid and cholesterol as surrogate markers of cardiovascular disease (CVD) risk

4. To determine effects on blood pressure as a surrogate marker of CVD and stroke risk.

5. To compare changes in body weight/waist/BMI/body composition during the 12 month intervention

6. To determine compliance (sachet count) during the 12 month intervention.

3.6. Changes in lifestyle factors

Some of the following lifestyle factors can alter weight (e.g. quitting or taking up smoking, or a marked change in physical activity), but would not be thought to have arisen from the intervention.

1. To compare changes in physical activity during the 12-month intervention.

2. To compare changes in other lifestyle factors during the 12-month intervention; smoking, drinking and recreational drugs.

3. To compare changes in diet during the 12-month intervention (via food diaries).

Notes:
Physical activity: Measure of level and time spent on activities by a person, according to the IPAQ
Smoking: Cigarettes, cigars or vaping
Drinking: Alcoholic drinks of any type
Recreational drugs: Legal and illegal drugs taken for recreational purposes in any form

3.7. Mechanistic sub-study objectives

1. To explore the effects of IPE on colonic metabolism using metataxonomic analysis of the 16S ribosomal ribonucleic acid (rRNA) gene in stool samples to identify the relative abundance of the bacterial component of the microbiome.

2. To explore the effects of IPE on the metabolite profile using nuclear magnetic resonance spectroscopic analyses. These data will be used to determine how these specific changes in the colonic environment influence L-cell differentiation using a human organoid model.

3. To explore the effects of IPE on anorectic gut hormones (GLP-1, PYY, gastrin and CCK) and subjective feelings of appetite via visual analogue scales (VAS), as measures of appetite regulation.

4. To explore the effects of IPE on energy expenditure and hepatic lipid metabolism as potential mechanisms involved in body weight maintenance.

Notes/clarification of terms:

Insulin: Hormone produced by the pancreas and lowers blood glucose levels.

Glucagon: Hormone produced by the pancreas that raises blood glucose levels. Type 2 diabetes affects and so limits the ability of insulin to maintain the balance – when glucose goes up after food, the insulin response is not there to reduce it shortly afterwards.

Homeostasis: The state of steady internal conditions. Glucose homeostasis is a surrogate marker of Type 2 diabetes risk.

Colonic metabolism: The breakdown of products in the colon (large intestine), into metabolites.

Metabolite profile: The measurement of products of metabolism, in the gut and/or blood.

GLP-1: Glucagon-like peptide-1 is secreted by L cells in the ileum and colon, in response to a meal and enhances the secretion of insulin to maintain blood sugar levels.

PYY: Peptide tyrosine tyrosine is secreted by L cells in the ileum and colon, in response to a meal, and reduces appetite.

Gastrin: A peptide hormone which stimulates the secretion of gastric acid, from G cells in the gut.

CCK: Cholecystokinin is a peptide hormone released by the small intestine, responsible for stimulating the digestion of fat and protein. Also suppresses appetite.

Feelings of appetite: The desire to eat food.

Energy expenditure: The total number of calories (energy) that a person needs to carry out a physical function such as breathing, digesting food etc.

Hepatic lipid metabolism: Lipid metabolism by the liver - the conversion of excess carbohydrates and proteins into fatty acids and triglycerides.

3.8. Supplementary investigation

As written in the grant application, "this trial has both a pragmatic element, to answer the question of whether the policy of prescribing and uptake of inulin propionate ester as specified in the trial will reduce further weight gain compared with control, and an explanatory element to understand the mechanisms on the causal pathway of such weight change and any limitations from compliance". Therefore, we will conduct a supplementary investigation to the primary objective of the trial, where we are interested to find out whether the treatment effect may be much larger amongst a complier group of subjects. According to ICH-E9 (R1) "*Supplementary analyses for an estimand can be conducted in addition to the main and sensitivity analysis to provide additional insights into the understanding of the treatment effect. They generally play a lesser role for interpretation of trial results"*(10).

4. DESIGN

4.1. Study Design

This is a two-centre, double-blind parallel group randomised controlled trial with one-year follow-up that will test the superiority of inulin propionate over inulin placebo on weight gain prevention at 12-months in 270 eligible participants aged 20-40 years who are overweight but not obese and at a high risk of weight gain.

4.2. Treatment Groups

The trial is randomised with two arms and with equal allocation of participants in a 1:1 ratio to these arms. **Active Arm**: IPE for 12 months; **Placebo Arm**: Inulin for 12 months.

4.3. Eligibility Criteria

Inclusion criteria:

1. Males and Females aged 20-40 years
2. Body Mass Index (BMI) of 24-27kg/m^2 if of South Asian ethnicity or 25-30kg/m^2 if non-South Asian, and at least one of the following:
    - A self-reported weight gain of 2kg over the last 12 months
    - Low self-reported physical activity ('low' activity as per International Physical Activity Questionnaire - IPAQ)
    - Low self-reported fruit and vegetable intake (<2 servings of fruit and vegetables per day)
    - Self-reported high intake of sugar sweetened beverages (>1 serving per day)
3. On stable medication (if taking any) at point of screening
4. Written informed consent

Exclusion criteria:

1. Diagnosed chronic disease; Type I and II diabetes, cancer, renal failure, heart disease, organic acidaemia (propionic acidaemia, methyl malonic acidaemia)
2. Diagnosed gastrointestinal condition including coeliac disease, inflammatory bowel disease and irritable bowel syndrome
3. Previous bowel reconstruction surgery
4. Pregnancy or lactation
5. Use of antibiotics at any time in the past 3 months
6. Anaemic and Vitamin B12 deficient (<160 ng/L)
7. Taking part in a weight loss program or consuming a weight loss product
8. Have lost 3kg in the last 3 months
9. Any other gastrointestinal upset (such as diarrhoea/constipation in the last 2 weeks, abdominal cramping etc.)
10. Any other reason in the opinion of the investigator

4.4. Blinding

Participants will be masked to group allocation. They will receive identical-looking and identically packaged trial intervention of either active IPE or inulin control. Clinicians/outcome assessors will also be blinded to patients' treatment assignment. The trial statistician will have access to the accumulating outcome data and trial arm that is required for reporting to the DMEC. The trial statistician will attend both open and closed DMEC meetings. The senior statistician will remain blinded and attend the open meeting.

4.5. Sample Size

On the basis outlined below, a sample size of 270 randomised participants (135 per arm) was chosen to provide 90% power to detect a 2kg difference between arms in mean body weight change over 12 months using a two-sided 5% significance test, assuming a 4.35kg SD and with 25% dropout allowance.

Determination of the primary outcome effect size:

In the randomised proof of concept trial, the difference between arms in the change in body weight over 24 weeks was 1.4kg (95% CI: -0.3 to 3.1), p=0.099. Using a Bayesian method recommended for preliminary trials in which evidence in the 95% CI is translated into probabilities, there was a 95% posterior probability of an underlying positive between-arm difference favouring the intervention (11). The posterior probability of intervention-favouring differences greater than 1kg, than 1.5kg, and 2kg were respectively 69%, 47% and 25% based on 24-week intervention. The difference increased in magnitude through successive 8-week, 16-week, and 24-week timepoints. By 24 weeks there were significant reductions in the proportion of intervention participants gaining ≥3%, and ≥5% of body weight from a mean baseline of 90kg. We therefore chose a 2kg between-arm 12-month effect size.

Determination of the primary outcome variability:

The chosen detectable effect size of 2kg agreed with that in a weight gain prevention trial conducted over 9 months in young adults (12) by also aiming to detect a 2kg effect, and achieved 4.3kg with 81% retention. The pooled standard deviation (SD) for body weight change was 4.35kg.

Power to detect the effect:

We set a 90% power to detect the primary outcome effect size of 2kg.

Determination of the sample size based on the primary outcome:

The sample size was set to be **270 participants, 135 per arm**. This involves a 25% allowance loss of 12-month primary outcomes, and two-sided statistical testing at the 5% significance level, implying that 202 subjects are required to be followed up with a valid primary outcome. Further adjustment for baseline body weight is expected to improve the precision of the estimated treatment effect.

Power calculation for Mechanistic outcomes:

For the mechanistic studies, 34 volunteers (17 per group) would provide sufficient statistical power to detect a 15 pmol/L effect size in GLP-1 concentrations between groups, with 90% power, 5% significance value, SD 13 pmol/L. These differences are based on our previous published findings that report enhanced gut hormone release following IPE supplementation (13, 14). We therefore planned to recruit a subsample of **52 volunteers (26 per group)** to allow a 70% retention rate.

Sample size calculations were performed using nQuery Advisor 4.0 software and independently validated using Stata software.

4.6. Schedule of Time and Events

A full schedule on the timing of measures is provided in below (as in section 6.5 of protocol). The plus or minus sign are a target/guide for when the visits should happen. If any fall below/above these, the closest measurement to the visit will be used.

Assessment	Screening	Baseline/ Randomisation (1-4 weeks after screening)	2 months (+/- 2 week)	6 months (+/- 4 weeks)	12 months (+/- 4 weeks)
Consent	X				
Demographics	X				
Randomisation		X			
Medical History	X	X	X	X	X
Concomitant medications	X	X	X	X	X
Pregnancy test - females only	X	X	X	X	X
Electrocardiogram (ECG)	X				
Vital signs: DBP, SBP and HR	X	X	X	X	X
Trial intervention		X	X	X	
Height, body weight, waist/hip measurements, BMI, body composition	X	X	X	X	X
Fasting blood test (glucose, insulin, lipid profile)		X		X	X
Full Blood Count & Vitamin B12	X				
Food diary		X	X	X	X
IPAQ	X	X	X	X	X
Lifestyle questions	X	X	X	X	X
Sachet count (compliance)			X	X	X
Adverse event tracking		X	X	X	X

Mechanistic evaluation (52 ppts; Imperial site only)					
Assessment	**Screening**	**Baseline/Randomisation (1-4 weeks after screening)**	**2 months (+/- 2 week)**	**6 months (+/- 4 weeks)**	**12 months (+/- 4 weeks)**
Energy expenditure (indirect calorimetry)		X			X
Appetite regulation (VAS, food diary, ad libitum test meal, and blood tests for anorectic gut hormones)		X			X
Substrate oxidation/DNL (via stable isotype tracers in water consumption) – ^{13}C breath, urine and blood samples		X			X
Gut microbiota (stool sample and hydrogen breath test)		X			X
Neuroendocrine cell number (stool sample)		X			X
Accelerometry		X			X

4.7. Timing of final analysis

Apart from DMEC reports, all analyses will be undertaken after the completion of the trial (last patient, last visit) (see DMP v1.0 (15)). Final analysis for the primary and secondary outcomes will be firstly reported collectively for the main report/publication of the trial. The mechanistic sub-sample study outcomes will be analyzed and reported at a later stage.

4.8. Randomisation

Participants will be randomised via a web-based randomisation system linked to the trial electronic data capture (EDC) system and database, called Sealed Envelope (https://www.sealedenvelope.com/). Randomisation will be undertaken using the method of minimisation with a random element in order to balance the arms by centre, sex, BMI within ethnicity (South Asians: 24.00-25.49 kg/m^2 and 25.50-27.00 kg/m^2 / non-South Asians: 25.00-27.49 kg/m^2 and 27.50 – 30.00 kg/m^2) and whether they volunteer to take part in the mechanistic sub study. This will tend towards giving equal numbers assigned to both arms in each category.

5. POPULATIONS OF ANALYSIS SETS

5.1. Target Population

The *target population*, to which inferences from the end of this trial are intended to generalise, is the population of adult patients at greatest risk of substantial increase in already overweight body weight.

5.2. Trial Population

The trial population, from which the study sample is drawn, is further defined to be those adults aged 20-40 recruited from trial centres who are either South Asian with a body mass index 24-27kg/m^2 or non-South-Asian with a BMI of 25-30kg/m^2, and who had a recent body weight gain and/or related lifestyle behaviours that are self-reported to be low, and met the remaining eligibility criteria.

5.3. Trial Samples

5.3.1 Intention To Treat (ITT)

The achieved trial sample comprises those patients who consent to participate and are actually randomised into this trial. These patients are the study subjects.

This randomised trial sample is also the trial Intention To Treat (ITT) population. The intention-to-treat principle states that every subject will be analysed according to the treatment group to which they were randomised. In this trial, subjects' data will be analysed according to the *Intention to Treat Strategy* (16), under which at least one analysis of the primary outcome is recommended to be based on all individuals in the ITT population.

The trial ITT population comprises all randomised participants, regardless of eligibility (inclusion/exclusion) error, post-randomisation withdrawal, and whether the correct study treatments were received or taken, or other interventions received.

5.3.2 Mechanistic sub-sample

The mechanistic sub-sample is not a random sample of the study participants, as the eligibility is restricted to those from the Imperial study site. Those in mechanistic sub-sample will need to consent by accepting an invitation to have additional assessments at baseline and 12 months in order additionally to address the part of the research relating to mechanisms. However, inclusion in the sub-sample is determined prior to randomization, enabling randomised comparisons within the mechanistic sub-sample.

5.3.3 Compliant sample for the supplementary investigation

This refers to the supplementary investigation where we are interested to find out whether the treatment effect may be larger amongst a complier group. This sample will comprise of all subjects in the trial who were confirmed to be compliant with eligibility criteria and those in the intervention arm who were at least 50% compliant with the IPE and control subjects who were otherwise similarly compliant.
Since there are no previous studies as to what should be considered 'adherence/or compliant to the intervention', we will investigate compliance not only on those taking 50% of the sachets provided, but across a range of compliance with the trial medication from 50% to 95%, in steps of 5%.

5.3.4 Safety sample

The safety analysis sample includes all randomised participants exposed to at least one dose of randomised treatment.

6. ANALYSIS VARIABLES

Many of the study variables are measured at more than one visit (section 4.6 above). In the reporting of the analysis the latest timepoint will be taken to be most important. Usually this will be at 12 months.

The trial is powered on the outcome measure of body weight at 12 months. A p-value will be provided for this hypothesis test. For this and other outcomes 95% confidence intervals will be presented.

6.1. Primary outcome variable

Change from baseline in body weight at 12-month follow-up.

6.2. Clinical and anthropometric outcome variables

The secondary outcome measures are listed as follows according to the type of the variable (continuous, categorical) and the sample (full sample, mechanistic sub-sample). Each will be analysed at the 12-month principal timepoint of follow-up but may also be presented at earlier follow-up timepoints (2 and 6 months)

for the purpose of secondary analysis. These will be listed at a later point when justifying the analysis method to be used.

Secondary objectives	Variables (all in full sample)	Type of variable	Timepoint(s) of evaluation of endpoint
1. To determine effects on **glucose homeostasis** as a surrogate marker of type-2 diabetes risk	Fasting biochemistry: -Glucose (mmol/L) -Insulin (mIU/L)	Continuous	6 and 12 months
2. To determine effects on **blood lipid and cholesterol** as surrogate markers of CVD risk)	Fasting biochemistry: - Triglycerides (mmol/L) - Total cholesterol (mmol/L) - LDL cholesterol (mmol/L) - HDL cholesterol (mmol/L)	Continuous	6 and 12 months
3. To determine effects on **blood pressure** as a surrogate marker of CVD and stroke risk	Blood pressure (mean of 3 measurements each): - SBP (mm/Hg) - DBP (mm/Hg)	Continuous	2, 6 and 12 months
4. To compare changes in **body weight/ waist/ BMI/body composition** during the 12 months intervention	- Body weight - Waist/hip - BMI (kg/m^2) - Body composition: Fat mass/Total body fat (FM) (kg) Fat Mass Index (FMI) (kg/m^2) Percent body fat (Fat%) (taken from Body Impedance machine) Fat free body mass (FFM) (kg) FM/FFM ratio	Continuous	2, 6 and 12* months

* Body weight already included as primary outcome at 12 months

6.3. Lifestyle factors' variables

Lifestyle factors	Variables	Type of variable	Timepoint(s) of evaluation of endpoint
A. To compare changes in **physical activity** during the 12 months intervention	**IPAQ** (International Physical Activity Questionnaire) scores (**all participants**) - Low - Moderate - High	Categorical	2, 6 and 12 months
	Accelerometer data **(Sub-study participants only**): - total energy expenditure - active energy expenditure - average metabolic equivalent of tasks (METs) - physical activity duration - step count - duration on body - lying down - sleep duration	Continuous	Baseline and 12 months

	Lifestyle factors	Variables	Type of variable	Timepoint(s) of evaluation of endpoint
B.	To compare changes in other **lifestyle factors** during the 12 months intervention	- Smoking status (current/ex-smoker/never) - Vaping status (current/ex-vaper/never) - Drinking status (current/ex-drinker/never) - Recreational substances status (yes/no)	Categorical	2, 6 and 12 months
C.	To compare changes in diet during the 12 month intervention	Food diaries: -Energy (in kcal or kJ) -Protein (g) -Fat (g) -Carbohydrate (g) -Fibre (g)	Continuous	2, 6 and 12 months

6.4. Safety Variables

Adverse events will be classified by "type" (serious/non-serious), severity (mild, moderate, severe), causality/relationship to study treatment (unrelated, unlikely, possibly, probably, definitely), expectedness (expected vs unexpected, SAEs only), outcome (resolved, resolved with sequelae, persisting, worsened, fatal) and MedDRA Structure (system organ class (SOC), high level group term (HLGT), preferred term (PT), lowest level term (LLT)).
According to the safety reporting manual of iPREVENT the causality is made by the investigator responsible for the care of the participant and the expected adverse events of the trial are gastrointestinal symptoms. All important abnormal laboratory test results occurring during the trial will be recorded as adverse events. Pregnancies at any point during the trial will also be recorded.

6.5. Demographic Variables

The demographic variables collected and to be summarized in Table 1 of reports are age, gender and ethnicity.

6.6. Mechanistic sub-sample variables

	Mechanistic sub-study objectives:	Variables	Type of variable	Timepoint(s) of evaluation of endpoint
a.	To compare changes in colonic metabolism to identify the relative abundance of the bacterial component of the microbiome	Gut microbiota abundance (stool sample - metataxonomic analysis of the 16S rRNA gene) Hydrogen breath test (ppm)	Continuous Continuous	12 months (changes from baseline)
b.	(i) To analyse the metabolite profile (nuclear magnetic resonance spectroscopic analysis) (ii) To determine how these specific changes in the colonic environment influence L-cell differentiation	Level of SCFA and metabolites profile (stool sample) (propionate, acetate and butyrate) Count of L-cells in intestinal organoid	Continuous Continuous	12 months (changes from baseline)

Mechanistic sub-study objectives:	Variables	Type of variable	Timepoint(s) of evaluation of endpoint
c. Appetite regulation:			
(i) To compare subjective feelings of appetite	-Visual Analogue Scales (VAS) of appetite (4 scales: hunger; desire to eat; fullness; nausea)	Continuous	12 months (changes from baseline)
(ii) To compare anorectic gut hormones	- PYY - GLP-1 - Gastrin - CCK	Continuous Continuous Continuous Continuous	
d. Mechanisms involved in weight maintenance: (i)To compare energy expenditure	- Energy expenditure (kcal/day) - Carbohydrate oxidation (g/min) - Fat oxidation (g/min) - RER (respiratory exchange ratio)	Continuous Continuous Continuous Continuous	12 months (changes from baseline)
(ii) To compare hepatic lipid metabolism	Stable isotope tracers of fat oxidation ($^{13}CO_2$ + ^{13}C-beta-hydroxybutyrate) -De Novo Lipogenesis (^2H_2O)	Continuous Continuous	
(iii) To compare Total Body Water	- ^2H_2O in body water	Continuous	

6.7. Compliance variables

This is of interest in itself but it is also involved in the supplementary investigation as mentioned above.

Secondary objective:	Variable	Type of variable	Timepoint(s) of evaluation of endpoint
To determine **compliance** during the 12-month intervention	Accountability of returned used for all subjects.	Continuous	2, 6 and 12 months

7. STATISTICAL METHODOLOGY

7.1. General Methodology

7.1.1 Data decisions made

In principle the data manager will make limited decisions about data variables and values so that issues such as missing data can be comprehensively handled by the trial statistician. Decisions which impact on the analysis will be recorded in the final clinical study report.

7.1.2 Outcomes requiring derivation

The International physical activity questionnaire (IPAQ) short form asks about three types of activity undertaken in four domains: a) leisure time physical activity, b) domestic and gardening activities, c) work-related physical activity and d) transport-related physical activity. It provides separate scores on walking,

moderate-intensity and vigorous-intensity activity which are summarized using median and interquartile range values and expressed in MET-minutes/week (17). According to the IPAQ guideline, "METs are multiples of the resting metabolic rate and a MET-minute is computed by multiplying the MET score of an activity by the minutes performed". A combined total physical activity score is also reported and calculated from the summation of the duration in minutes and frequency (days) of walking, moderate-intensity and vigorous-intensity activities. In the current study, the categorical variable of IPAQ will be used, which describe the physical activity in three levels: Low, Moderate and High. The data cleaning and processing (including missing value guidance) of IPAQ will be followed according to the guideline(17).

7.13 Use of data transformation

It is anticipated that some continuous outcomes, mainly in the mechanistic sub-study, will need to be considered for transformation, such as the accelerometer data or measurements expressed as percentages. Assumptions of normality and constant variance required by the models will be examined using residual and other diagnostic plots. If it is relevant, such as when there is too much skewness for the sample size such as triglycerides, a log transformation will be considered, because this retains a sensible interpretation for inferences; in relative terms between arms. If an absolute interpretation is needed, then data transformation may not be undertaken, but a nonparametric Bootstrap method for obtaining confidence intervals may be considered (18).

7.1.4 Defining and handling outliers

Outliers are observations that have extreme values relative to other observations observed under the same conditions. An outlier will be defined and handled in two ways:

i) For the purpose of checking data at interim and final analyses (for querying), defined as data-points being at least three standard deviations away from the mean of its distribution of values (either cross-sectional, change from baseline, or pairs of successive serial data-points ('bivariate outliers') observed across other patients. These definitions will apply to the transformed scale for those outcomes that have been log transformed. If many outliers are identified in this way, other thresholds (eg. 3.5SD or 4SD) will be used for querying only those that may be more highly influential.

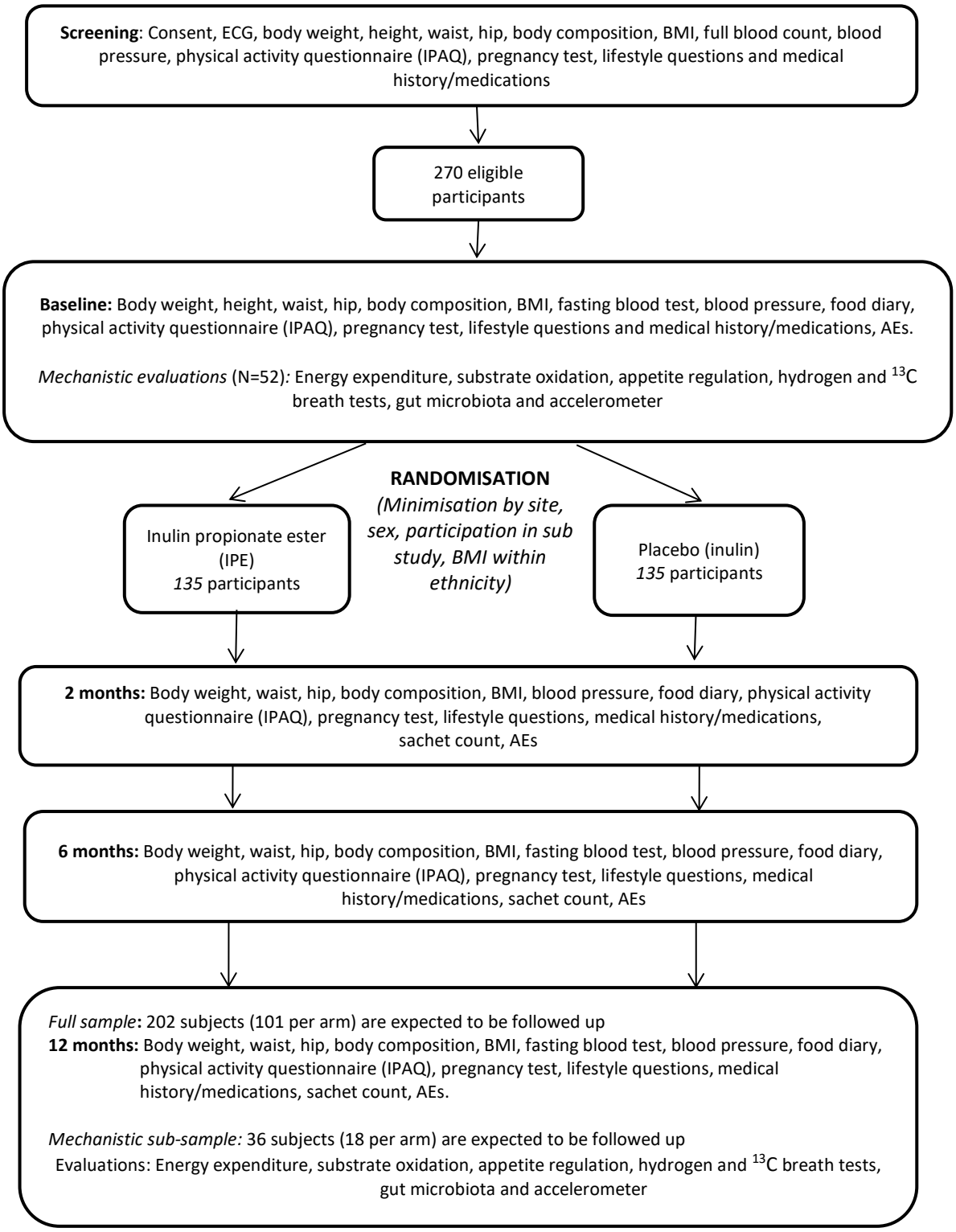
Outliers will be identified for further investigation by looking at the distributions of the data through histograms, scatter plots or box-plots. Univariate tests for the compatibility of the distribution with a normal distribution will not be undertaken since they can be too sensitive to departures that are often not relevant for the comparison of means (Central Limit Theorem).

Once an outlier is found, a blinded member of the team with enough clinical experience will be involved in the decisions as to whether a data value is impossible versus implausible versus plausible. If the outlier is impossible, then it will be set to missing. If an outlier is clinically plausible, the outlier will remain. If an outlier is in between these extremes, i.e. clinically implausible (but possible), it will not be ignored or deleted but will be retained for ITT analysis. If outliers remain in the distribution of a variable, then data transformations or nonparametric methods of analysis may be considered.

ii) For comparative analysis, outliers will be defined as observations with model residuals that are at least four standard deviations away from the mean of its distribution of values, considered 'potentially implausible'. In this situation, sensitivity analysis will be undertaken to check whether the outlier is influential by obtaining results primarily with and then without inclusion of the outlier. If the conclusions are changed, then this will be noted.

7.1.5 Flow diagram/recruitment

The flow diagram of the study is the one below. This will include the number assessed for eligibility, randomised, who comprise the intention to treat population, and the numbers followed-up to be in the analyses of the primary outcome as well as the main reasons for missing data by stages of the trial (CONSORT diagram(19)).


Screening: Consent, ECG, body weight, height, waist, hip, body composition, BMI, full blood count, blood pressure, physical activity questionnaire (IPAQ), pregnancy test, lifestyle questions and medical history/medications
270 eligible participants
Baseline: Body weight, height, waist, hip, body composition, BMI, fasting blood test, blood pressure, food diary, physical activity questionnaire (IPAQ), pregnancy test, lifestyle questions and medical history/medications, AEs.
Mechanistic evaluations (N=52)*:* Energy expenditure, substrate oxidation, appetite regulation, hydrogen and ^{13}C breath tests, gut microbiota and accelerometer
RANDOMISATION
(Minimisation by site, sex, participation in sub study, BMI within ethnicity)
Inulin propionate ester (IPE)
135 participants
Placebo (inulin)
135 participants
2 months: Body weight, waist, hip, body composition, BMI, blood pressure, food diary, physical activity questionnaire (IPAQ), pregnancy test, lifestyle questions, medical history/medications, sachet count, AEs
6 months: Body weight, waist, hip, body composition, BMI, fasting blood test, blood pressure, food diary, physical activity questionnaire (IPAQ), pregnancy test, lifestyle questions, medical history/medications, sachet count, AEs
Full sample: 202 subjects (101 per arm) are expected to be followed up
12 months: Body weight, waist, hip, body composition, BMI, fasting blood test, blood pressure, food diary, physical activity questionnaire (IPAQ), pregnancy test, lifestyle questions, medical history/medications, sachet count, AEs.
Mechanistic sub-sample: 36 subjects (18 per arm) are expected to be followed up
Evaluations: Energy expenditure, substrate oxidation, appetite regulation, hydrogen and ^{13}C breath tests, gut microbiota and accelerometer

7.1.6 Withdrawals/follow-up

Alongside the descriptions of the number/percentage of patients withdrawing from the study by arm in the CONSORT diagram as mentioned above, the number and percentage of patients withdrawing just from intervention (who continue data collection) will be summarized by treatment arm, reasons and timing. Baseline demographics, lifestyle factors and participants' weight will be described and compared between those withdrawing from the study and those completing the study using means (standard deviations (SD)) or number (percentages) and compared using difference in means or proportions with 95% confidence intervals.

7.1.7 Baseline demographics

Baseline descriptions of participants by treatment and overall will be summarised (into Table 1 of the report). No significance testing will be carried out as any differences found may be chance-generated and not for hypothesized reasons. This applies also for baseline in those in the mechanistic study, who consent to this prior to allocation to study arm.

7.2. Primary outcome analysis

7.2.1 Description

The primary outcome measure is the change from baseline in body weight at 12 months. As the analysis approach for continuous outcomes below makes advantage of linear covariate-adjustment for the baseline of the outcome, the primary endpoint can equivalently be regarded to be each participant's 12-month measurement. This is convenient because then those with a 12- month outcome, but whose baseline measurement is missing, are not regarded to be missing the endpoint. The primary outcome is therefore operationalised within the statistical model as the 12-month body weight, rather than the change in this from baseline to 12 months. It is not thought that this outcome will be missing at baseline, as it is involved in eligibility criteria and is a minimisation factor. This approach may be more useful for other outcomes.
The primary outcome will be analysed using a repeated measures ANCOVA via a linear mixed effects (LME) model incorporating the 3 post-baseline measurements of the body weight (2, 6 and 12 months). All outcome measurement from all randomised subjects who provide at least one post-baseline valid measurement will be included.

This mixed model will include fixed effect terms for arm (1 parameter), centre (1 parameter), sex (1 parameter), age (1 linear parameter), BMI within ethnicity (3 parameters), whether or not subjects volunteer to take part in the sub study (1 parameter), "*time*" (2 parameters), the baseline of the outcome (continuous body weight -1 parameter) and its missing indicator required for the missing indicator method (1 parameter , if needed) (20). Additionally, the model will include a fixed effect to indicate whether the primary outcome was measured by the participant at each timepoint. The other effects to be included in the model will be the interactions between "*time*" and each of the other fixed effects in the model (20 parameters). This model allows the treatment effect to be formally tested, at the primary timepoint of 12 months, and estimated at 2 and 6 months. Age is included because this was found to predict drop-out in a related trial (21).

Models will be fitted using a restricted maximum likelihood estimation using an unstructured variance-covariance matrix for the within-subject residuals (since there are only three follow-up timepoints, and so there is a manageable enough number of covariance parameters to expect this type of matrix to be supported by the sample size).

Model assumptions will be checked through the examination of the residuals and other diagnostic plots. Their distribution will be checked to see if there may be any evidence of large departures from a normal distribution, and to understand any influence of outliers (as explained above). It is highly unlikely that the distributional assumption would not be met for the purpose of estimating the parameters, as other studies of a similar sample size with primary outcome of body weight have reported analyses that did not require transformation or non-parametric approaches such as the Bootstrap (21).

In the event of data entry errors occurring in the randomisation (subjects entered in the wrong category of a minimisation factor), the corrected baseline category will be used in the analysis. Moreover, if other model assumptions do not hold such as model not converging, we will follow the steps written in section 8.2.

7.2.2 Primary estimand

Recently there has been a new guideline about strategies on how to deal with events occurring in the middle of the trial and how these relate to the treatment effect that reflects the clinical question posed by a clinical trial objective(10).

The following primary estimand, defined by its five attributes, is then considered:

Attributes of the estimand[1]:	
1 Population of interest:	Described in section 5.0 (see ITT trial sample)
2 Variable (endpoint) of interest:	Described in section 6.1
3 Treatment of interest:	Described in section 2.3
4 Intercurrent events[2]:	**Strategy[3] for addressing the intercurrent event**:
a) Death	Treatment policy
b) Pregnancy/ Treatment discontinuation due to pregnancy	Hypothetical strategy
c) Concomitant treatment: Eg. weight loss program /bariatric surgery/ gym personal trainer	Treatment policy†
d) Development of an illness, say cancer or COVID-19 or other disease/AE that affects a participants' weight	Treatment policy†
e) Treatment discontinuation due to an AE not affecting participants weight	Treatment policy
f) Treatment discontinuation due to lack of efficacy (increase in weight)	Treatment policy
g) Treatment discontinuation due to run out of supply	Treatment policy
h) Treatment discontinuation due to being indirectly affected by COVID-19 (preventing from coming to clinical visits)	Treatment policy
5 Population-level summary for the variable: Difference in weight gain between those taking IPE and control inulin after 12 months	

[1]Citing the ICH E9-R1, "*An estimand is a precise description of the treatment effect reflecting the clinical question posed by a given clinical trial objective"*(10).

[2]Intercurrent events are: "*Events occurring after treatment initiation that affect either the interpretation or the existence of the measurements associated with the clinical question of interest.(…) Unlike missing data, intercurrent events are not to be thought of as a drawback to be avoided in clinical trials*" (10).

[3]This is the specification of how to account for the intercurrent events that reflect the scientific question of interest. To note is that *'clarity in the estimand gives a basis for planning which data need to be collected and hence which data, when not collected, present a missing data problem to be addressed in the statistical analysis'*(10).

†This will be followed by a hypothetical strategy as sensitivity analysis.

To summarize, the treatment policy strategy (where the value of a subject's weight will used regardless of whether the intercurrent event occurred) will be used for all intercurrent events apart from pregnancy, where a hypothetical strategy will be considered (10). This means that we are not interested in the weight after a woman gets pregnant (since weight is evidently changed after pregnancy), but only in the 'hypothetical' weight had the subject not become pregnant. Therefore, the weight after a subject is known to be pregnant will not be considered for the ITT analysis, but we will make use instead of the linear mixed effects model which will indirectly impute the 'missing' data (see next section).

7.2.3 Missing data

An expert missing-data group concluded that rather than statisticians reacting to missing data at the end of a trial, there should be comprehensive, proactive planning for handling missing data at the stage of designing trials (22). The group recommended there should be consideration of missing data mechanisms (e.g Missing At Random), and, if the missing data may be informative that appropriate sensitivity analyses should be undertaken to investigate the robustness of the inferences to the different assumptions made by the main analysis.

It is anticipated that there may be approximately 25% missing data at the 12-month timepoint. The linear mixed effects model, above, will assume that the missing data is "missing at random" (MAR). This is an advantage over the assumption that the missing data is "missing completely at random" because it corrects for contribution to missing data from the covariates in the model and the correlated body weight outcome over time. The robustness to this assumption will be challenged using the sensitivity analyses. Multiple imputation is an alternative approach that also makes a MAR assumption, however we will use the likelihood approach of the linear mixed effects model, above, that gives similar resulting estimates(23).

7.2.4 Sensitivity analysis to missing data – 'Tipping point' analysis

A sensitivity analysis will be undertaken to assess the possibility of alternative plausible values of treatment effect arising from potential mishandling of missing data in the primary analysis model. The LME model for the primary outcome analysis described above is the first of a two-part approach called the Intention to Treat Strategy (16) in which a second analysis examines the sensitivity of the results to missing data in the full randomised, Intention to Treat, population. This meets the ideal of ITT to include everyone randomised. The approach to missing data follows the recently published implementation paper of the ITT strategy (23).

For the sensitivity analysis, we pre-specify a range for body weight from -5kg to +5kg over which the mean of the "unobserved outcome data" might depart (or be different) from the mean of the "observed outcome data" (23). In terms of a typical individual subject with unobserved (i.e. missing) data, this range can be thought of as the amount by which s/he may on average have had a different estimated treatment effect compared to the corresponding subject with the outcome data observed (given the same baseline covariates and follow-up data in the LME model). The range (-5 to +5) is chosen to represent both negative and positive departures that could potentially arise as the "net effect" of alternative reasons which may be unknown; such as dropout due to no anticipated further improvement, or dropout due to no improvement so far together with no anticipated achievable improvement. This range of 10Kg (from -5 to +5) is set generously wide for exploring sensitivity of the main results to departures from the MAR assumption, because 5kg (as the maximum departure in either direction) is larger than the detectable between-arm treatment effect of 2kg seen in superiority trials (difference in means) which is a sizeable shift in the mean of the distribution for dropouts compared to completers.
At the end of the trial, the fractions of individuals with missing data for body weight at 12 months will be available in each arm fi (for intervention) and fc (for control). The parameter representing excess weight in those with unobserved data compared to those with observed data, δ, will take values by passing across the range -5 to +5. Three scenarios will be undertaken within the sensitivity analysis (23, 24). These reflect whether departures from the MAR assumption apply within the intervention arm only (IPE), within the control arm only (inulin), or within both arms equally and in the same direction (thereby potentially cancelling out across the sensitivity range, if the dropout rate were to be the same in both arms).
Scenario 1: the treatment effect from the LME model will be increased by fiδ
Scenario 2: the treatment effect from the LME model will be increased by -fcδ
Scenario 3: the treatment effect from the LME model will be increased by (fi-fc$)\delta$

The number of kg within the range of -5kg to +5kg at which the results change in their statistical significance will be determined (if there is such a number), and this is known as the tipping point. This approach has been used in a recent trial (LEAVO) (25).

7.2.5 Sensitivity analysis for the primary estimand where a hypothetical strategy is adopted for the intercurrent events: AE or concomitant treatments affecting participants' weight

For participants who have a disease such as cancer, COVID-19 or another disease/AE that affects (directly or indirectly) a person's weight, or who starts a diet program/gym personal trainer or something similar, a hypothetical strategy could be an alternative, where we would just be interested in estimating the treatment effect of IPE had these events not happened. Thus, in this sensitivity analysis, measurements occurring after these events will instead be removed from the analysis and the LME model refitted (as done for pregnancy).

7.2.6 Subgroup analysis

It is not intended to do any subgroup analysis.

7.3. Clinical and anthropometric secondary outcome analysis

As for the primary outcome, the analysis of each continuous secondary outcome measured across the study timepoints will be compared between arms at 12-months using a linear mixed effects model adjusting for all minimisation factors, age (continuous) and, where collected, the baseline of the outcome and the associated missing indicator (if there is missing data in the baseline of outcome). *Time* will be represented as categorical contrasts in main effect form and in interaction with all other fixed effects.

To assess the arm effect on the study population 'during' the 12 months, the area under the curve of the estimated between-group intervention effects at 2, 6 and 12 months (where appropriate) will be estimated as well as the respective standard errors and covariances. This is possible because the intervention effect parameters from the linear mixed effects model have a population average interpretation as well as subject-specific. This will be supported by estimating cross-sectional differences between arms at these timepoints.

Where possible, study analyses will be used to provide estimates and two-sided 95% confidence intervals. For the secondary endpoints presented in section 3.9 of the protocol, the following methods of analyses are planned:

Objectives		Variables	Main summary statistic	Statistical method for comparison between arms and inference type
2.	To determine effects on **glucose homeostasis** as a surrogate marker of type-2 diabetes risk	Fasting biochemistry: - Glucose (mmol/L) - Insulin (mIU/L)	Means(SE)	LME incorporating the 2, 6 and 12-month timepoints described with 95% Confidence intervals.
3.	To determine effects on **blood lipid and cholesterol** as surrogate markers of CVD risk	Fasting biochemistry: - Triglycerides (mmol/L) - Total cholesterol (mmol/L) - LDL cholesterol (mmol/L) - HDL cholesterol (mmol/L)	Means(SE)	LME incorporating the 6 and 12-month timepoints described with 95% Confidence intervals.
4.	To determine effects on **blood pressure** as a surrogate marker of CVD and stroke risk	Blood pressure: - SBP (mm/Hg) - DBP (mm/Hg)	Means(SE)	LME incorporating the 2, 6 and 12-month timepoints described with 95% Confidence intervals.

	Objectives	Variables	Main summary statistic	Statistical method for comparison between arms and inference type
5.	To compare changes in **body weight/ waist/ BMI/body composition** during the 12 months intervention	- Body weight - Waist/hip - BMI (kg/m^2) Body composition: - Fat mass (FM) (kg) - Fat Mass Index (FMI) (kg/m^2) - Percent body fat (Fat%) - Fat free body mass (FFM) (kg) - FM/FFM ratio	Means(SE)	LME incorporating the 2, 6 and 12-month timepoints described with 95% Confidence intervals.

7.4. Lifestyle factors analysis

Lifestyle factors will be summarized over time using simple descriptive statistics.

Objectives	Variables	Main summary statistic	Statistical method for comparison between arms and inference type†
A. To compare changes in **physical activity** during the 12 months intervention	IPAQ (ordered) categorical score (all patients) -Low -Moderate -High	Unadjusted patient proportions at 2, 6 and 12months and change from baseline Cross-tabulations of the ordered categorical IPAQ with baseline	Linear-by-linear association test at 12 months
	Accelerometer data (for sub study participants only): - total energy expenditure (KJ) - active energy expenditure (KJ) - average metabolic equivalent of tasks (METs) - physical activity duration - step count - duration on body - lying down - sleep duration	Means (SD) / Medians (IQR) Mean and SD of change in accelerometry from baseline	95% CI for difference analogous to a T-test
B. To compare changes in other **lifestyle factors** during the 12 months intervention	- Smoking status (current/ex-smoker/never) - Vaping status (current/ex-vaper/never) - Drinking status (current/ex-drinker/never) - Recreational substances status (yes/no)	Unadjusted patient proportions at 2, 6 and 12months and change from baseline Cross-tabulations with baseline	95% CI for difference in current status analogous to a chi-squared test
C. To compare changes in **diet** during the 12 months intervention	Food diaries: -Energy (in kcal or kJ) -Protein (g) -Fat (g) -Carbohydrate (g) -Fibre (g)	Means(SD)/ Medians(IQR) Mean and SD of change from baseline	95% CI for difference analogous to a T-test

†Differences in these lifestyle factors between arms would lead to sensitivity analyses.

7.5. Safety Analysis

Adverse events and SAEs will be reported in terms of the number of participants with at least one event per arm and using descriptive statistics, as unadjusted patient proportions within and between arms with 95% confidence intervals using exact methods where appropriate, such as the Wilson method (26) (with none versus with at least one event). Numerators and denominators will be presented. For the most prominent gastrointestinal disorders (in at least 10% of participants in either treatment group), as well as AEs that lead to treatment discontinuation, details of the overall number of events, timing to reporting of first such event by participant (median (IQR) time since randomisation), severity, and duration will be described as per preferred term. Denominators will generally be participants over the trial period, although between-visit stratification will be considered, as will a person-years denominator depending on the variability in follow-up. *P*-values will not be used taking account of guidance given in the CONSORT statement extension for harms (27, 28). Volcano (space-saving graphical summary) or dot plots may also be presented if frequent adverse events occur(29).

7.6. Compliance Analysis

Objective		Variable	Main summary statistic	Statistical method for comparison between arms and inference type
7	To determine **compliance** during the 12-month intervention	Firstly, by accountability of returned used sachets. Secondly, communicated verbally by the participant (such during COVID periods where visits were done remotely) Percentage compliance for a subject will then be defined as number of used sachets divided by the total number of sachets expected to have been taken over the 12-month study period)*100%.† In periods where the sachets have not been returned it will be assumed that the compliance was the same as in the average of the other periods (0-2months, 2 -6 months, 6 -12months). For participants reporting treatment discontinuation or that were lost to follow-up, we will consider zero used sachets from these timepoints onwards. Those with no treatment compliance data will be regarded as below 50% compliance for purpose of analysis.	Proportions of 50% to 95% compliance (in steps of 5%) Medians (IQR)	Difference in proportions with 95% CIs Comparison of compliance distributions Difference in medians

†If return of sachets is low (median of used and unused <50% of those dispensed) we will check for the correlation between this objective measurement of compliance (accountability of unused sachets) and the self-reported measurement 'How many times did subject miss a sachet since the previous study visit?' and if it turns to be high (≥0.7), and this self-reported measurement has less missing data, than the self-reported measure will be used to assess compliance. For subjects who become pregnant, only the time up to this event will count when assessing compliance.

7.7. Supplementary analysis: The effect of IPE in a more compliant population

Attributes of the supplementary analysis estimand:	
1 Population of interest:	Defined by those participants eligible to be in the trial *
2 Variable (endpoint) of interest:	Already described in section
3 Treatment of interest:	Already described in section 2.3
4 Intercurrent events*:	**Strategy to deal with the intercurrent event**:
a) Death	Treatment policy
b) Pregnancy/ Treatment discontinuation due to pregnancy	Hypothetical strategy
c) Concomitant 'treatment': Eg. weight loss program /bariatric surgery/ gym personal trainer	Hypothetical strategy
d) AE/Development of an illness, say cancer or COVID-19 or other disease that affects participants' weight	Hypothetical strategy
e) Treatment discontinuation due to an AE not affecting participants weight	Treatment policy
f) Treatment discontinuation due to lack of efficacy (increase in weight	Treatment policy
g) Treatment discontinuation due to run out of supply	Treatment policy
h) Treatment discontinuation due to being indirectly affected by COVID-19 (preventing from coming to clinical visits)	Treatment policy
5 Population-level summary for the variable: Difference in weight gain between those taking IPE and control inulin after 12 months	

*Just to emphasise this population is different from that addressed in section 7.2.2

In this analysis, the strategies to address the intercurrent events are the same as the ones described in section 7.2.2, except for AE or concomitant 'treatments' affecting a participant's weight, for which the hypothetical scenario is envisage. Thus, as before, measurements occurring after these events will be removed from the analysis and the LME model will be refitted.

7.7.1 Sensitivity analysis for non-compliance

White et al. also recommend that analyses allowing for non-response and low intervention uptake (or compliance) are specified in advance in the analysis plan (24). "*Incomplete uptake of trial interventions often means that randomised groups have more similar experience than the investigators had intended, which usually causes the difference in outcomes to be smaller than it would have been with better uptake*." (24). This means that low compliance with the IPE may bring about two groups that have more similar intervention experience to each other than was planned (i.e. intervention more similar to control through low compliance).

Further to the previous sensitivity analysis we will conduct an analysis estimating the effect of IPE versus Inulin control on the primary outcome in a more highly compliant population, whilst respecting randomisation. This approach should provide a better estimate of the true effect without suffering from potential biases seen in a per-protocol analysis.

Therefore a complier average causal effect (CACE) analysis will be carried out as recommended and outlined by Dunn et al. (30). The Complier-Average Causal Effect (CACE) estimate is the comparison of the average outcome of the compliers in the IPE arm with the average outcome of the comparable group ("would-be compliers") in the Inulin Control arm. Since it is not well studied as to the number of sachets needed to be taken over a 12 month period for the IPE to have an effect, if there is one, alternative levels of compliance will considered, ranging from 50% to 95%, and is therefore more exploratory than the planned primary analysis. The outline of the approach to be taken is given here:

Sample sizes (N) and means (M) are deduced for the Inulin control "would-be compliers" and "would-be-noncompliers" in the following table by assuming that the proportion of intervention group compliers, and control group would-be compliers, is the same under randomisation, and that would-be non-compliers in the control group would have the same mean outcome as non-compliers in the intervention group (the exclusion

restriction assumption). The sample sizes refer to those followed-up with primary outcome data (body weight at 12 months).

Arm	Compliers (≥ chosen threshold %)	Noncompliers (< chosen threshold %)	All
IPE	N_{I1} M_{I1}	N_{I2} M_{I2}	N_I M_I
Inulin Control	$= N_C - (N_{I2}/N_I)*N_C$ $= (M_C-(N_{I2}/N_I)*M_{I2})/(N_{I1}/N_I)$	$= (N_{I2}/N_I)*N_C$ $= M_{I2}$	N_C M_C

(Statistics preceded by an "=" are unobserved, and are estimated from the observed statistics.)
The method is adapted to a more plausible MAR assumption by replacing the sample sizes at follow-up by those at baseline. In the presence of missing compliance, it will be primarily assumed that the participant is a non-complier.

The CACE estimate will instead be obtained from the analysis LME model applying the strategies in table 7.7, as this adjusts for baseline and refers to the population for the supplementary investigation. The CACE estimate is the ratio of the estimated treatment effect to the proportion compliant, following the rule of thumb (31) (estimate LME/ proportion compliant).

7.8. Mechanistic sub-study analysis

Extending the table in section 3.9 of the protocol, the planned analysis for the mechanistic sub-study endpoints are in table below. Since the distribution of all the variables is not clearly defined *a priori*, and the size of the subsample is relatively small, the statistical methods used at final analysis may not be exactly the same as the ones described. For example, a log-transformation may be needed, or a non-parametric method of analysis (such as Mann-Whitney test of the change score, or of the final score if less variable). Bootstrap 95% confidence intervals may also need to be computed.

Mechanistic sub-study objectives:	**Variables**	**Main summary statistic**	**Statistical method for comparison between arms and inference type**
a. To compare changes in colonic metabolism to identify the relative abundance of the bacterial component of the microbiome	-Gut microbiota abundance (stool sample) -Hydrogen breath test (ppm)	Means(SD)/Medians(IQR) of the AUC and/or the change over time in the AUC Means(SD)/Medians(IQR) of the AUC and/or the change over time in the AUC	ANCOVA with 95% CI ANCOVA with 95% CI
b. (i) To analyse the metabolite profile (ii) To determine how these specific changes in the colonic environment influence L-cell differentiation	Level of SCFA and metabolites profile (stool sample) Proliferation of L-cells in intestinal organoid	Means(SD)/ Medians(IQR) Means(SD)/ Medians(IQR)	ANCOVA, with 95% CI ANCOVA, with 95% CI

c. Appetite regulation: (i) To compare subjective feelings of appetite	-Visual Analogue Scales (VAS) of appetite (4 scales: hunger; desire to eat; fullness; nausea)	Means(SD) or Medians(IQR) of the AUC or the change over time in the AUC	ANCOVA, with 95% CI
(ii) To compare anorectic gut hormones	- PYY - GLP-1 - Gastrin - CCK	Means(SD) or Medians(IQR) of the AUC and/or of the change over time in the AUC	ANCOVA, with 95% CI
d. Mechanisms involved in weight maintenance:			
(i) To compare energy expenditure	- Energy expenditure (kcal/day) - Carbohydrate oxidation (g/min) - Fat oxidation (g/min) - RER (respiratory exchange ratio)	Means(SD) or Medians(IQR) of the AUC and/or of the change over time in the AUC	ANCOVA with 95% CI
(ii) To compare hepatic lipid metabolism	-Stable isotope tracers of fat oxidation (13CO2 + 13C-beta-hydroxybutyrate) -De Novo Lipogenesis (^2H_2O)	Means(SD)/ Medians(IQR)	ANCOVA with 95% CI
(iii) To compare Total Body Water	-^2H_2O in body water	Means(SD)/ Medians(IQR)	ANCOVA with 95% CI

7.9. Interim Analysis

As described in the DMEC charter, formal interim analysis of the primary outcome for early stopping is not planned for this study. Regular interim reports will be prepared for DMEC meetings every 6 to 12 months. There will be a stop-go decision on the recruitment which will not involve any interim analysis of the primary outcome (described in study protocol).

7.10. Handling multiple comparisons

Significance tests will be used sparingly and restricted where possible to addressing stated hypotheses. Secondary outcomes, as well as the primary outcome, will be summarised using an effect size with a 95% confidence interval. Interpretation for those secondary outcomes that do not directly address the stated study hypotheses will be more cautious. Where multiple outcomes address an objective, the interpretation will account for the role of the outcome within the objective and the strength and direction of effects in all outcomes.

7.11. Handling protocol deviations

The number (percentage) and type (major/minor) of protocol deviations with potential to affect the primary outcome of the trial will be summarized by treatment group. No formal statistical test comparing the groups will be undertaken as it is not expected for this to happen in many patients. Nevertheless, if there are at least 5% of study subjects that were unblinded during the study or were known to have taken part in a weight loss program or consumed a weight loss product then a sensitivity analysis removing these subjects from the primary statistical model will be carried out to assess robustness of the results.

7.12. Software

The principal statistical software package will be IBM SPSS Statistics 25, and R software will be available if required.

8. IMPACT OF COVID-19

8.1. On data quality and usage of data

The possibility of subjects having COVID-19 was already addressed in tables 7.2.2 and 7.7.

As written in the iPREVENT guidance document of COVID-19 restrictions, all in-person clinic study visits have been suspended until further notice, with follow up visits being conducted only remotely. This means that body weight (primary outcome) will be self-reported during the restriction time.

When subjects return to clinic study visits, these will be asked to measure their weight before going into the clinic so that the self-reported weight can be compared with the weight measured at the clinic using the Bland and Altman limits of agreement and correlations. Since the self-reported weight was collected at the pre-screening questionnaire it is possible to measure change in measured weight and self-reported weight. These will be discussed with independent committees.

Since it is hoped that all subjects reach the 12-month visit in a non-restricted/covid-19 period, only the weight measured at the clinic will be considered for the primary outcome of the trial.

8.2. On model assumptions not being met

In the eventuality of low recruitment, or the trial having to stop early due to COVID pandemic or high dropout, the sample size may be smaller than expected. Also, the model may be over-parameterised. In these situations, the model may not converge. In this eventuality, the reason for the non-convergence will be explored. It may be due to the covariance structure and/or too many fixed-effect parameters. The non-convergence may disappear after simplifying the covariance structure (to being distinct variances per timepoint but a common correlation) and/or by removing the interactions of time with all fixed effects except for arm and baseline (and missing indicator if needed) – the "reduced set" of covariates. If only one of these two actions were necessary, then we would prefer to retain the complex covariance structure. If neither action removed the non-convergence, ANCOVA models would be fitted for each timepoint, involving ideally all planned fixed-effects, or otherwise the "reduced set". The same model was fitted successfully in a previous trial that had more arms and timepoints, and similar sample size per arm (25). If any these situations occur, we will check which of these several options are feasible just before database lock.

9. AMENDMENTS TO SAP VERSIONS

A list of amendments to SAP version 1.0 will be listed here. The list will be cumulative and identify the changes from the preceding document versions.

Protocol version	Updated SAP version no.	Section number changed	Description of and reason for change	Date changed
5.0	2.0	7.2.1	The following sentence was added: *Additionally, the model will include a fixed effect to indicate whether the primary outcome was measured by the participant at each timepoint*. This was recommended by the DMC Chair/Statistician to accommodate the way that weight is measured.	07/01/2022
5.0	2.0	7.6	Sachets accountability was clarified as well as compliance definition (where sachets have not been returned). This was recommended by the DMC Chair/Statistician as a more realistic assumption for missing data rather then assuming conservatively low compliance of zero.	13/01/2022

10. REFERENCES

1. (MHRA) MaHpRA. Good Clinical Practice Guide: TSO (The Stationery Office).
2. Group IEW. ICH Harmonised tripartite guideline-Statistical principles for clinical trials-E9. 1998. Report.
3. Guideline IHT. Structure and Content of Clinical Study Reports E3. Current Step. 1995;4:24.
4. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
5. Committee for Proprietary Medicinal P. Committee for Proprietary Medicinal Products (CPMP): points to consider on adjustment for baseline covariates. Stat Med. 2004;23(5):701-9.
6. Obesity: we need to move beyond sugar. Lancet. 2016;387(10015):199.
7. Fildes A, Charlton J, Rudisill C, Littlejohns P, Prevost AT, Gulliford MC. Probability of an obese person attaining normal body weight: cohort study using electronic health records. Am J Public Health. 2015;105(9):e54-e9.
8. Hebden L, Chey T, Allman-Farinelli M. Lifestyle intervention for preventing weight gain in young adults: a systematic review and meta-analysis of RCTs. Obes Rev. 2012;13(8):692-710.
9. Chambers ES, Viardot A, Psichas A, Morrison DJ, Murphy KG, Zac-Varghese SE, et al. Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults. Gut. 2015;64(11):1744-54.
10. Committee for Proprietary Medicinal Products (CPMP). ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. 2019
11. Burton PR, Gurrin LC, Campbell MJ. Clinical significance not statistical significance: a simple Bayesian alternative to p values. J Epidemiol Community Health. 1998;52(5):318-23.
12. Allman-Farinelli M, Partridge SR, McGeechan K, Balestracci K, Hebden L, Wong A, et al. A Mobile Health Lifestyle Program for Prevention of Weight Gain in Young Adults (TXT2BFiT): Nine-Month Outcomes of a Randomized Controlled Trial. JMIR Mhealth Uhealth. 2016;4(2):e78.
13. Polyviou T, MacDougall K, Chambers ES, Viardot A, Psichas A, Jawaid S, et al. Randomised clinical study: inulin short-chain fatty acid esters for targeted delivery of short-chain fatty acids to the human colon. Aliment Pharmacol Ther. 2016;44(7):662-72.
14. Chambers ES, Viardot A, Psichas A, Morrison DJ, Murphy KG, Zac-Varghese SE, et al. Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults. Gut. 2015;64(11):1744-54.
15. Anjum A. iPREVENT Data Management Plan. Unit ICT.
16. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ. 2011;342:d40.
17. Committee IR. Guidelines for the data processing and analysis of the International Physical Activity Questionnaire. 2005. 2016.
18. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat Med. 2000;19(9):1141-64.
19. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Med. 2010;8(1):18.
20. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med. 2005;24(7):993-1007.
21. Astbury NM, Aveyard P, Nickless A, Hood K, Corfield K, Lowe R, et al. Doctor Referral of Overweight People to Low Energy total diet replacement Treatment (DROPLET): pragmatic randomised controlled trial. BMJ. 2018;362:k3760.
22. Burzykowski T, Carpenter J, Coens C, Evans D, France L, Kenward M, et al. Missing data: discussion points from the PSI missing data expert group. Pharm Stat. 2010;9(4):288-97.
23. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clinical trials. 2012;9(4):396-407.

24. White IR, Kalaitzaki E, Thompson SG. Allowing for missing outcome data and incomplete uptake of randomised interventions, with application to an Internet-based alcohol trial. Stat Med. 2011;30(27):3192-207.
25. Hykin P, Prevost AT, Vasconcelos JC, Murphy C, Kelly J, Ramu J, et al. Clinical Effectiveness of Intravitreal Therapy With Ranibizumab vs Aflibercept vs Bevacizumab for Macular Edema Secondary to Central Retinal Vein Occlusion: A Randomized Clinical Trial. JAMA Ophthalmology. 2019.
26. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. Statistics in medicine. 1998;17(8):873-90.
27. Ioannidis JP, Evans SJ, Gøtzsche PC, O'neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Annals of internal medicine. 2004 Nov 16;141(10):781-8.
28. Lineberry N, Berlin JA, Mansi B, Glasser S, Berkwits M, Klem C, Bhattacharya A, Citrome L, Enck R, Fletcher J, Haller D. Recommendations to improve adverse event reporting in clinical trial publications: a joint pharmaceutical industry/journal editor perspective. BMJ. 2016 Oct 3;355.
29. Zink RC, Wolfinger RD, Mann G. Summarizing the incidence of adverse events using volcano plots and time intervals. Clin Trials. 2013;10(3):398-406.
30. Dunn G, Maracy M, Dowrick C, Ayuso-Mateos JL, Dalgard OS, Page H, et al. Estimating psychological treatment effects from a randomised controlled trial with both non-compliance and loss to follow-up. The British Journal of Psychiatry. 2003;183(4):323-31.
31. White IR. Uses and limitations of randomization-based efficacy estimators. Stat Methods Med Res. 2005;14(4):327-47.

11. SIGNATURE PAGES

SIGNATURE PAGE 1 (Study Statistician)

The signature of the below constitutes approval of this SAP.

Study Title: Increase in colonic PRopionate as a method of prEVENTing weight gain in adults aged 20-40 years: iPREVENT

SAP version: 2.0

Signed: Joana Vasconcelos

Print Name: JOANA VASCONCELOS

Date: 13/01/2022

SIGNATURE PAGE 2 (Senior Statistician)

The signature of the below constitutes approval of this SAP.

Study Title: Increase in colonic PRopionate as a method of prEVENTing weight gain in adults aged 20-40 years: iPREVENT

SAP version: 2.0

Signed: A.T. Prevost

Print Name: A.T. PREVOST

Date: 25/1/22

SIGNATURE PAGE 3 (Chief Investigator - CI)

The signature of the below constitutes approval of this SAP.

Study Title: Increase in colonic PRopionate as a method of prEVENTing weight gain in adults aged 20-40 years: iPREVENT

SAP version: 2.0

Signed: 

Print Name: Gary Frost

Date: 28-02-2022

SIGNATURE PAGE 4 (Chair of Trial Steering Committee)

The signature of the below constitutes approval of this SAP.

Study Title: Increase in colonic PRopionate as a method of prEVENTing weight gain in adults aged 20-40 years: iPREVENT

SAP version: 2.0

Signed: 

Print Name: NICHOLAS FINER

Date: 25 JAN 2022

APPENDIX: Record of data decisions during the blinded review

Record of data and/or analysis decisions (such as arising from distributions of variables)

From the data prepared for the first DMC meeting on 7th January 2020, in which data from the first 35 trial participants was available, it was clear that triglycerides were markedly skewed and therefore a log transformation may be appropriate. This has been added to section 7.1 'Use of data transformation'.