# Personalised REsponses to DIetary Composition Trial (PREDICT): an intervention study to determine inter-individual differences in postprandial response to foods

**Sarah Berry** ( sarah.e.berry@kcl.ac.uk )
Kings College London https://orcid.org/0000-0002-5819-5109
**David Drew** ( DADREW@mgh.harvard.edu )
Harvard Medical School https://orcid.org/0000-0002-8813-0816
**Inbar Linenberg**
Zoe Global Ltd
**Jonathan Wolf**
Zoe Global Ltd
**George Hadjigeorgiou**
Zoe Global Ltd
**Richard Davies**
Zoe Global Ltd
**Haya Al Khatib**
King's College London, Zoe Global Ltd
**Deborah Hart**
King's College London
**Gabriella Surdulescu**
King's College London
**Darioush Yarand**
King's College London
**Ayrun Nessa**
King's College London
**Alyce Sheedy**
King's College London
**Amrita Vijay**
Nottingham University
**Francesco Asnicar**
University of Trento
**Nicola Segata**
University of Trento

**Andrew Chan**
Harvard University
**Paul Franks**
Lund University
**Ana Valdes**
University of Nottingham
**Tim Spector**
King's College London

---

**Method Article**

**Keywords:** Postprandial metabolism, Lipemia, Glycemia, Machine learning, Personalized nutrition, Microbiome

**Posted Date:** January 14th, 2020

**DOI:** https://doi.org/10.21203/rs.2.20798/v1



---

# Abstract

The Personalised REsponses to DIetary Composition Trial (PREDICT) was a single-arm, single-blind intervention study that utilized both standardized “test” meals and captured “free-living” non-standardized food consumption, to predict an individual’s metabolic response to foods based on the person’s characteristics (including their metabolomic and gut microbiome profiles), meal composition (macronutrients and energy content) and meal context (including time of day, sleep and exercise). The initial study commenced in June 2018 and targeted 1,000 generally healthy participants between the ages of 18 and 65 years for enrollment in the United Kingdom, and recruited from the existing and ongoing TwinsUK research cohort and the general population. In December 2018, an independent trial mirroring the UK protocol was launched in parallel at Massachusetts General Hospital to serve as a US-based validation cohort including 100 healthy individuals in the general population. The overall objective of these trials was to predict glucose, insulin, lipid and other postprandial responses to foods based on the individual’s characteristics, including molecular biomarkers and lifestyle factors as well as the nutritional composition of the food. The studies, collectively termed PREDICT 1, were completed in May 2019. The study has delivered highly-dimensional multi-omic data from which the predictors of individual postprandial responses have been determined. These findings will be used to design personalized dietary interventions focused on key health outcomes. This protocol is registered with the US National Institutes of Health trial registry, ClinicalTrials.gov: NCT03479866.


# Introduction

Effective prevention strategies are required to reduce the immense global burden of nutrition-related non-communicable diseases (NCD)[1]. The population-wide nutritional recommendations intended to guide our diet decisions are typically based on epidemiological and clinical trial data, using average nutritional responses for populations. However, it is increasingly evident that one-size nutritional recommendations do not fit all. Moving beyond this to a precision nutrition approach would require appropriately designed large-scale studies from which multi-dimensional, high-resolution phenotypic datasets could be analysed to assess individual responses to food[2]. New technologies offer the promise of accurately measuring a wide variety of postprandial (non-fasting) responses in many people in real-world settings, allowing the numerous factors shaping individual nutritional responses to be identified, quantified and understood[2].

Although fasting blood assays are the standard for many clinical diagnoses, people spend the majority of their waking hours in the postprandial state. Postprandial lipid, glucose and insulin dyshomeostasis are independent risk factors for NCDs and obesity[3,4,5], attributable in part to oxidative stress, inflammation, beta-cell failure, endothelial damage and lipoprotein remodelling caused by repeated elevated or prolonged blood glucose and lipid excursions. While postprandial glycemic responses are important health determinants, glycemic control is just one part of a more complex metabolic equation involving triglycerides (TG) (from which fatty acids are derived, the primary alternative energy substrate to glucose) and insulin (a key hormone regulating glucose and TG transport

and metabolism)[6]. Thus, also characterizing postprandial regulation of lipids and identifying the factors responsible for individual variations could help to optimize diet recommendations targeting broader improvements in cardiometabolic health.

To address these challenges, we undertook the PREDICT 1 study (NCT03479866) to quantify and predict individual variations in postprandial TG, glucose and insulin responses to standardized meals and free-living foods. We integrated data from a cohort of twins and unrelated adults from the UK (n=1,000) to interrogate genetic, metabolic, microbiome composition, meal composition and meal context data to distinguish predictors of individual responses to meals; we then validated these predictions in an independent cohort of adults from the US (n=100). Within we describe the detailed clinical trial protocol for the PREDICT 1 study as carried out in the UK. The intervention study was identically repeated in the US, however, any site-specific differences are noted.

The full protocol can be found in Supplementary files 'PREDICT 1 Protocol'.

# Reagents

# Equipment

# Procedure

### 1. Trial design

The study was a single-arm intervention consisting of test meal challenges, including one clinic baseline visit (at the Clinical Research Facility, St Thomas's Hospital, London, UK or the Translational and Clinical Research Center, MGH, Boston, MA, US), followed by a 13-day home-based study (Days 2-14). A summary of the study design appears in **Figure 1**.

### 2. Participants and interventions

*Study setting*

The PREDICT 1 study was a two-country study conducted between June 2018 and May 2019, with the primary cohort based at St. Thomas' Hospital in London, UK and a validation cohort (that underwent the same profiling as in the UK) assessed at Massachusetts General Hospital (MGH) in Boston, Massachusetts.

*Eligibility Criteria*

Potential participants were identified within the TwinsUK cohort and general population in the UK, and the general population in the US. Eligible adults, aged ≥18 years and <65 years, who were generally healthy were recruited to the study. Full inclusion and exclusion criteria are included in **Table 1**.

*Interventions and participant timeline*

A timeline summary is provided in **Table 2**.

*Pre-enrollment procedures*

Participants were mailed a pre-visit study pack with a stool collection kit and a health and lifestyle questionnaire (amended Twins Research health and lifestyle questionnaire[7] and food frequency questionnaire (European Prospective Investigation into Cancer and Nutrition (EPIC) Food-Frequency Questionnaire (FFQ)). In the US, minor modifications were made to the health and lifestyle questionnaires to conform to a US population and the Harvard semi-quantitative FFQ, a validated US dietary assessment instrument, was substituted for the EPIC FFQ.

The day before the visit, participants were asked to refrain from taking part in any strenuous exercise and to limit fat, fiber and alcohol intake, and to abstain from caffeine from 6pm. They were instructed to avoid eating or drinking anything except still water from 9pm the night before their visit; they were encouraged to drink plenty of still water to help with cannulation on the day of the visit. They were also asked to avoid taking any over-the-counter medication on the day of their appointment.

*Baseline clinic visit (Day 1)*

Stool collection and questionnaires were completed at home and subsequently returned to study staff at the baseline visit. Those who could not collect a stool sample prior to their baseline visit completed the collection as soon as possible during the home-phase. Participants arrived at 8:30am for their visit and their written consent for participation was taken, after which they were cannulated in the forearm (antecubital vein) to collect a fasted blood sample before being fitted with wearable devices (continuous glucose monitor (CGM; Freestyle Libre Pro, Abbott, Abbott Park, IL, US) and wrist-based triaxial accelerometer (AX3, Axivity, Newcastle, UK)). Heart rate and blood pressure were measured using an automated blood pressure monitor while fasted (in triplicate, with mean of second and third measurements recorded). In a subgroup (n=456), ambulatory blood pressure was also monitored throughout the baseline visit (TM-2430, A&D Instruments Limited, Abingdon, UK). Participant weight,

height, hip and waist circumference were measured using standard clinical techniques. Participants underwent a dual-energy, x-ray absorptiometry (DEXA) scan using a QDR Horizon W. Hologic Bone Densitometer (Vertec Scientific Ltd., Reading, UK). Fasting blood glucose level was checked using HemoCue Glucose 201 + System (Radiometer, Crawley, UK) and Stat Strip (Nova Biomedical, Waltham, MA, US) in the UK and US, respectively.

Following the baseline blood draw and finger prick blood spot card sampling, participants consumed a breakfast consisting of a Metabolic Challenge Meal (Meal 1), consisting of two high fat muffins (with blue food coloring for participants to provide an estimate of transit time by reporting the time of the first bowel movement where blue color was observed) and a milkshake. This meal had a total nutrient profile of 890 kcal, 86 g of carbohydrates, 53 g of fat, 16 g of protein, and 2 g of fibre. Participants also consumed a Medium Fat & Carb lunch meal (Meal 2) exactly 4 hours after the breakfast (at 240 min) consisting of three muffins (delivering 500 kcal, 71 g of carbohydrates, 22 g of fat, 10 g of protein and 2 g of fibre). Participants were instructed to consume the entire meal within 10 minutes from the first bite. Meal response data was included in analysis for all participants, including those consuming the meal in more than 10 minutes.

Additional venous blood was collected via cannula at 15, 30, 60, 120, 180, 240, 270, 300 and 360 minutes following the breakfast meal. Participants were allowed to sip water at regular intervals throughout the visit. Finger-prick dried blood spots (DBS) were collected on cards before and 5 hours after the breakfast, to provide a clinic-based comparison to the at-home DBS tests. Between blood sampling, participants were trained in how to complete the home-phase of the study, including when and how to consume standardised test meals, perform DBS tests, and use the Zoe study app (see below). Upon completing their baseline visit, participants received all the components necessary to complete the home-phase as a take-home pack. The take-home pack contained: 1) Standardised meals for the home-phase; 2) DBS equipment and return envelopes (collection equipment plus 4x aluminium bags and one master pre-paid return envelope for returning these bags); 3) Stool sample collection kit; 4) Pre-paid return kit (box or envelope with 2 separate packs for stool and equipment) with postage-paid packs to return the activity monitor, the CGM and the stool sample.

*Home-phase (Days 2-14)*

During the study home-phase, participants consumed multiple standardised test meals over a 9-11 day period, differing in macronutrient composition (carbohydrate, fat, protein and fiber), outlined in **Table 3,** while wearing the CGM and activity monitor. Participants recorded all of their dietary intake, satiety and exercise on the Zoe study app as outlined below throughout the study. DBS tests were completed on Days

2, 3, 4, and 5 before and after test meals. Participants also provided a stool sample upon completion of the study on Day 14. Following completion of the home-phase, participants returned all study samples and devices to study staff via standard mail.

Standardized test meals

Upon completing their clinical visit, participants received a pack of standardized test meals to take home for consumption, and were instructed on how to store and consume the meals. The study had a total of 3 test meal protocol groups (described in **Table 4**). Standard test meals were consumed at breakfast during the first 9-11 days (Days 2-12) of the home period, and additionally for lunch on Days 2 and 3. The at-home meals were:

**Meal 1. Metabolic Challenge Meal:** High fat, low carbohydrate muffins and milkshake (as was used for clinical visit breakfast)

**Meal 2. Medium Fat & Carb:** average level of fat and carbohydrate muffins

**Meal 3. High Fat 1:** high fat, low carbohydrate muffins

**Meal 4. High Carb:** high carbohydrate, low fat muffins

**Meal 5. OGTT**: oral glucose tolerance test (liquid high carbohydrate drink)

**Meal 6. High Fiber:** high fiber meal consisting of two fiber bars and muffins

**Meal 7. High Fat 2:** high fat, low carbohydrate muffins

**Meal 8. High Protein:** high protein meal consisting of a milkshake and muffins.

Participants received clear instructions (verbal and written) on how and when the meals were to be consumed and recorded in the Zoe study app, as described below. Standardized meals were provided frozen and participants were therefore instructed to store these in the freezer upon arriving at home. Each evening, participants were instructed to transfer the meals for the following day from the freezer to the refrigerator, so they were defrosted in time for consumption. Participants were instructed to fast for a minimum of 8 hours prior to consuming a test breakfast meal and to limit exercise and drink only still, not sparkling, water in moderation during fasting periods (including during the overnight fast).

In addition to water, participants were allowed to have one tea or coffee with their muffin-based set breakfasts. Tea or coffee could contain up to 40 mL 0.1% (skimmed) fat cow's milk, and no sugar or other

sweeteners were not allowed. If participants consumed coffee or tea with their breakfasts, they were instructed to consume this drink consistently, in the same strength and amount, alongside all muffin-based test meals throughout the study. Participants could choose to consume a single tea or coffee alongside their lunch meal, subject to the same restrictions as the breakfast drink, and were asked to keep this drink consistent between test lunches on the study. They were instructed to consume all muffin-based meals within 10 minutes and the OGTT meal within 5 minutes, and to notify study staff if this was not achieved. After consuming their breakfast, participants were instructed to fast for 3 or 4 hours after meal consumption (depending on test meal protocol; protocol Group 1 had a fasting period of 3 hours for Meal 5, and 4 hours for all other meals; in Groups 2 and 3 had a fasting period of 3 hours for all breakfast meals, excluding combinations of breakfast and lunch, where fasting periods were 4 and 2 hours, respectively). They were advised to limit exercise and drink only plain, still water and when possible and safe, to avoid taking any over-the-counter medication during fasting periods.

Participants were instructed to not consume any food or drink other than water alongside the OGTT (Meal 5), to refrain from movement or activity as much as possible for the first hour following the breakfast, and avoid physical activity during the 3-hour fasting period that followed it. Participants were allowed to consume other foods as they normally would once they had completed the fasting period on the study day. Test meals and any dietary intake consumed within fasting periods, including accompanying drinks, were recorded in the Zoe study app by participants with the exact time at consumption and ingredient quantities so that compliance could be monitored by study staff. Only test meals that were completed according to instructions were included in analysis.

*Test meal preparation*

Test meals were prepared and packaged in the Dietetics Kitchen (Department of Nutritional Sciences, King's College London, London, UK) using standard ingredients; plain flour, sugar, baking powder, vanilla essence, milk, egg, salt, high-oleic sunflower oil, whey protein powder, chocolate milkshake powder (Nesquik, Nestle, Gatwick, UK), and commercially available fiber bars (Chocolate Fudge Brownie, Fibre One, General Mills, MN, US; Goodness Bar Apple & Walnut, The Food Doctor, Hessle, UK). All test meal muffins were stored frozen until delivery to participants, who were instructed to store them frozen at home. Powder sachets and fiber bars were stored at room temperature until consumption, while long-life milk was stored at room temperature until delivery to the participant and then refrigerated in their home-phase. Test meals baked for use by the US cohort were shipped frozen, under temperature controlled-conditions, to the US to limit variability of the intervention. Test meal drinks were prepared by the participant at home by mixing pre-portioned powder sachets with long-life milk provided (Meal 1: Metabolic Challenge Meal, 220 mL 0.1% fat milk; Meal 8: High Protein, 200 mL 1.6% fat milk). The OGTT consisted of a pre-portioned powdered glucose sachet which participants mixed with 300 mL water in the UK. In the US, participants were provided with pre-mixed OGTTs ready for consumption (Cat# 82028-512; VWR, US).

*Blinding and assignment of interventions*

The PREDICT study followed a total of 3 different test meal protocols. The first 196 participants (Group 1, UK only) received 9 meals to be consumed for the first 9 days of the at-home phase, and completed DBS corresponding to the Medium Fat & Carb meal (Meal 2) and the "High Fat 1" meal (Meal 3) in duplicate. Following a pre-planned interim analysis in August 2018, the study team revised the protocol for Group 2 of the study (n=380 UK, n=2 US) with 2 additional breakfasts (High Protein (Meal 8) in duplicate, randomised between Days 10-12) and lunches (Medium Fat & Carb (Meal 2), on Days 2 and 3). The High Fat 1 breakfast (Meal 3) was substituted with a new higher fat breakfast, "High Fat 2" (Meal 7) followed by a Medium Fat & Carb lunch (Meal 2) to increase the fat content of the existing meal and elicit a higher TG response, in order to reach detectable limits for DBS analysis. The High Protein meal (Meal 8) was added to investigate the effect of protein on glycaemic responses, given that high study compliance up to this date made an additional meal possible. This group completed DBS corresponding to the High Fat 2 meal (Meal 7) and Medium Fat & Carb (Meal 2) as the breakfast and lunch combination, and the High Carb (Meal 4) stand-alone breakfast. The final test meal protocol, Group 3 (n=405 UK, n=93 US), mirrored Group 2 in both meals and DBS, with the exception of two days; (1) On Day 3, participants consumed the a Metabolic Challenge breakfast (Meal 1) instead of Meal 7 to compare TG response between clinic- and home-based DBS, followed by a Medium Fat & Carb lunch; (2) the displaced High Fat 2 breakfast (Meal 7) was moved to between Days 10-12 and one High Protein breakfast (Meal 8) was eliminated. A full summary of the meals consumed by each Group appears in **Table 4**. Meal order in each group was randomised using Microsoft Access for each participant, using a 2-block randomisation and 1 non-randomised block as denoted in **Table 4.**

Participants were blinded to the nutrient composition of test meals. Test meals were labelled with a barcode and randomization code to allow study staff to identify the composition of the meal consumed, but were also labelled with text for the participant as "Day 1", "Day 2", etc. so that they consumed meals on the proper day according to their randomisation scheme. To log their consumption of the test meal, participants were asked to scan the barcode with the Zoe study app immediately before consumption.

*Food logging mobile study app.*

The Zoe study app was developed to support the PREDICT 1 study by serving as an electronic notebook of study tasks, a tool for recording all dietary intake and a portal for communication with study staff. The app sent participants notifications and reminders to complete tasks at certain time-points, such as when their test lunch meals and DBS were due. The app also prompted participants to report their hunger and alertness levels on visual analogue scales truncated from Flint *et al*[8], by displaying the questions "how hungry are you?" and "how alert are you?" above the scales, at 0 minutes (time of logging) and regular intervals thereafter following the logging of a breakfast, lunch or dinner meal. Participants were asked to

log in the study app any exercise which would not be well captured by a wrist-affixed accelerometer, such as cycling or weight-lifting, where their wrists were stationary during the exercise. Participants logged their full dietary intake using the study app over the 14-day study period, including all standardized test meals and free-living foods, beverages (including water) and medications. Data logged into the Zoe study app was uploaded onto a digital dashboard in real time and reviewed and assessed for logging accuracy and study guideline compliance by study staff.

Study staff trained all participants at their baseline clinic visit on how to accurately weigh and record dietary intake through the Zoe study app, using photographs, product barcodes, product-specific portion sizes, and digital scales. The Zoe study app contained a database of generic and branded food items with nutritional information sourced from generic data sources (Composition of Food Integrated Dataset (COFID), https://www.gov.uk/government/publications/composition-of-foods-integrated-dataset-cofid; United States Department of Agriculture (USDA), https://ndb.nal.usda.gov/ndb/), commercial food databases under license (Nielsen; Tesco; Nutritionix), back-of-pack information from commercial providers, and publicly available restaurant nutritional data. It also allowed participants to photograph back of pack labels in cases where this information was missing from the nutritional database, and where possible the photographed information was entered into the database by study staff. Study staff reviewed all data logged into the study app by participants and were able to identify potential food logging errors by comparing the logged portion sizes with attached photographs. Any uncertainties were actively clarified with the participant through the app messaging system or via phone while the participant was on the study in real time. Free-living dietary data logged through the app was categorised into separate meals for analysis against continuous glucose data.

## 3. Participant recruitment and compliance

*Recruitment and eligibility screening*

In the UK, participants (target n = 1,000) were recruited from the TwinsUK cohort, an ongoing research cohort described elsewhere[11] and the general population using online advertising (**Figure 2a**). In the US, participants (target n = 100) were recruited through online advertising, research participant databases and Rally for Research (https://rally.partners.org/), an online recruiting portal for research trials (**Figure 2b**). During an initial recruitment interaction with study staff, participants were screened minimally for eligibility criteria (e.g. phone compatibility for the logging app, pregnancy, food allergies to test meal components) and to answer any questions from participants. If potential subjects were interested in participating and remained eligible, they then underwent a detailed eligibility assessment performed by a study coordinator over the phone. Participants were provided with an informational booklet as a “Participant Information Sheet” that fully described the study including both the clinical and at-home phase of the study to assist in the informed consent process.

*Participant compliance*

In the event of possible protocol non-compliance or if the CGM was damaged, fell off, or experienced a general failure, the following approaches were taken and priority was given to those completing and obtaining glucose data for the at-home meals. If a CGM was damaged or fell off within Days 2-12 of the home-phase (when test meals were in progress), study staff provided a replacement CGM. If this occurred during the free-living Days 13-14 (after test meals were complete), the monitor was not replaced. Whenever possible, CGMs were replaced within 24 hours, and the participant was asked to pause their test meals until the new CGM arrived. Free-living days were changed to test meal days where necessary for participants to complete all test meals. As long as the CGM was replaced within a maximum of 48 hours, the participant was able to complete all test meals within the standard at-home period of the trial.

If for any unforeseen reason a replacement CGM was not provided within 48 hours, the participant was given the option to continue past the 14 days to complete all test meals, such that their responses for all meals would be recorded. In the event of a test meal error (e.g. muffins were not consumed with their designated milkshake or fiber bars as per labelled instructions), participants were provided with the option to repeat the meal on either Day 13 or Day 14 if the research team was able to provide the participant with a replacement meal, within the 14 day period.

All participants must have consumed the baseline breakfast and provided baseline fasting bloods. Failure to meet these two requirements resulted in termination of continued study participation as other endpoints would have been of no value. If any other sample collections, dietary interventions, or blood timepoints were missed or unable to be completed, the subjects continued with participation and these missing items were logged by the study staff.

*Adverse events (AE) monitoring*

As the study was a dietary intervention study, safety monitoring focused on unanticipated events involving risks to participants, including any unanticipated medical occurrence in a patient or study subject. Serious AEs (SAEs) were considered those that resulted in death, were life threatening, required or prolonged hospitalization, or resulted in persistent or significant disability or incapacity. AEs were recorded and reported to the Principal Investigator (PI) at each site and the sponsor of the study. All SAEs, both related and unrelated, were recorded and reported immediately to the PI and the sponsor as well as appropriate research ethics committees.

*Participant retention*

Given the complexity of the study, there was a high-risk of non-compliance without study staff support. As a result, we tried to ensure that we kept participants motivated and supported to be able to comply. The key objective of the clinical research coordination staff at St. Thomas's Hospital and MGH and the in-study engagement team (through study app interactions) was to anticipate and troubleshoot any issues before they arose by designing engagement such that it increases compliance; striking a balance between overcontact and support; emphasizing the motivation for completing the study and to continuously motivate and engage. In addition to the study booklet and the training sessions delivered during the baseline clinical visit, two methods were used: pro-active engagement – calls and in-app messages to participants at predefined points in time in order to remind them, congratulate/motivate them or support them in certain actions; and reactive engagement – calls and/or in-app messages to participants based on certain real-time compliance triggers to help educate and support them in logging certain things better (particularly around dry blood spots and meals). The engagement team used a monitoring dashboard which summarized compliance measures by participant, specifically the recordings of their test meals, DBS, and free-living consumption with the aid of photographs, participant-written notes and nutritional information where available. This dashboard was a key tool in ensuring high study compliance as the engagement team actively corrected miscompliance and clarified any uncertainties in real time. The engagement team was available 7 days a week and had a process to ensure that even out of hours inquiries were answered expediently. They utilized a 2-level escalation process for questions that involved scientific input and/or any AE reporting.

## 4. Ethics and dissemination

Ethical approval for the study was obtained in the UK from the Research Ethics Committee and Integrated Research Application System (IRAS 236407) and in the US from the Institutional Review Board (Partners Healthcare IRB 2018P002078). The trial was registered on ClinicalTrials.gov (registration number: NCT03479866) and was run in accordance with the Declaration of Helsinki and Good Clinical Practice. Study procedures were only carried out after having received written informed consent from each participant.

### *Informed consent process*

Interested participants were sent the participant information sheet and informed consent form by mail or by email, at least 2 weeks before their baseline visit was scheduled to allow adequate time to read it, understand the protocol, the risks and burdens and benefits, and be provided an opportunity to ask questions. Only participants who had the capacity to provide consent for themselves were recruited.

Consent was collected by the designated HSP/GCP trained researchers responsible for conducting the study at each site, who were experienced in patient communication and consent taking. The researcher

discussed in detail the procedures that would happen as part of the study prior to collecting consent. The researcher also explained to the participants that they are completely free to refuse to enter the study or to withdraw from it at any time without obligation to give a reason. Participants were explicitly given the opportunity to ask any additional questions before they signed the consent form. The participant was provided with a copy of their signed and dated consent form and any other written information and were instructed to retain these documents. It was essential that participants fully understand the requirements of the study, and therefore participants without a good understanding of verbal or written English were not included in the study. Translation and interpretation was not accommodated as the at-home phase of the study required that participants use the food logging app which was unavailable in other languages.

## 5. Protocol versions and amendments

The UK PREDICT study followed 3 major protocols between June 2018 – May 2019, with the first version being approved by the Hampstead Research Ethics Committee on 27 March 2018. The second version (v1.1) contained a simplified protocol with the removal of using bio-impedance to measure body composition. The first participants were enrolled in June 2018 under Protocol v1.1.

A second amendment (Protocol v1.2) allowed additional test meals to be included in the home-phase (described earlier, **Table 4**) and participants' logging of gut transit time by using a Metabolic Challenge Breakfast (Meal 1) at clinic dyed blue with food coloring. The DBS protocol was also changed; the original DBS at-home protocol tested both TG and C-peptide on Days 2-5 (Group 1). Following an interim study break in August 2018, it was found that the DBS protocol should be optimised by changing the timing of DBS tests according to physiological peaks in biomarkers (TG or C-peptide) and by increasing the fat and carbohydrate content of test meals on DBS days to exaggerate the response. Starting on 28 Aug 2018, TG was measured on Days 2-3 at fasting, 300 and 360 minutes post-prandially following test meals as described above, while C-peptide was quantified on Days 4-5 at fasting, 30 and 120 minutes post-prandially as described for Group 2. Protocol v1.2 also included a second saliva sample collection on the clinic day, at 30 minutes after the metabolic challenge breakfast, to measure salivary amylase production post prandially and provide a comparison to fasted amylase levels. A further amendment (v1.3) was approved to lower the lower BMI limit for eligibility to 16.5kg/$m^2$ (originally 20 kg/$m^2$). Meal changes which resulted in Group 3 were implemented in January 2019.

Multiple other minor changes not impacting the intervention or primary study outcomes were made to the protocol internally. Ambulatory blood pressure monitoring during the clinical visit was introduced on 8 Aug 2018 as an exploratory measure to potentially use as an input variable in predicting metabolic responses to meals.

A 'cracker test' (using Carr's Table Water or Rakusen's Snackers) was added to the protocol on 9 July 2018 as an exploratory test for potential use an input variable in data analysis to predict metabolic

response to meals (based on the link with salivary amylase). The test was done in triplicate and involved recording the time taken to perceive a sweet taste while chewing a plain cracker.

In the US, on 15 November 2018, UK protocol v1.3 was mirrored and approved by the Partners Healthcare IRB as PREDICT-US v1.0 The first two participants were enrolled in December 2018. On 3 January 2019, the IRB approved an amendment (PREDICT-US v2.0) to address meal changes introduced in the UK for Group 3 and to allow the use of multiple CGMs on the same participant. No other major amendments to the intervention protocol were made during the study period in the US. Minor amendments were made to improve recruitment (addition of direct advertisements and patient repositories, for example) or administrative changes to study staff.

# Troubleshooting

# Time Taken

The initial study commenced in June 2018 in the United Kingdom. In December 2018, an independent trial mirroring the UK protocol was launched in parallel at Massachusetts General Hospital to serve as a US-based validation cohort. The studies were completed in May 2019.

# Anticipated Results

### 1. Outcomes

This trial aimed to examine the factors that predict individual postprandial responses in context of the individual and meals consumed. For primary outcome reporting, the study focused on the following postprandial metabolic response between 0 and 6 hours for blood triglycerides, glucose, and insulin concentrations to the sequential mixed-nutrient dietary challenges during the clinical visit on Day 1. Secondary main outcomes focused on additional data collected over the subsequent 13-day at-home phase. Lipemic responses and C-peptide concentrations (as a surrogate for insulin) were measured using dried blood spot (DBS) cards collected at home at three post-prandial time points. In addition, glycemic responses to the test meals in context of their different macronutrient content were examined using CGM data.

A summary of all endpoints that were collected and considered as *a priori* outcomes of the study are presented in **Table 5** in no particular order.

*Sample collection, handling and analysis*

Venous blood sample collection

Participants were cannulated in the forearm antecubital vein. Venous blood was collected at baseline (prior to a test breakfast) and at 9 time-points postprandially (15, 30, 60, 120, 180, 240, 270, 300, and 360 minutes). Plasma glucose was analyzed from blood samples collected into 2 mL fluoride oxalate tubes and centrifuged at 1900 *g* for 10 min at 4 °C. Serum C-peptide, insulin, TG, fasting lipid profile, metabolomics, thyroid stimulating hormone, alanine aminotransferase, and liver function panel were analyzed from blood samples collected into 5 mL gel separator serum tubes and allowed to stand at room temperature before centrifuging at 1900 *g* for 10 min at 4 °C. Samples were aliquoted and stored at -80 °C. Blood, for complete blood count (CBC) analysis, was collected into 6 mL EDTA tubes, kept at 4 °C and analyzed within 12 hours of collection. Spare serum aliquots were stored for future analysis.

Serum biomarkers

In the UK, insulin, glucose, TG and C-peptide analysis was conducted by Affinity Biomarkers Labs (London, UK). Glucose and TG analyses were conducted on a Siemens ADVIA 1800 using Siemens assay kits (Siemens Healthcare Diagnostics Ltd, Surrey, UK). TG was analyzed using the ADVIA chemistry TG method based on the Fossati three-step enzymatic reaction with a Trinder endpoint. Glucose was analyzed using the ADVIA chemistry glucose oxidase (GLUO) method (based on the modified method of Keston). C-peptide and insulin were analyzed using the Siemens ADVIA Centaur XP systems using a two-site sandwich immunoassay. Complete blood count (CBC) was measured by Viapath (London, UK) for the UK cohort using standard automated clinical chemistry techniques. The inter-assay coefficient of variation for PREDICT samples analyzed by Affinity were: insulin 3.4%, C-peptide 7.9%, TG 3.7%, and glucose 2.6%.

In the US, CBC was established using fresh blood samples in the MGH Core Laboratory. HbA1C tests were performed by the MGH Diabetes A1c lab. Glucose, insulin, TG, and C-peptide were conducted by Quest Diagnostics (Boston, MA) using standard automated clinical chemistry techniques.

Upon completion of the US study, frozen serum and plasma samples were sent from the US to the UK and the entire cohort had liver function panel, full lipids (TC, HDL-C LDL-C and TG), thyroid stimulating hormone and alanine aminotransferase measurements performed by Affinity Biomarkers Labs. Details described elsewhere[9].

Continuous glucose monitoring

Interstitial glucose was measured every 15 minutes using Freestyle Libre Pro continuous glucose monitors (Abbott, Abbott Park, IL, US). Monitors were fitted by trained clinical practitioners on the upper, non-dominant arm at participants' baseline clinical visit and covered with Opsite Flexifix adhesive film (Smith & Nephew Medical Ltd, Hull, England) for improved durability. CGMs were worn for the entire study

duration (14 days) and removed on Day 15, after which they were mailed back to study staff. Given the CGM device requires time to calibrate once fitted to a participant, CGM data collected 12 hours and onwards after activating the device was used for analysis. To understand the accuracy of the CGM device, a subgroup of participants (n=377) were fitted with two monitors on the same arm. The Coefficient of Variation (CV =11.75%) was small and the correlation (r = 0.97) was high (for $glucose_{iAUC0-2h}$).

Activity and sleep

Energy expenditure was measured using a triaxial accelerometer (AX3, Axivity, Newcastle Upon Tyne, UK) fitted by clinical practitioners at the baseline clinic visit on the non-dominant wrist and worn for the duration of the study (except during water-based activities, including showers and swimming), after which they were removed on Day 15 and mailed back to study staff. Accelerometers were programmed to measure acceleration at 50 Hz with a dynamic range of ±8 *g* (where g refers to local gravitational force equal to 9.8 $m/s^2$). Non-wear periods were defined as windows of at least 1 hour with less than 13 *mg* for at least 2 out of 3 axes, or where 2 out of 3 axes measured less than 50 *mg*. Windows of sleep were measured using methods described elsewhere [10].

Urine samples

Urine samples were collected in standard clinical sterile 10 mL collection containers at home prior to the baseline visit and brought to the clinical site by the participants. Upon receipt, urine samples were stored at 4 °C immediately and aliquoted and stored at -80 °C within 4 hours of receipt for future analysis of metabolites.

Saliva samples

A saliva sample of approximately 2 mL was collected in 5 mL sterile universal tubes at two timepoints during the baseline clinical visit; the first when fasted at baseline, and the second 30 minutes after consumption of breakfast. Samples were immediately aliquoted and stored at -80 °C for future analysis of metabolites.

DBS collection

TG and C-peptide were quantified from DBS tests completed by participants at the baseline visit, with the help of study staff (while fasted and at 300 minutes post-breakfast; for method validation) and on the

first 4 days of the home-phase while consuming test meals, at timepoints described in **Table 6**.

The Zoe study app sent participants reminders to complete their DBS tests at due times, which participants then logged in the app by recording the time at testing and a photo of the completed card for quality assessment by study staff. Test cards not meeting the quality protocol (multiple small spots or inadequate coverage) were not included for analysis. Test cards were stored in aluminium sachets with desiccant once completed and placed in the fridge at the end of the study day or until participants mailed them back to the study site. DBS cards were frozen at -80 °C upon receipt in the laboratory until being shipped to Vitas for analysis (Vitas Analytical Services, Oslo, Norway).

DBS validation

DBS C-peptide and TG concentrations were validated during PREDICT 1, against venous serum concentrations collected during the baseline clinic visit while fasted and at 300 minutes following the clinic breakfast. Correlations between the two methods were found to be high; for TG (1,772 pairs) Pearson's r=0.94; for C-peptide (1,679 pairs) Pearson's r=0.91.

Quantification of total TG from DBS

From the DBS sample, 2 punches were taken and transferred into a HPLC vial and lipids extracted with methanol at 600 rpm and 25 °C for 3 hours. The resulting extract was processed with a TG kit (FUJIFILM Wako Chemicals GmbH, Neuss, Germany) at 600 rpm and 37 °C for 2.5 hours and the reaction products were subsequently analyzed by HPLC-UV. HPLC was performed with a HP 1260/1290 infinity liquid chromatograph (Agilent Technologies, Palo Alto, CA, US) using UV detection. The analyte was separated from matrix components on a 4.6 mm x 100 mm reversed phase column at 40 °C. A one-point calibration curve was made from analysis of TG standard after enzymatic reaction with the kit. The analytical method is linear from 0.5-6 mmol/L with a quantification limit of 0.3 mmol/L.

Quantification of C-peptide from DBS

C-peptide in DBS were assayed using a Mercodia solid phase two-site enzyme immunoassay (ELISA; Mercodia AB, Uppsala, Sweden). Three spots were punched into the kit plate with anti-C-peptide antibodies bound to the well. Assay buffers were added and C-peptide extracted from the spots at 4 °C. After washing, peroxidase-conjugated anti-C-peptide antibodies were added and after the second

incubation and a washing step, the bound conjugate was detected by reaction with 3,3, 5, 5-tetramethylbenzidine (TMB). The reaction was stopped by adding acid to give a colorimetric endpoint that was read spectrophotometrically at 450 nm.

Stool sample collection

Participants collected two stool samples at home; the first prior to their clinical visit and the second once they have completed the study, on Day 15. In the UK, baseline samples were collected using the EasySampler collection kit (ALPCO, NH, US), while the post-study sample and both US samples were collected using the Fecotainer collection kit (Excretas Medical BV, Enschede, The Netherlands), and sampled by sterile spatula into fecal collection tubes. For baseline samples one fresh unfixed sample was placed into a sterile universal collection container (Sarstedt, Australia, Cat #L0263-10) and one sample was placed into a tube containing DNA/RNA Shield buffer (Zymo Research, CA, US, Cat #R1101) and stored at ambient temperature until return to the study staff. At the end of the home-phase, samples were collected similarly, but only sampled into a DNA/RNA Shield buffer tube and sent by standard mail to study staff. Upon receipt in the laboratory, samples were homogenised, aliquoted and stored at -80 °C in Qiagen PowerBeads 1.5 mL tubes (Qiagen, Germany) for later analysis of 16S rRNA sequencing and metagenomic sequencing. The sample collection procedure was tested and validated internally comparing different storage conditions (fresh, frozen, buffer), different DNA extraction kits (PowerSoilPro, FastDNA, ProtocolQ, Zymo), and different sequencing technologies (16S rRNA and arrays).

Microbiome 16S rRNA sequencing

We performed 16S rRNA sequencing to profile the gut microbiome in fecal samples preserved in DNA/RNA shield buffer. The DNA was isolated by QIAGEN Genomic Services using DNeasy® 96 PowerSoil® Pro. Optical density measurement was done using Spectrophotometer Quantification (Tecan Infinite 200). For 16S sequencing, the V4 hyper-variable region of the 16S rRNA gene was then amplified at Genomescan, Leiden, Netherlands. Libraries were sequenced for 300 bp paired-end reads using the Illumina NovaSeq6000 platform. In total, 9.6 Pbp were generated and raw reads were rarefied to 360k reads per sample.

Microbiome metagenomic sequencing

For metagenomic sequencing intended for future analyses, before library preparation and sequencing, the quality and quantity of the samples were assessed using the Fragment Analyzer (Agilent Technologies, Inc.) according to manufactures guidelines. Samples with a high-quality DNA profile were further processed. The NEBNext® Ultra II FS DNA module (cat# NEB #E7810S/L) was used for DNA

fragmentation, end-repair, and A-tailing. For adapter ligation, the NEBNext® Ultra II Ligation module (cat# NEB #E7595S/L) was used. The quality and yield after sample preparation were measured with the Fragment Analyzer. The size of the resulting product was consistent with the expected size of approximately 500-700 bp. Libraries were sequenced for 300 bp paired-end reads using the Illumina NovaSeq6000 platform according to manufacturer's protocols. 1.1 nM library was used for flow cell loading. NovaSeq control software NCS v1.5 was used. Image analysis, base calling, and the quality check was performed with the Illumina data analysis pipeline RTA3.3.5 and Bcl2fastq v2.20.

### 2. Sample size

1050 individuals were targeted for recruitment at baseline, which assumed a 5% loss-to-follow-up to correspond to 1000 evaluable individuals at the end of the study. The study was designed to be able to predict lipid and glucose postprandial responses based on individual characteristics, correlations of a magnitude of r=0.13 ($R^2$=0.017) will be detected as statistically significant with p<0.005 and 80% power. Effects of r=0.165 ($R^2$=0.027, i.e. explaining 2.7% of interindividual variation) are necessary to achieve 80% power with p<0.00001, i.e. accounting for 5000 independent hypothesis tests.

### 3. Statistical analysis

The primary outcome statistical analysis is described in detail in the primary report. Briefly, to reduce the dimensionality of the data and thus the risk of data overfitting, we began by undertaking principal component analysis (PCAs) with orthogonal transformation (varimax procedure). From this we derived PCs for each set of key explanatory variables: individual characteristics (20 PCs), microbiome (40 PCs), meal composition (1 PC), habitual diet (5 PCs) and meal context (5 PCs). Additional PCs were generated for each set of anthropometric, biochemical/clinical, physical activity and sleep data to investigate their role in predicting the postprandial responses. Using the PCs derived from the key explanatory variables, we predicted the dependent variables ($TG_{6h\text{-}rise}$, $glucose_{iAUC0\text{-}2h}$, $C\text{-}peptide_{1h\text{-}rise}$) using multivariable regression (GLM) and report adjusted r-squared values for each dependent variable and set of explanatory variables. Data analyses were performed using Python and sklearn library.

## References

1. Collaborators, G.B.D.D. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **393**, 1958-1972 (2019).

2. Tebani, A. & Bekri, S. Paving the Way to Precision Nutrition Through Metabolomics. *Front Nutr* **6**, 41 (2019).

3. Kolovou, G.D., *et al.* Assessment and clinical relevance of non-fasting and postprandial triglycerides: an expert panel statement. *Curr Vasc Pharmacol* **9**, 258-70 (2011).

4. Astley, C.M., *et al.* Genetic Evidence That Carbohydrate-Stimulated Insulin Secretion Leads to Obesity. *Clin Chem* **64**, 192-200 (2018).

5. Blaak, E.E. *et al.* Impact of postprandial glycaemia on health and prevention of disease. *Obes Rev* **13**, 923-84 (2012).

6. Stumvoll, M., Goldstein, B.J. & van Haeften, T.W. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet* **365**, 1333-46 (2005).

6. Verdi, S. *et al.* TwinsUK: The UK Adult Twin Registry Update. *Twin Res Hum Genet*, 1-7 (2019).

8. Flint, A., Raben, A., Blundell, J.E. & Astrup, A. Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. *Int J Obes Relat Metab Disord* **24**, 38-48 (2000).

9. Soininen, P., Kangas, A.J., Wurtz, P., Suna, T. & Ala-Korpela, M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ Cardiovasc Genet* **8**, 192-206 (2015).

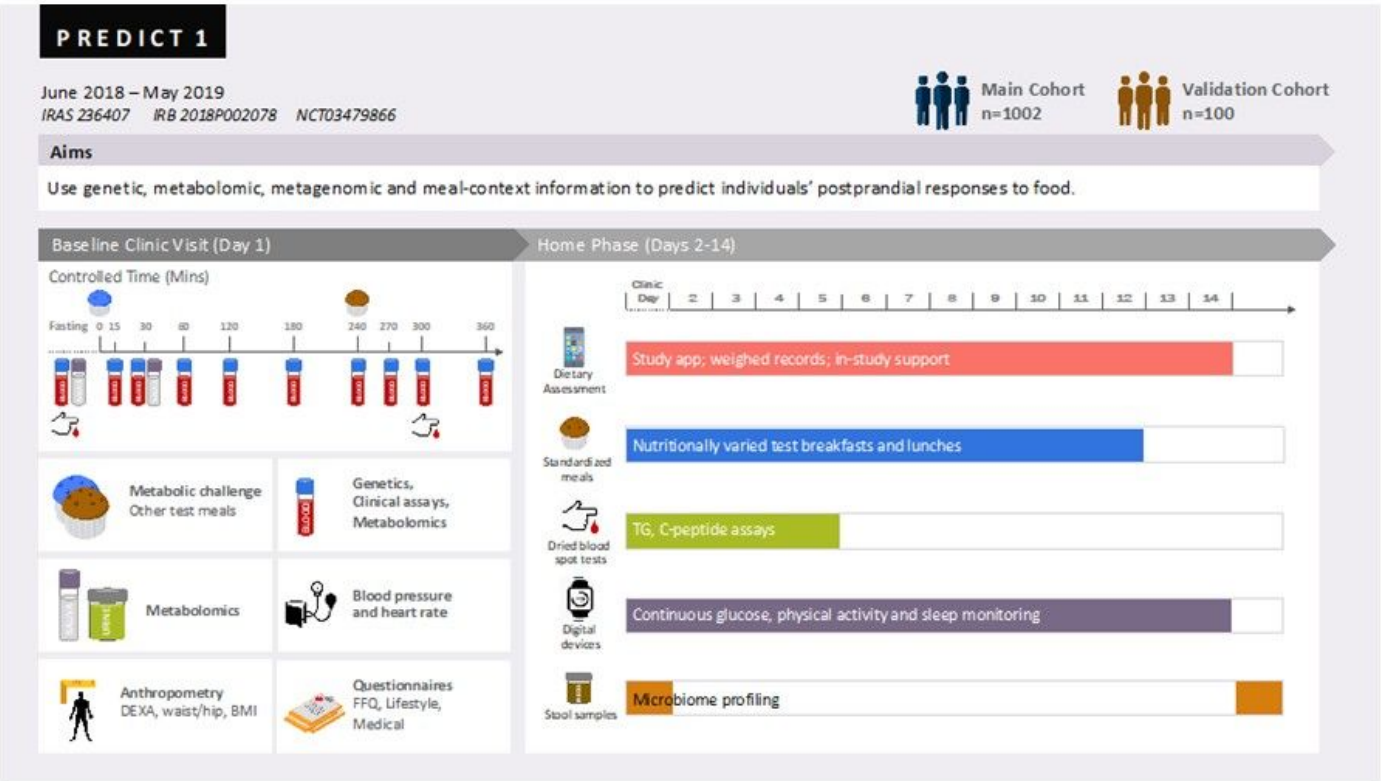
10. van Hees, V.T., *et al.* Estimating sleep parameters using an accelerometer without sleep diary. *Sci Rep* **8**, 12975 (2018).

Moayyeri, A., Hammond, C.J., Hart, D.J. & Spector, T.D. The UK Adult Twin Registry (TwinsUK Resource). *Twin Res Hum Genet* **17**, 144-9 (2013).

# Acknowledgements

We express our sincere thanks to the participants of the PREDICT 1 study. We also thank Carbs & Cals and Nielsen nutritional food databases for their contribution to the Zoe study app. We thank the staff of Zoe Global, especially Kate Creedon, and the staff of the Department of Twin Research, the Clinical and Translational Epidemiology Unit and the Translational and Clinical Research Center at Massachusetts General Hospital, especially Marjorie A. Noone, Jane Hubbard, and Amanda Griffin, for their tireless work in contributing to the running of the study and data collection.

# Figures




## Figure 1

Figure 1. Experimental design. The PREDICT 1 study included a primary UK-based cohort (nmax=1,002) and an independent US-based validation cohort (nmax=100). The study consisted of a 1-day clinical visit at baseline followed by a 13-day at-home period. At baseline (Day 1), participants arrived fasted and were given a standardised metabolic challenge meal for breakfast (0h; 86g carbohydrate, 53g fat) and lunch (4h; 71g carbohydrate, 22g fat). Fasting and postprandial (9 timepoints; 0-6h) venous blood was collected to determine serum concentrations of glucose, triglycerides (TG), insulin and C-peptide (a surrogate for insulin). Stool, urine and saliva samples, anthropometry, and a questionnaire querying habitual diet, lifestyle and medical health were obtained at baseline. During the home-phase (Days 2-14), participants consumed standardized test meals varying in macronutrient composition (Table 3) and order (Table 4) while wearing digital devices to continuously monitor their blood glucose (continuous glucose monitor; CGM), physical activity and sleep. Capillary blood was collected using dried blood spot cards, during the clinic visit and at home, to analyze fasting and postprandial concentrations of TG and C-peptide. Participants were supported throughout the study with reminders and communication from study staff delivered through the ZOE study app. A second stool sample was collected at home by

participants following completion of the study and all devices and samples were mailed back to study staff.

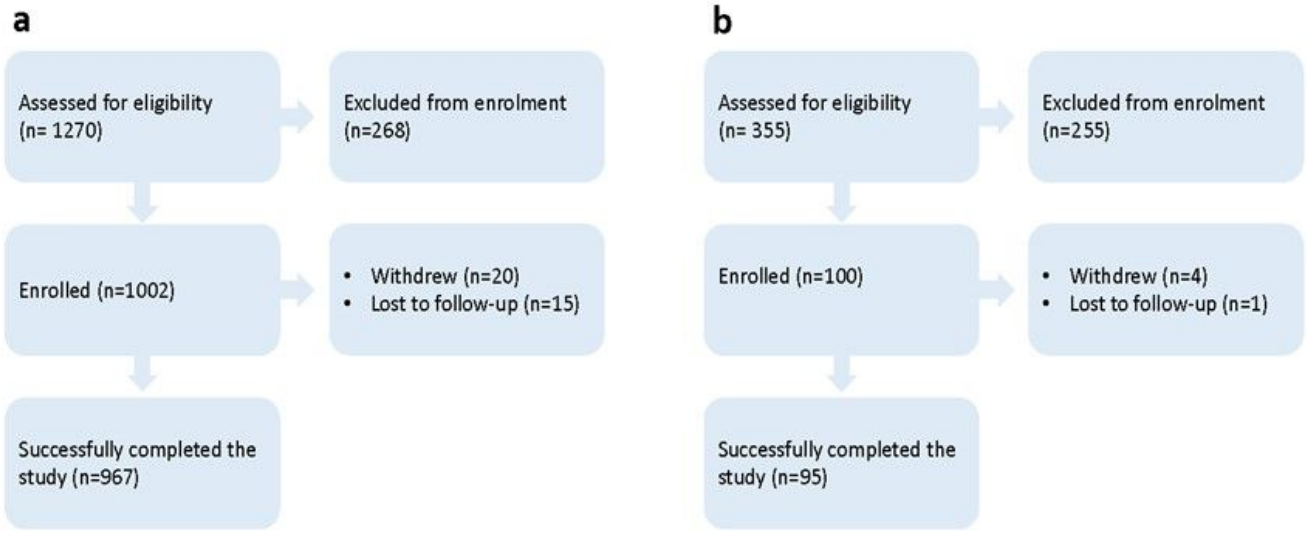



**Figure 2**

Figure 2. Consort Diagrams for (a) UK and (b) US study populations

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table4.Testmealordering.pdf
- Table5.PREDICT1aprioriouteomevariables.pdf
- Table6.DriedBloodSpotprotocol.pdf
- PREDICT1Protocol.pdf
- Table1.PREDICT1eligibilitycriteria.pdf

- Table2.Studytimeline.pdf
- Table3.Testmealcomposition.pdf