



# Human postprandial responses to food and potential for precision nutrition

Sarah E. Berry<sup>1,15</sup>, Ana M. Valdes<sup>2,3,15</sup> ✉, David A. Drew<sup>4</sup>, Francesco Asnicar<sup>5</sup>, Mohsen Mazidi<sup>6</sup>, Jonathan Wolf<sup>7</sup>, Joan Capdevila<sup>7</sup>, George Hadjigeorgiou<sup>7</sup>, Richard Davies<sup>7</sup>, Haya Al Khatib<sup>1,7</sup>, Christopher Bonnett<sup>7</sup>, Sajaysurya Ganesh<sup>7</sup>, Elco Bakker<sup>7</sup>, Deborah Hart<sup>6</sup>, Massimo Mangino<sup>6</sup>, Jordi Merino<sup>4,8,9</sup>, Inbar Linenberg<sup>7</sup>, Patrick Wyatt<sup>7</sup>, Jose M. Ordovas<sup>10,11</sup>, Christopher D. Gardner<sup>12</sup>, Linda M. Delahanty<sup>4</sup>, Andrew T. Chan<sup>4</sup>, Nicola Segata<sup>5,15</sup>, Paul W. Franks<sup>6,13,14,15</sup> and Tim D. Spector<sup>15</sup> ✉

**Metabolic responses to food influence risk of cardiometabolic disease, but large-scale high-resolution studies are lacking. We recruited  $n = 1,002$  twins and unrelated healthy adults in the United Kingdom to the PREDICT 1 study and assessed postprandial metabolic responses in a clinical setting and at home. We observed large inter-individual variability (as measured by the population coefficient of variation (s.d./mean, %)) in postprandial responses of blood triglyceride (103%), glucose (68%) and insulin (59%) following identical meals. Person-specific factors, such as gut microbiome, had a greater influence (7.1% of variance) than did meal macronutrients (3.6%) for postprandial lipemia, but not for postprandial glycemia (6.0% and 15.4%, respectively); genetic variants had a modest impact on predictions (9.5% for glucose, 0.8% for triglyceride, 0.2% for C-peptide). Findings were independently validated in a US cohort ( $n = 100$  people). We developed a machine-learning model that predicted both triglyceride ( $r = 0.47$ ) and glycemic ( $r = 0.77$ ) responses to food intake. These findings may be informative for developing personalized diet strategies. The ClinicalTrials.gov registration identifier is [NCT03479866](https://clinicaltrials.gov/ct2/show/study/NCT03479866).**

Effective prevention strategies are required to reduce the immense global burden of nutrition-related non-communicable diseases (NCDs)<sup>1</sup>. Nutritional research and the corresponding guidelines<sup>2–4</sup> focus on population averages. However, the high inter-person variability in response to foods and weight-loss diets<sup>5</sup> demands development of more personalized approaches. Precision nutrition that is based on empirical evidence requires research using multidimensional, high-resolution time-series data from adequately powered studies<sup>6</sup>. The application of technologies to accurately and precisely quantify many postprandial (non-fasting) traits in large cohorts and in real-world settings is extending capabilities in this field of research.

Although fasting blood assays are used in many clinical diagnoses, most people are predominantly in the postprandial state during waking hours. Postprandial lipid, glucose and insulin dys-homeostasis are independent risk factors for NCDs and obesity<sup>7–9</sup>. Postprandial hyperglycemia raises the risk of cardiovascular disease (CVD), coronary heart disease (CHD)<sup>10</sup> and cardiovascular mortality, even in individuals with normal fasting glucose<sup>11</sup>, and postprandial triglyceride level is more predictive of CVD than are fasting concentrations<sup>12,13</sup>, highlighting the relevance of diet and its metabolic consequences in cardiovascular risk.

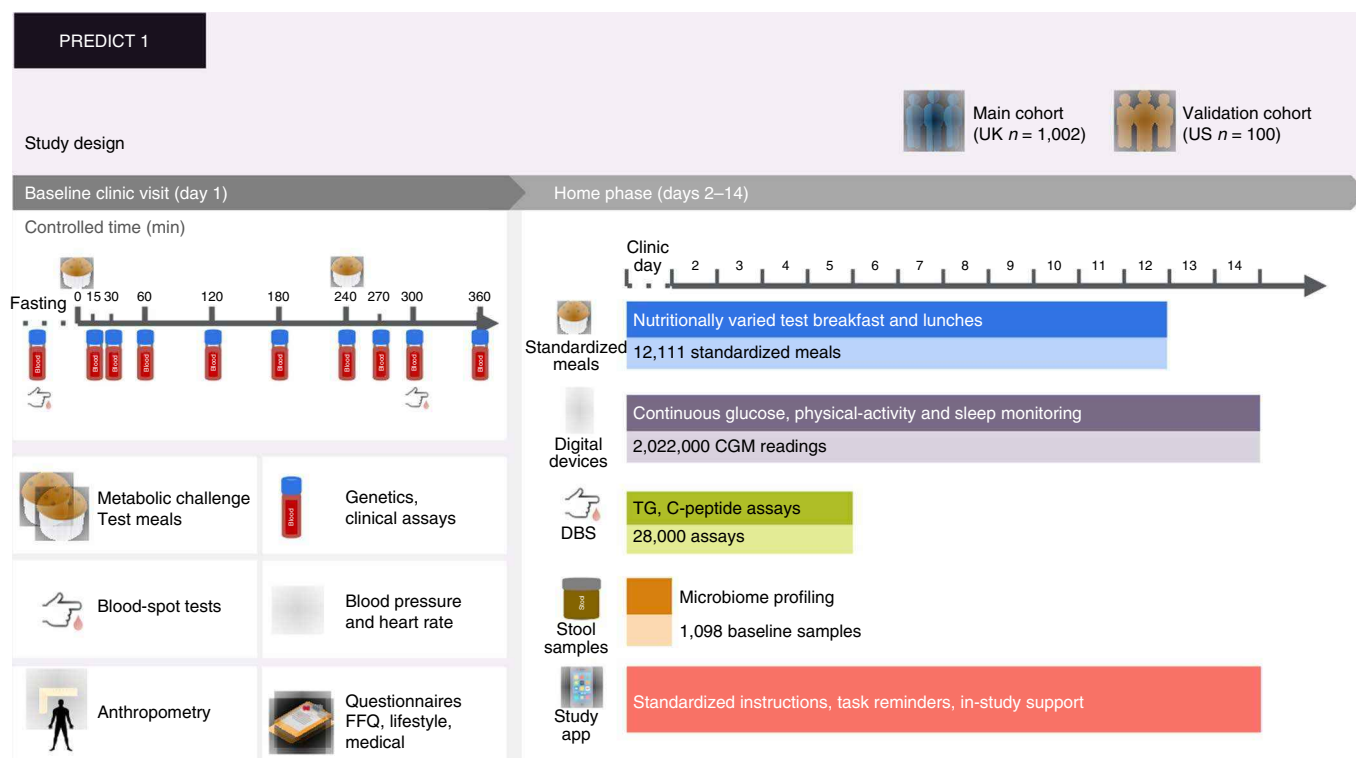
A person's unique postprandial glycemic and lipidemic responses are likely attributable to their biological (for example, microbiome

and nuclear DNA variation) and lifestyle characteristics<sup>2,14</sup>, as has previously been demonstrated in research using specific meals<sup>5</sup>. Although postprandial glycemic responses are important health determinants, glycemic control is just one part of a more complex metabolic equation involving triglyceride (the primary alternative energy substrate to glucose) and insulin (which regulates glucose and triglyceride transport and metabolism)<sup>15</sup>. Thus, characterizing postprandial regulation of lipids and identifying the factors responsible for individual variation could help optimize diet recommendations to target broader improvements in cardiometabolic health.

The personalized responses to dietary composition (PREDICT 1) clinical trial ([NCT03479866](https://clinicaltrials.gov/ct2/show/study/NCT03479866)) was designed to quantify and predict individual variations in postprandial triglyceride, glucose and insulin responses to standardized meals. PREDICT 1 enrolled twins and unrelated adults from the United Kingdom in whom genetic, metabolic, microbiome-composition, meal-composition and meal-context data were obtained to distinguish predictors of individual responses to meals. These predictions were validated in an independent cohort of adults from the United States.

Our findings show wide variations in postprandial responses between people, even identical twins, attributable in large part to modifiable factors. We found that people who experience poor metabolic responses to a given meal are likely to respond poorly to other meals with the same macronutrient profile, and that the overall

<sup>1</sup>Department of Nutrition, King's College London, London, UK. <sup>2</sup>School of Medicine, University of Nottingham, Nottingham, UK. <sup>3</sup>Nottingham NIHR Biomedical Research Centre, Nottingham, UK. <sup>4</sup>Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA. <sup>5</sup>Department CIBIO, University of Trento, Trento, Italy. <sup>6</sup>Department of Twins Research and Genetic Epidemiology, King's College London, London, UK. <sup>7</sup>Zoe Global Ltd, London, UK. <sup>8</sup>Broad Institute of MIT and Harvard, Cambridge, MA, USA. <sup>9</sup>Institut d'Investigació Sanitària Pere Virgili, Universitat Rovira i Virgili, Reus, Spain. <sup>10</sup>JM-USDA-HNRCA at Tufts University, Boston, MA, USA. <sup>11</sup>IMDEA Food Institute, CEI UAM + CSIC, Madrid, Spain. <sup>12</sup>University of Stanford, Stanford, CA, USA. <sup>13</sup>Department of Clinical Sciences, Lund University, Malmö, Sweden. <sup>14</sup>Department of Nutrition, Harvard Chan School of Public Health, Boston, MA, USA. <sup>15</sup>These authors contributed equally: Sarah E. Berry, Ana M. Valdes, Nicola Segata, Paul W. Franks, Tim D. Spector. ✉e-mail: [ana.valdes@nottingham.ac.uk](mailto:ana.valdes@nottingham.ac.uk); [tim.spector@kcl.ac.uk](mailto:tim.spector@kcl.ac.uk)



**Fig. 1 | Experimental design.** The PREDICT 1 study comprised a primary UK-based cohort ( $n_{\max} = 1,002$  participants) and an independent US-based validation cohort ( $n_{\max} = 100$  participants). TG, triglyceride.

correlation between postprandial glucose and triglyceride responses is weak. The postprandial prediction models we have developed could help optimize personalized diet recommendations.

## Results

Baseline clinical measurements were collected from 1,002 healthy adults from the United Kingdom. These consisted of postprandial metabolic responses (0–6 h; blood triglyceride, glucose and insulin concentrations) to sequential mixed-nutrient dietary challenges. Findings were validated a US cohort of 100 healthy adults. Additional data were collected over the subsequent 13-d period at home: postprandial responses to eight meals (seven in duplicate) of differing macronutrient (fat, carbohydrate, protein and fiber) content were measured using continuous glucose monitors (CGMs) and dried-blood-spot (DBS) analysis. The study design is described in detail in the Methods and Fig. 1, and the inclusion criteria and descriptive characteristics of study subjects are presented in Supplementary Table 1. Further information on research design is available in the Reporting Summary.

### Inter- and intra-individual variation in postprandial responses.

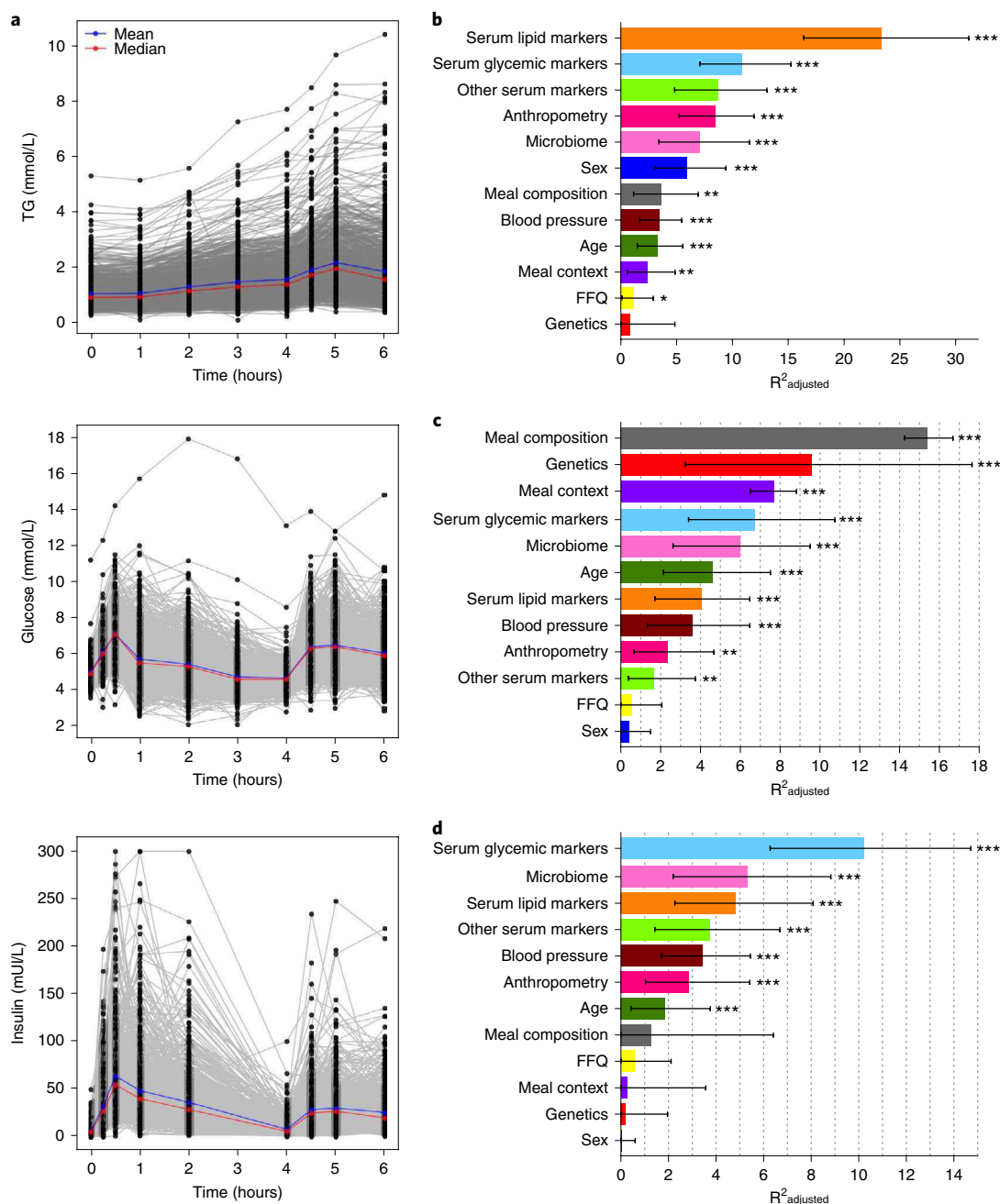
Inter-individual variability in postprandial responses was examined in a tightly controlled clinical setting following the sequential standardized test-meal challenge after fasting (Fig. 2a). The inter-individual patterns of response for each outcome were assessed using Levene's test of variance. Heterogeneity across all postprandial time points (during fasting for 6 h) varied greatly for triglyceride ( $P = 3.931 \times 10^{-11}$ ), glucose ( $P = 2.91 \times 10^{-194}$ ) and insulin ( $P = 2.45 \times 10^{-17}$ ) concentrations. In serum, the population coefficient of variation was higher for postprandial triglyceride<sub>6h-rise</sub> (6 h–0 h triglyceride concentration) (103%) and glucose<sub>iAUC0-2h</sub> (0 h–2 h incremental area under the curve) (68%) compared with fasting values (50% and 10%, respectively). This was not true for insulin<sub>iAUC0-2h</sub> (59%) compared with the fasting value (69%; Fig. 2a),

suggesting that these measures of postprandial triglyceride and glucose concentrations, but not of insulin, provide better discrimination of an individual's metabolic tolerance than fasting values do.

A key assumption when developing personalized prediction algorithms is that an individual's unique response to the same meal is reproducible. Much of the between-person phenotypic variability observed in studies examining response to diet interventions that include only a single test–response scenario could be a result of regression to the mean or other sources of error. Repeated measures (multiple measures taken within an individual at a single time point and across multiple time points) can be used to partition error from true biological variability, thereby improving the precision of the estimate. Accordingly, we administered test meals of varying macronutrient composition in duplicate per participant, under similar conditions (see Methods and Supplementary Table 2 for details). We also used CGMs, which provided sequential measures of blood glucose at 15-min intervals during the study period. Intra-individual variability (repeatability) was assessed using intra-class correlation coefficients (ICCs) for triglyceride, connecting peptide (C-peptide, a surrogate for insulin secretion) (from DBS assays) and glucose (from CGMs) measurements. The ICCs were: triglyceride<sub>6h-rise</sub> = 0.46 (95% confidence interval (CI), 0.37–0.54); glucose<sub>iAUC0-2h</sub> = 0.74 (95% CI, 0.72–0.75); C-peptide<sub>2h-rise</sub> = 0.62 (95% CI 0.54, 0.69) (Supplementary Table 3). The differences in ICCs between triglyceride, C-peptide and glucose measurements partly reflect the different assays used (DBS and CGM) (Methods).

### Predicting individual postprandial responses within a population.

We assessed the overall extent to which input variables (Supplementary Table 3) predict personal postprandial responses (Fig. 2b–d), initially using multivariable linear regression. Input variables include: (1) baseline characteristics (age, sex, clinical biochemistry (lipid, glycemic and other measures), anthropometry); (2) genetics (single-nucleotide polymorphisms

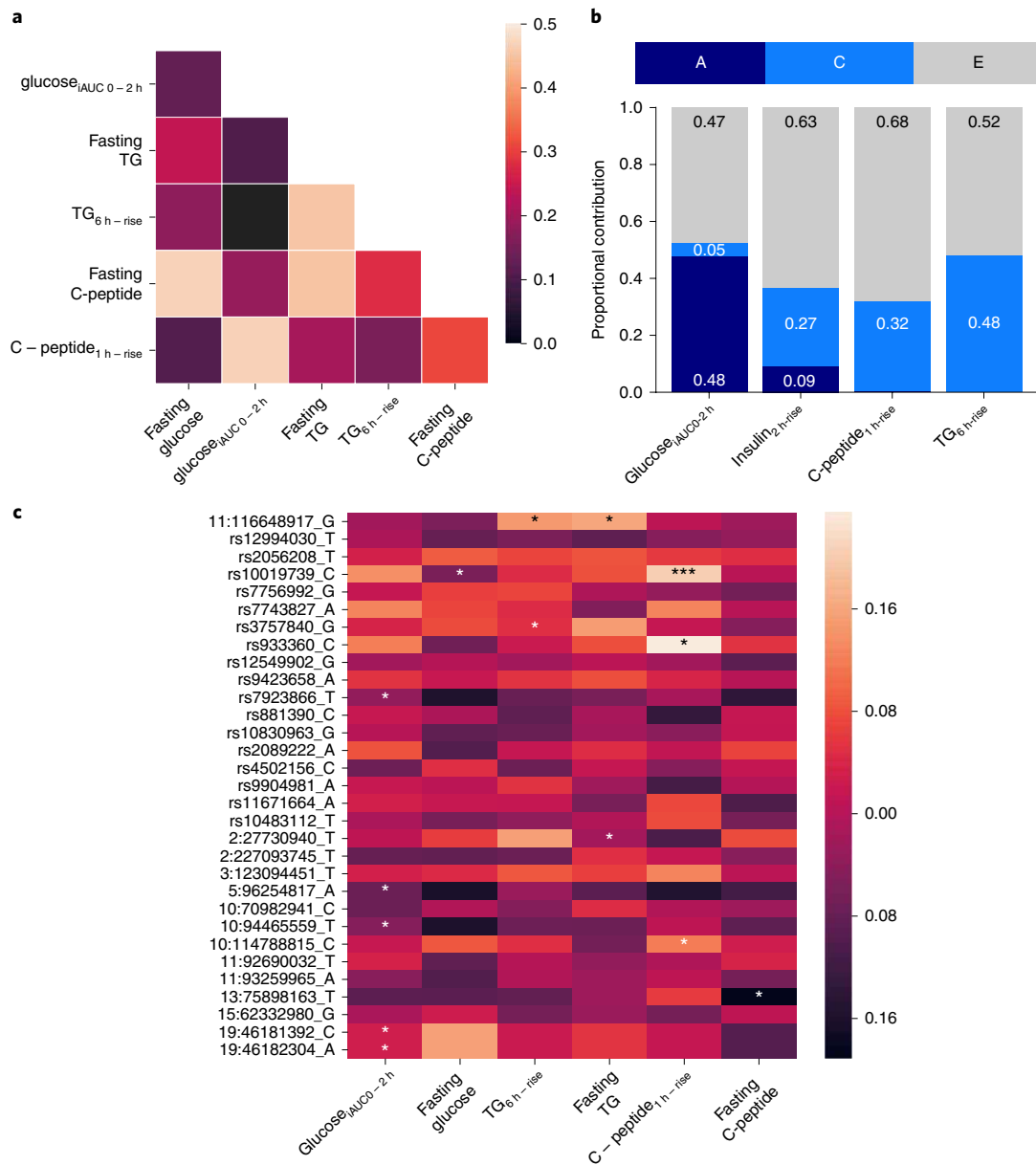


**Fig. 2 | Variation in postprandial responses. a**, Inter-individual variation in triglyceride, glucose and insulin postprandial responses to the breakfast and lunch meal challenges in the clinic ( $n=1,002$ ). **b**, Determinants of triglyceride<sub>6hr-rise</sub> measured from DBS (comparison of meals 1 and 7). **c**, Determinants of glucose<sub>iAUC0-2h</sub> measured by CGM (comparison of 7 test meals; 1, 2, 4, 5, 6, 7 and 8). **d**, Determinants of C-peptide<sub>1hr-rise</sub> measured using DBS analysis as a proxy for insulin (comparison of meals 2 and 3). Trait variations explained for each input variable are derived from separate (non-hierarchical) regression models. Values represent adjusted proportion of variance explained ( $R^2$ ), and error bars show the 95% CIs.  $R^2$  values adjusted for meal composition and meal context were derived from the following meal sample sizes: triglyceride<sub>6hr-rise</sub>  $n=712$ ; glucose<sub>iAUC0-2h</sub>  $n=9,102$ ; C-peptide<sub>1hr-rise</sub>  $n=186$ . All other determinant values were derived from the following meal sample sizes: triglyceride<sub>6hr-rise</sub>  $n=920$ ; glucose<sub>iAUC0-2h</sub>  $n=958$ ; C-peptide<sub>1hr-rise</sub>  $n=960$ . \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  using multivariable linear regression.

(SNPs)); (3) gut-microbiome features; (4) habitual diet (from Food Frequency Questionnaire (FFQ)); (5) meal context (sleep, previous meals, physical activity, meal sequence and/or timing); and (6) meal composition (energy from carbohydrates, sugar, fat, protein and fiber). Postprandial glycemic responses were determined from serum and CGM measurements in the clinic and at home (from seven standardized meals; Methods). Postprandial C-peptide and triglyceride were determined (from two standardized meals) from serum and DBS assays on material collected during the clinic and

home phases. We also tested the correlation between fasting and postprandial characteristics and found that the correlation between postprandial triglyceride and both postprandial glucose and postprandial C-peptide measures was low (Fig. 3a).

**Individual baseline characteristics.** The proportions of trait variance explained by individual baseline characteristics are shown in Fig. 2b–d for triglyceride<sub>6hr-rise</sub>, glucose<sub>iAUC0-2h</sub> and C-peptide<sub>1hr-rise</sub> (Supplementary Table 3).



**Fig. 3 | Relationship of baseline values, genetic and microbiome factors with postprandial responses.** **a**, Pearson correlations between baseline values and postprandial prediction measures of 980 participants from the UK cohort. **b**, Heritability of postprandial responses (the ACE model was fitted on log-scaled postprandial responses for triglyceride, glucose, insulin and C-peptide) in 183 monozygotic and 47 dizygotic twin pairs. In the ACE model, A represents the additive genetic component; C the shared environmental component; and E the individual environmental component. **c**, SNP associations with postprandial measures focusing on SNPs identified in published postprandial trait GWAS<sup>17-21</sup> ( $n=241$ ;  $*P < 0.05$ ,  $***P < 0.001$ , using two-sided Chi-squared test).

**Genetic factors.** The heritability of postprandial responses in the UK cohort was examined using classical twin methods (variance components analyses) to establish the upper bound of what might be predicted by directly measured genetic variation. Two-thirds of the cohort was recruited from the TwinsUK registry<sup>16</sup>, of which 230 twin pairs ( $n=460$ ; 183 monozygotic and 47 dizygotic) were studied for heritability. Additive genetic factors explained 48% of the variance in glucose<sub>AUC0-2h</sub>, whereas 0% of the variance in triglyceride<sub>6h-rise</sub> and 9% of the variance in insulin<sub>2h-rise</sub> were explained in this way (Fig. 3b). The estimated genetic variances in insulin<sub>1h-rise</sub> and C-peptide<sub>1h-rise</sub> were close to 0 (Supplementary Table 4).

**SNP-based genetic factors.** In a subgroup of participants who are part of the TwinsUK cohort, had genome-wide genotyping

previously done and had available genome-wide association study (GWAS) data ( $n=241$ ), we tested whether 32 SNPs derived from previous genome-wide scans of postprandial glucose, insulin or triglyceride concentrations<sup>17-21</sup> were associated with the postprandial variables studied here. Several SNPs were significantly ( $P < 0.05$ ) associated with these variables (Fig. 3c and Supplementary Table 4), but they collectively explained only ~9% of observed variation in glucose<sub>AUC0-2h</sub> (Fig. 2c), and less than 1% of variation for postprandial triglyceride and postprandial C-peptide (Fig. 2b,d).

**Gut microbiome (16S ribosomal RNA).** We estimated the contribution of gut-microbiome composition using relative bacterial taxonomic abundances and measures of community diversity and richness, derived from 16S rRNA high-throughput sequencing



of baseline stool specimens (Supplementary Table 4). We found that, without adjusting for any other individual characteristics, gut-microbiome composition explained 7.5% of postprandial triglyceride<sub>6h-rise</sub>, 6.4% of postprandial glucose<sub>iAUC0-2h</sub> and 5.8% of postprandial C-peptide<sub>1h-rise</sub>.

**Meal composition, habitual diet and meal context.** To determine the impact of the macronutrient composition of meals, we measured triglyceride<sub>6h-rise</sub> and C-peptide<sub>1h-rise</sub> for two standardized home-phase meals (i.e., consumed at home) of differing macronutrient compositions (for triglyceride, meals 1 and 7: 85 versus 28 g of carbohydrate and 50 versus 40 g of fat at breakfast, both followed by a lunch of 71 g carbohydrate and 22 g fat; for C-peptide, meals 2 and 3: 71 versus 41 g of carbohydrate and 22 versus 35 g of fat; Supplementary Table 2) in subsets of participants ( $n=712$  and  $n=186$ , respectively). Glucose<sub>iAUC0-2h</sub> was measured for seven standardized meals (comparison of meals 1, 2, 4, 5, 6, 7 and 8: 28–95 g carbohydrate and 0–53 g fat), totaling 9,102 meals in 920 individuals. The proportions of variance explained by meal composition, habitual diet and meal context are shown for triglyceride<sub>6h-rise</sub> in Fig. 2b, for glucose<sub>iAUC0-2h</sub> in Fig. 2c and for C-peptide<sub>1h-rise</sub> in Fig. 2d. A multivariate regression model (meals 1, 2, 4, 5, 6, 7 and 8) revealed that glucose<sub>iAUC0-2h</sub> (mmol per l per s) was significantly ( $P<0.001$ ) reduced by 79, 142 and 185 for every 1 g fat, fiber and protein, respectively, after adjustment for carbohydrate consumption.

**Machine-learning model.** To estimate the unbiased predictive utility of the analyzed factors, we used a machine-learning approach robust to overfitting<sup>22</sup>. Random forest regression models<sup>23</sup> were fitted using all informative features (meal composition, habitual diet, meal context, anthropometry, genetics, microbiome, clinical and biochemical parameters) to predict triglyceride<sub>6h-rise</sub>, glucose<sub>iAUC0-2h</sub> and C-peptide<sub>1h-rise</sub> in the UK cohort dataset. The predicted values were compared with the observed values for each trait using Pearson correlation coefficients ( $r$ ); these correlations were  $r=0.47$ ,  $r=0.77$  and  $r=0.30$  for triglyceride<sub>6h-rise</sub>, glucose<sub>iAUC0-2h</sub> and C-peptide<sub>1h-rise</sub>, respectively. Similar correlations were observed in the held-out validation set (US cohort). The model predictions for triglyceride<sub>6h-rise</sub> and glucose<sub>iAUC0-2h</sub> were  $r=0.42$  and  $r=0.75$ , respectively, but were much weaker for C-peptide<sub>1h-rise</sub> ( $r=0.14$ ) (Fig. 4). The features used to fit the models are reported in Supplementary Table 5. The repeatability and robustness of the machine-learning model are presented in Extended Data Fig. 4.

**Postprandial responses in relation to surrogate scores of clinical outcomes.** We compared the extent to which fasting and postprandial concentrations for the different biomarkers could be used to predict impaired glucose tolerance (7.8–11.0 mmol l<sup>-1</sup> 2 h after an oral glucose tolerance test (OGTT)) and atherosclerotic cardiovascular disease (ASCVD) 10-year risk score (Methods) by comparing the area under the receiver operator characteristics (ROC-AUC) curves (Fig. 5). We found that fasting triglyceride and triglyceride<sub>6h-rise</sub> contributed similarly to the ROC-AUC for ASCVD risk, and that including both was more informative than including only one of them (Fig. 5a). We also found that, although postprandial glucose was not as informative as fasting glucose, adding glucose<sub>iAUC0-2h</sub> to fasting glucose resulted in a slightly higher ROC-AUC (0.72 versus 0.69) for ASCVD 10-year risk. Fasting C-peptide and fasting glucose were as effective (ROC-AUC=0.69) as fasting triglyceride was in ASCVD prediction, whereas postprandial C-peptide (ROC-AUC=0.63) and postprandial glucose (ROC-AUC=0.62) were weaker than postprandial triglyceride (ROC-AUC=0.71). Fasting and postprandial triglyceride concentrations were weakly predictive (ROC-AUC=0.55 and 0.59, respectively) of impaired glucose tolerance (IGT), whereas fasting and postprandial C-peptide were moderately predictive (ROC-AUC=0.64 and 0.65,

respectively), although with no added predictive value in combination. We did not include here the prediction of IGT using glucose data from CGM. This is because IGT is defined solely on the basis of the blood glucose concentration at 2 h during an OGTT, which is captured by the CGM glucose recording, and so the derivation of the predictor and the clinical-score variables would be heavily dependent upon one another. Results were similar in the UK and US cohorts (Fig. 5).

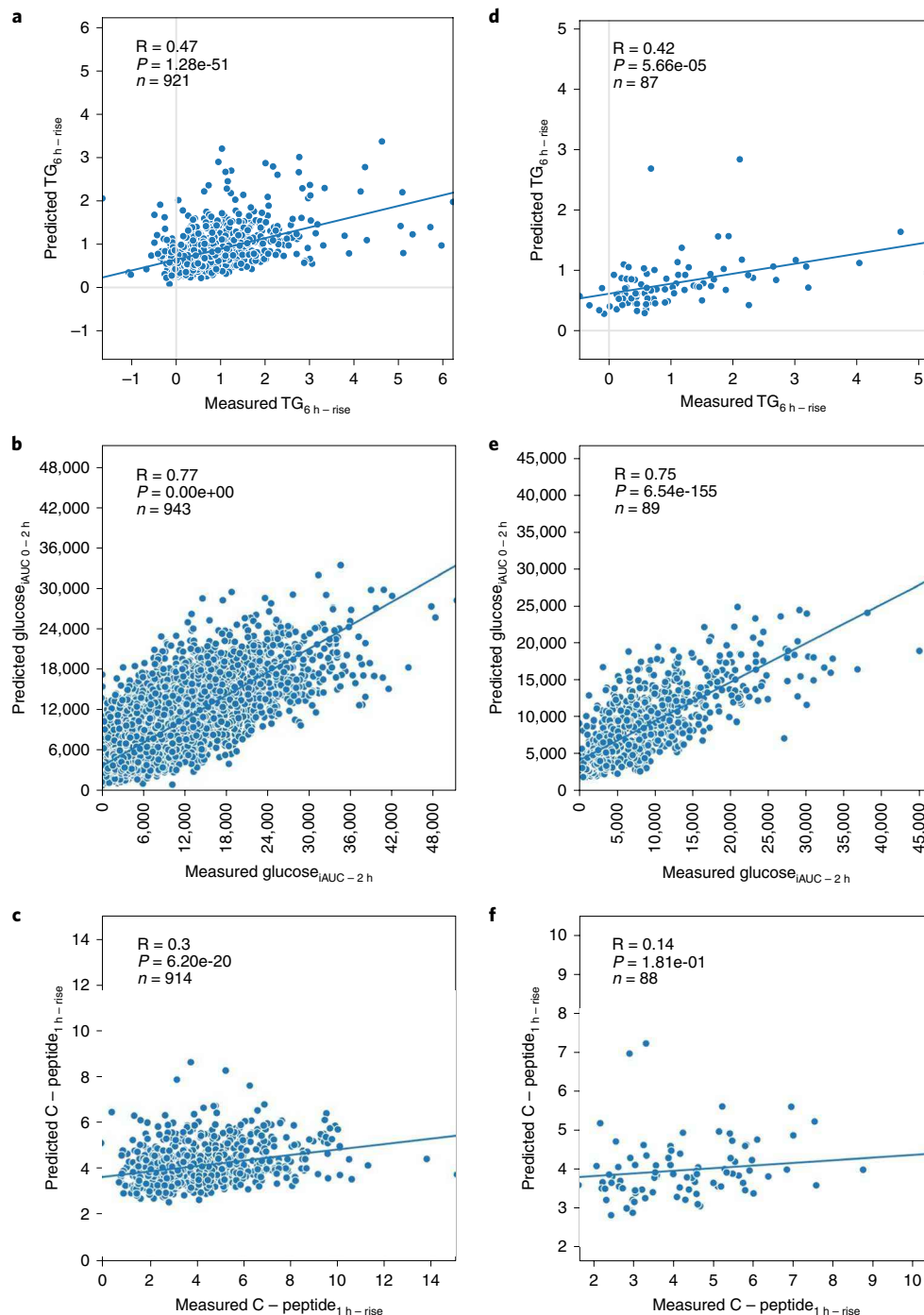
**Decoding individual responses.** Having investigated postprandial responses within the population, we next explored the responses at the individual level. We examined glycemic responses, as the granular CGM data collected during the at-home phase enabled us to assess real-world effects in detail, which was not possible for triglyceride or C-peptide. We investigated how much of an individual's postprandial response is attributable to a meal's glycemic properties, compared with how the variation results from other modifiable factors, such as meal timing, exercise and sleep.

We first examined the contribution of the meal. Although it is a widely held notion that, for an individual, variations in meal composition are primarily responsible for the variation in responses to food and that ranking of meal responses should be the same for all people<sup>24,25</sup>, we explored whether meal-specific responses that are unique to the individual exist. We ranked the order of each participant's glucose<sub>iAUC0-2h</sub> for every possible pair of standardized meals consumed at home. We then determined how frequently these rankings differed for each participant. For most pairs of meals, the ranking was the same for all individuals (for example, the glucose administered in the OGTT elicits a higher glucose<sub>iAUC0-2h</sub> than the carbohydrate in the high-fiber muffins, in all participants) (Fig. 6a). However, for select pairs of meals, the ranking was reversed in up to 48% of participants, such as between the medium-fat and -carbohydrate lunch versus the high-carbohydrate breakfast (350 of 727 participants) (meal 2 versus meal 4; Supplementary Table 2). In 186 out of 498 (37.3%) participants, discrepancies were also seen between the high-fat and the high-protein meals (meals 7 and 8). The distribution of how these meals were ranked for the participants of the PREDICT study is presented in Extended Data Fig. 2.

We note that the reordering of meal rankings could have been the result of noise. We therefore used analysis of variance (ANOVA) to estimate the effect size for the different factors explaining glycemic response (Fig. 6b), including person-specific effects (effects that vary between people, but not between meals). As described in the Methods, we considered not only the effect of the meal macronutrient and energy content in the response (meal composition), but also how each individual responded on average to all their set meals relative to the population (individual glucose scaling), as well as the effect of the individual's meal-specific response, the error attributable to the glucose measurement and other sources of variation (including modifiable sources of variation, such as sleep, circadian rhythm and exercise).

We found that, consistent with the linear models described earlier, the ANOVA models showed that there were three meal-related factors explaining individual glycemic responses. Meal macronutrient composition alters iAUC by 16.73% (CI, 15.37%–18.92%), but the individual glucose scaling is larger, altering iAUC by roughly 18.74% (17.96%–19.46%). The individual's meal-specific response is much smaller, affecting the final meal iAUC by 7.63% (6.11%–8.96%). Other modifiable sources of variation not directly related to the meal composition, such as meal timing, exercise and sleep, contributed amounts of variance similar to that of the meal's composition (Fig. 6b,c).

To investigate whether modifying the order in which meals are consumed and the time of the day affect glycemic responses, we looked at participants who ate an identical meal (meal 2) for breakfast and lunch. The average glycemic response for the same



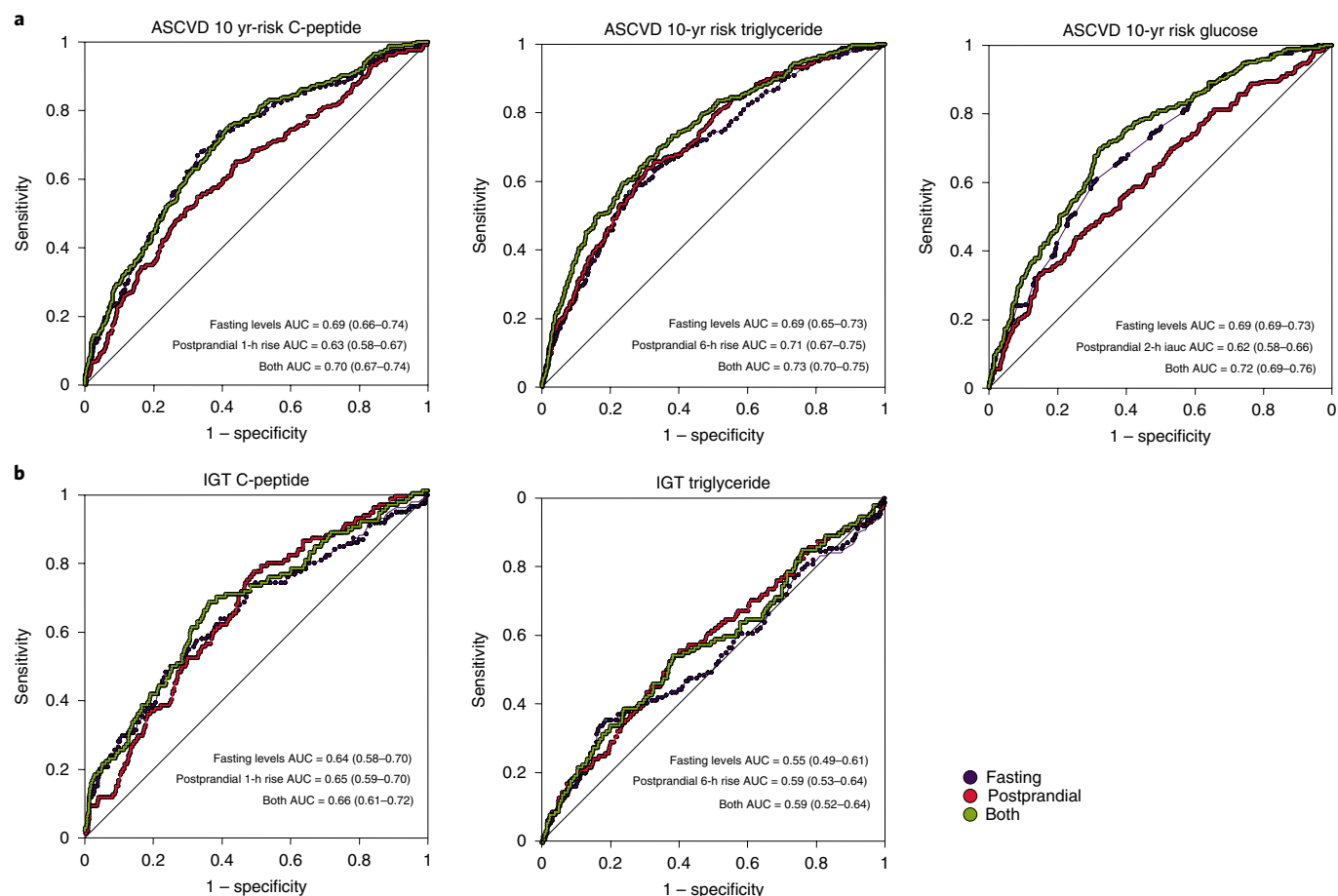
**Fig. 4 | Machine-learning models fitted to postprandial measures.** **a–c**, Results from the machine-learning model for triglyceride<sub>6h-rise</sub> (**a**), glucose<sub>iAUC0–2h</sub> (**b**) and C-peptide<sub>1h-rise</sub> (**c**) postprandial responses in the UK cohort. **d–f**, Results for the machine-learning models applied to the US validation cohort. The relationship between variables is expressed as Pearson’s correlation coefficient ( $r$ ) and denoted with a regression line;  $n$  represents the number of participants. The features used to predict each value are the same as those listed in the linear models in Fig. 2b–d.

individuals was on average twofold higher ( $t = -35.7$ , 2,721 d.f.;  $P < 0.001$ ) when the meal was ingested for lunch (mean glucose<sub>iAUC0–2h</sub> = 14,254, s.d. = 6,593) (4h following the metabolic-challenge breakfast) than when ingested for breakfast (mean glucose<sub>iAUC0–2h</sub> = 7,216, s.d. = 4,157), although with wide inter-individual variation (Fig. 6c).

## Discussion

Nutrition and health are intimately linked. Each day, people make diet-related decisions that are influenced by perceived enjoyment

and satiation, as well as health benefits and harm attributed to specific foods and beverages. Standard nutritional guidelines<sup>2–4</sup> are typically based on population averages. However, it is increasingly evident that one-size nutritional recommendations do not fit all, which is exemplified by the variable efficacy of tightly controlled lifestyle-intervention trials<sup>26–29</sup>. To address these challenges, we undertook a 2-week interventional trial, including a tightly controlled in-clinic day and a 2-week at-home phase, in which postprandial metabolic responses to a series of standardized meals were obtained in more than 1,000 healthy adults from the



**Fig. 5 | Associations between fasting and postprandial values for triglyceride, C-peptide and glucose concentrations with clinical measures in the UK cohort. a,b**, ROC curves illustrating the predictive utility of fasting and postprandial triglyceride, glucose and C-peptide measures to distinguish the bottom 70% from the top 30% of the cohort (cut-off ASCVD 10-year risk, 0.0183) for ASCVD 10-year risk ( $n=951$  independent samples from the UK cohort) (**a**) and impaired glucose tolerance (IGT) ( $n=826$  independent samples from the UK cohort) (**b**). The same analyses were performed in the US cohort ( $n=92$  independent samples) resulting in ROC-AUC (with 95% CI) values for ASCVD 10-year risk of: C-peptide fasting AUC = 0.68 (0.56–0.80), postprandial AUC = 0.66 (0.54–0.77), both AUC = 0.69 (0.58–0.81); triglyceride fasting AUC = 0.73 (0.63–0.84), postprandial AUC = 0.75 (0.65–0.85), both AUC = 0.77 (0.67–0.88); and glucose fasting AUC = 0.74 (0.63–0.85), postprandial AUC = 0.64 (0.52–0.76), both AUC = 0.76 (0.64–0.85). For IGT, values were: C-peptide fasting AUC = 0.66 (0.53–0.80), postprandial AUC = 0.59 (0.46–0.72), both AUC = 0.67 (0.54–0.80); and triglyceride fasting AUC = 0.66 (0.53–0.80), postprandial AUC = 0.59 (0.46–0.72), both AUC = 0.61 (0.54–0.80).

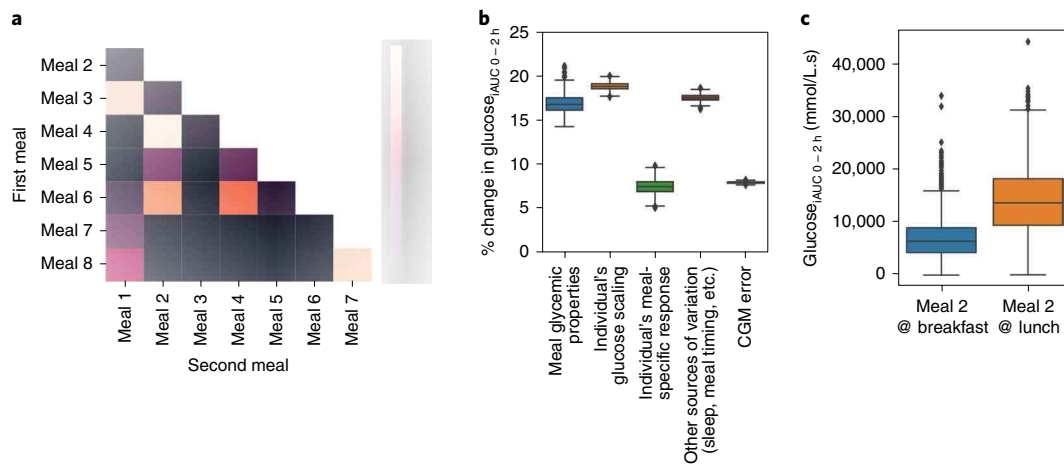
United Kingdom and United States. The primary aim was to derive algorithms that predict an individual's postprandial metabolic responses to specific foods. The core outcomes were variations in blood concentrations of triglyceride, glucose and insulin (or C-peptide), as these biomarkers work in concert to affect cardio-metabolic risk<sup>8,30</sup>.

In many cases, we observed responses that contrast with those reported in traditional clinic-based studies, thereby reshaping conclusions about the key factors influencing responses to foods. For example, genetic influence was lower than expected, especially for triglyceride, whereas modifiable factors such as meal timing conveyed larger effects than anticipated.

Meal composition has large effects on postprandial insulinemic and lipemic response<sup>31</sup>. Some small studies have suggested that meals with high-fat and/or high-protein content elicit very different postprandial responses than lower-fat and/or lower-protein meals with identical carbohydrate content (reviewed in ref. <sup>31</sup>). The type of fat in a meal also alters the lipemic response<sup>32</sup>. However, measuring postprandial triglyceride and C-peptide at home in large cohorts is both logistically challenging and places a considerable burden on the participants. Thus, for pragmatic reasons, only two pairs of

meals (high fat and high carbohydrate) were used to calculate postprandial triglyceride and C-peptide responses, and the difference in macronutrient content of these meals was low. This limited number of different meals and their relatively similar macronutrient content might explain why the effects seen for postprandial triglyceride and C-peptide were lower than expected.

In addition to fasting concentrations of triglyceride and glucose, we found that postprandial triglyceride and glucose concentrations were informative for determination of IGT and CVD risk. However, postprandial C-peptide measurements provided no additional information over fasting concentrations. We found that, although postprandial triglyceride and glucose responses were highly variable between individuals, a person's response to the same meals was often similar and therefore predictable. Any given individual generally responds comparably to different meals of the same macronutrient profile. Some people experience large postprandial excursions across most meals, whereas others consistently experience modest responses. This is important for individualized prediction and recommendations, as it suggests that once an individual's postprandial response to specific foods is known, their response to other foods could be inferred.



**Fig. 6 | Person-specific diversity in postprandial response.** **a**, Proportion of times in the PREDICT 1 study that the ranking of the glycemic response (glucose<sub>iAUC0-2h</sub>) to pairs of set meals was altered ( $n = 828$ , UK cohort). **b**, Effect size for factors explaining glycemic response. The different sources of variation were estimated using two-way ANOVA, as described in Supplementary Table 3 ( $n = 483$  individuals). **c**, Time-of-day effects on glycemia following identical meals. ( $n = 920$ , UK cohort). Boxes show quartiles (25th, 50th, 75th percentiles), and whiskers show the 95% interval.

We show not only that a person's glycemic response is the result of glucose scaling specific to the individual, which determines whether a person is a high or low responder to all meals, but also that there are meal-specific responses unique to an individual. Possible explanations include individual genetic differences in the ability to digest high-starch meals<sup>33</sup>. Zeevi et al.<sup>5</sup> reported an example in which one participant had an exaggerated glycemic response to a banana but not to a cookie, whereas the second participant had the opposite response. We assessed this phenomenon in our data and found that individual glucose scaling and meal-specific responses both exist, but individual meal-specific responses are generally much more effective than scaling.

People differ greatly in their responses to diet interventions. The DIETFITS study, for example, randomized 609 people to either a healthy low-fat or a healthy low-carbohydrate diet for 12 months<sup>34</sup>. By the end of the study, average weight loss was similar between groups (~5–6 kg), but wide variations were seen within groups (–30 kg to +10 kg). Elsewhere, the Diabetes Prevention Program showed that, although a standardized intensive lifestyle intervention focusing on changes in diet (tailored only to the energy requirements of the individual) lowered diabetes risk substantially<sup>28</sup>, its efficacy varied greatly across the study population<sup>26,27</sup> and was determined to some extent by genetic factors<sup>29</sup>. Although the response to diet interventions will depend partly on adherence, findings from the PREDICT trial and elsewhere<sup>35,36</sup> suggest that, even in highly adherent participants, substantial variations in response exist, which might be predictable. In PREDICT, non-food-specific factors (for example, meal timing, sleep and activity) were highly informative of these person-specific responses.

Previous large-scale studies of postprandial responses have focused solely on glycemic outcomes because assessing postprandial triglyceride and insulin concentrations in free-living conditions is challenging<sup>2,25</sup>. Here, we assessed glycemic responses with CGMs, and also triglyceride and C-peptide concentrations during the at-home period using a validated DBS method and support from a specifically designed mobile app (Methods). The low correlation between triglyceride and glucose suggests that prediction algorithms relying solely on glucose would be insufficient for the detection of dysregulated triglyceride responses.

The prediction algorithms we developed are likely to have been strengthened by the use of randomized, mixed meals that contained combinations of macronutrients reflecting those seen in real-world

settings, rather than supraphysiological lipid or carbohydrate challenges as have been used in previous studies.

In general, genetics, contrary to our expectations, was not a predominant determinant of most of these responses; we found that the heritable fraction (the trait variance explained by additive genetic factors) was low for postprandial triglyceride (6-h rise, 0.0%), as well as C-peptide and/or insulin concentrations at 1 h (0.3%) and at 2 h (9.1%). The heritable fractions for postprandial glucose (2-h iAUC) responses were considerably higher (48%). Despite the wealth of publicly available SNP data ([www.type2diabetesgenetics.org](http://www.type2diabetesgenetics.org)), there are no robust data for these specific postprandial traits, as almost all published GWASs of serological traits have focused on fasting values. Nevertheless, in exploratory analyses, we examined the predictive value of loci previously linked to post-challenge triglyceride, glucose or insulin concentrations<sup>17–21</sup>, but found that the predictive utility of these variants was poor, particularly for triglyceride and C-peptide (Fig. 3c). The modest heritability of postprandial traits means that, even in an unrealistically optimistic scenario in which most of this trait variance is explained by known DNA variants, it is unlikely that prediction algorithms using DNA variant data alone, which many direct-to-consumer nutrigenomics companies advocate, would succeed.

The lack of a major genetic component to these traits highlights the likely involvement of modifiable environmental exposures. Indeed, we found that meal composition and context (for example, meal timing, exercise, sleep and circadian rhythm) were core determinants of postprandial metabolism. These predictions were strengthened using data on gut-microbiome diversity. Using machine learning that combined all relevant data, an individual's postprandial triglyceride and glycemic responses could be meaningfully predicted, with similar results in the US validation cohort. For C-peptide, the prediction was much weaker in the validation cohort ( $r = 0.30$  UK,  $r = 0.14$  US), possibly reflecting the lower number of test meals relative to the number of input variables, which could adversely affect the reliability of the prediction<sup>37</sup>. The postprandial glycemic predictions were similar to those reported by Zeevi et al.<sup>5</sup>, although the analysis methods and input features are not directly comparable.

Despite having developed these prediction algorithms, there is scope for improvement, such as the inclusion of a more diverse array of meal interventions, and more detailed assessments of contextual factors than were used in the current study. Technological advances



could also help to improve predictions. For example, although glucose can be continuously assessed with CGMs, no commercially available devices suitable for free-living assessments of continuous insulin and triglyceride concentrations currently exist. Moreover, owing to the differences in tolerability and the lower limit of detectable responses of dietary carbohydrates compared with fats<sup>38</sup>, our trial suggests that the prediction of postprandial glucose is methodologically superior to that for triglyceride responses (Fig. 2b–d). Difficulties in directly comparing changes in triglyceride and glucose were a limitation of our study. Continuous, accurate measures of these traits could substantially improve predictions owing to reductions in model error and the ability to study non-linear patterns of response, which may be important. The inclusion of deep ‘-omics’ data may further enhance the predictive ability of these algorithms; for example, here we used microbiome data derived from 16S rRNA sequencing, which were valuable for prediction (explaining 6.4% and 7.5% of the variances for glucose and triglyceride responses, respectively), but data may be even more informative if derived from higher-resolution metagenomic sequencing. The nutritional signatures detectable within the metabolome, both in blood<sup>39</sup> and feces<sup>40</sup>, suggest that including a larger metabolomics panel—and, quite probably, other -omics data, for example meta-transcriptomics, transcriptomics or proteomics—in our algorithms would add costs but also enhance predictions. Using FFQs, we found that habitual diet explains a small proportion (<2%) of an individual’s postprandial responses. However, FFQs have well-known limitations, and other objective approaches may be considerably less biased and less error prone<sup>27</sup>. Pairing this with short-term assessments, such as the weighed dietary record included in the PREDICT study app, may help mitigate these limitations. More comprehensive challenge tests might also reveal new aspects of postprandial metabolism; here, we used a 6-h test meal challenge, as this was deemed the maximum duration that most participants were likely to accept. Data from challenge tests of longer durations (up to 8 h) may provide valuable information on both glucose and triglyceride responses.

For postprandial triglyceride and glucose responses, the prediction models derived in the UK cohort performed almost as well in the independent US validation cohort, which is reassuring given the differences in environmental factors; nevertheless, both cohorts comprised younger healthy adults of European ancestry. Thus, the generalization of our findings would require validation in people of non-European ancestry, older adults and people with diseases that affect metabolism, such as diabetes. The clinical implications of our predictions will require appropriately powered longitudinal studies.

In conclusion, this is the most comprehensive assessment to date of metabolic responses to nutritional challenges in a rigorous intervention setting. We observed considerable inter-individual differences in postprandial metabolic responses to the same meals, challenging the logic of standardized diet recommendations. These findings, in addition to the scalability of the assessment methods and the accuracy of the prediction algorithms described here, mean that, at least from a cardiometabolic health perspective, population-wide personalized nutrition has potential as a strategy for disease prevention.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-020-0934-0>.

Received: 28 October 2019; Accepted: 11 May 2020;  
Published online: 11 June 2020

### References

- GBDD 2017 Diet Collaborators. 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).
- US Department of Health and Human Services & US Department of Agriculture. *2015–2020 Dietary Guidelines for Americans* 8th edn (2015).
- Karpyn, A. *Food and Public Health: A Practical Introduction* (Oxford University Press, 2018).
- European Food Safety Authority. Dietary reference values for nutrients summary report. *EFSA Supporting Publications* **14**, e15121E (2017).
- Zeevi, D. et al. Personalized nutrition by prediction of glycemic responses. *Cell* **163**, 1079–1094 (2015).
- Tebani, A. & Bekri, S. Paving the way to precision nutrition through metabolomics. *Front. Nutr.* **6**, 41 (2019).
- Kolovou, G. D. et al. Assessment and clinical relevance of non-fasting and postprandial triglycerides: an expert panel statement. *Curr. Vasc. Pharmacol.* **9**, 258–270 (2011).
- Astley, C. M. et al. Genetic evidence that carbohydrate-stimulated insulin secretion leads to obesity. *Clin. Chem.* **64**, 192–200 (2018).
- Blaak, E. E. et al. Impact of postprandial glycaemia on health and prevention of disease. *Obes. Rev.* **13**, 923–984 (2012).
- DECODE Study Group, European Diabetes Epidemiology Group. Glucose tolerance and cardiovascular mortality: comparison of fasting and 2-hour diagnostic criteria. *Arch. Intern. Med.* **161**, 397–405 (2001).
- Ning, F. et al. Cardiovascular disease mortality in Europeans in relation to fasting and 2-h plasma glucose levels within a normoglycemic range. *Diabetes Care* **33**, 2211–2216 (2010).
- Bansal, S. et al. Fasting compared with nonfasting triglycerides and risk of cardiovascular events in women. *JAMA* **298**, 309–316 (2007).
- Lindman, A. S., Veierod, M. B., Tverdal, A., Pedersen, J. I. & Selmer, R. Nonfasting triglycerides and risk of cardiovascular death in men and women from the Norwegian Counties Study. *Eur. J. Epidemiol.* **25**, 789–798 (2010).
- Jackson, K. G., Poppitt, S. D. & Minihane, A. M. Postprandial lipemia and cardiovascular disease risk: Interrelationships between dietary, physiological and genetic determinants. *Atherosclerosis* **220**, 22–33 (2012).
- Stumvoll, M., Goldstein, B. J. & van Haeften, T. W. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet* **365**, 1333–1346 (2005).
- Moayyeri, A., Hammond, C. J., Hart, D. J. & Spector, T. D. The UK Adult Twin Registry (TwinsUK Resource). *Twin Res. Hum. Genet.* **16**, 144–149 (2013).
- Prokopenko, I. et al. A central role for GRB10 in regulation of islet function in man. *PLoS Genet.* **10**, e1004235 (2014).
- Saxena, R. et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. *Nat. Genet.* **42**, 142–148 (2010).
- Wojczynski, M. K. et al. Genome-wide association study of triglyceride response to a high-fat meal among participants of the NHLBI Genetics of Lipid Lowering Drugs and Diet Network (GOLDN). *Metabolism* **64**, 1359–1371 (2015).
- Sanna, S. et al. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. *Nat. Genet.* **51**, 600–605 (2019).
- Almgren, P. et al. Genetic determinants of circulating GIP and GLP-1 concentrations. *JCI Insight* **2**, e93306 (2017).
- Wojcikowski, M., Ballester, P. J. & Siedlecki, P. Performance of machine-learning scoring functions in structure-based virtual screening. *Sci. Rep.* **7**, 46710 (2017).
- Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer New York, 2013).
- Russell, W. R. et al. Impact of diet composition on blood glucose regulation. *Crit. Rev. Food Sci. Nutr.* **56**, 541–590 (2016).
- Jeong, Y. et al. A review of recent evidence from meal-based diet interventions and clinical biomarkers for improvement of glucose regulation. *Prev. Nutr. Food Sci.* **25**, 9–24 (2020).
- Delahanty, L. M. et al. Effects of weight loss, weight cycling, and weight loss maintenance on diabetes incidence and change in cardiometabolic traits in the Diabetes Prevention Program. *Diabetes Care* **37**, 2738–2745 (2014).
- Diabetes Prevention Program Research Group. et al. 10-year follow-up of diabetes incidence and weight loss in the diabetes prevention program outcomes study. *Lancet* **374**, 1677–1686 (2009).
- Knowler, W. C. et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N. Engl. J. Med.* **346**, 393–403 (2002).
- Papandonatos, G. D. et al. Genetic predisposition to weight loss and regain with lifestyle intervention: analyses from the diabetes prevention program and the look AHEAD randomized controlled trials. *Diabetes* **64**, 4312–4321 (2015).
- Taylor, R. Banting Memorial Lecture 2012: reversing the twin cycles of type 2 diabetes. *Diabet. Med.* **30**, 267–275 (2013).

31. Bell, K. J. et al. Impact of fat, protein, and glycemic index on postprandial glucose control in type 1 diabetes: implications for intensive diabetes management in the continuous glucose monitoring era. *Diabetes Care* **38**, 1008–1015 (2015).
  32. Hansson, P. et al. Meals with similar fat content from different dairy products induce different postprandial triglyceride responses in healthy adults: a randomized controlled cross-over trial. *J. Nutr.* **149**, 422–431 (2019).
  33. Atkinson, F. S., Hancock, D., Petocz, P. & Brand-Miller, J. C. The physiologic and phenotypic significance of variation in human amylase gene copy number. *Am. J. Clin. Nutr.* **108**, 737–748 (2018).
  34. Gardner, C. D. et al. Effect of low-fat vs low-carbohydrate diet on 12-month weight loss in overweight adults and the association with genotype pattern or insulin secretion: The DIETFITS randomized clinical trial. *JAMA* **319**, 667–679 (2018).
  35. Viitasalo, A. et al. Genetic predisposition to higher body fat yet lower cardiometabolic risk in children and adolescents. *Int. J. Obes. (Lond.)* **43**, 2007–2016 (2019).
  36. Hughes, R. L., Kable, M. E., Marco, M. & Keim, N. L. The role of the gut microbiome in predicting response to diet and the development of precision nutrition models. Part II: results. *Adv. Nutr.* **10**, 979–998 (2019).
  37. Lee, J. G. et al. Deep learning in medical imaging: general overview. *Korean J. Radiol.* **18**, 570–584 (2017).
  38. Dubois, C. et al. Effects of graded amounts (0–50 g) of dietary fat on postprandial lipemia and lipoproteins in normolipidemic adults. *Am. J. Clin. Nutr.* **67**, 31–38 (1998).
  39. Garcia-Perez, I. et al. Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. *Lancet Diabetes Endocrinol.* **5**, 184–195 (2017).
  40. Zierer, J. et al. The fecal metabolome as a functional readout of the gut microbiome. *Nat. Genet.* **50**, 790–795 (2018).
- Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
- © The Author(s), under exclusive licence to Springer Nature America, Inc. 2020

## Methods

**Study population, study design, recruitment criteria, meal challenges and Zoe study app.** *Study population.* The PREDICT 1 study was a multinational study conducted between 5 June 2018 and 8 May 2019. The primary cohort was recruited at St Thomas' Hospital in London, UK, and a validation cohort (that underwent the same profiling as the UK cohort) was assessed at the Massachusetts General Hospital (MGH) in Boston, Massachusetts, as described in the online protocol<sup>41</sup>. In the United Kingdom, participants (target enrollment, 1,000 participants) were recruited from the TwinsUK cohort, an ongoing research cohort described elsewhere<sup>16</sup> and through online advertising (Extended Data Fig. 1a). In the United States, participants (target enrollment, 100 participants) were recruited through online advertising, research-participant databases and Rally for Research (<https://rally.partners.org/>), an online recruiting portal for research trials (Extended Data Fig. 1b). Ethical approval for the study was obtained in the United Kingdom from the Research Ethics Committee and Integrated Research Application System (IRAS 236407), and in the United States from the institutional review board (Partners Healthcare IRB 2018P002078). The trial was registered on ClinicalTrials.gov (registration number: NCT03479866) as part of the registration for the PREDICT program of research, which also includes two other study protocol cohorts. The trial was run in accordance with the Declaration of Helsinki and Good Clinical Practice.

Study participants were healthy individuals aged 18–65 years who were able to provide written informed consent. Criteria used to assess eligibility are listed in Supplementary Table 1. Exclusion criteria included ongoing inflammatory disease; cancer in the last three years (excluding skin cancer); long-term gastrointestinal disorders including irritable bowel disease or Celiac disease (gluten allergy), but not including irritable bowel syndrome; taking immunosuppressants or antibiotics as daily medication within the last three months; capillary glucose level of  $>12\text{ mmol l}^{-1}$  (or  $216\text{ mg dl}^{-1}$ ), or type 1 diabetes mellitus, or taking medication for type 2 diabetes mellitus; currently experiencing acute clinically diagnosed depression; heart attack (myocardial infarction) or stroke in the last 6 months; pregnant; and vegan or experiencing an eating disorder or unwilling to consume foods that are part of the study.

*Study design.* For the study, 1,002 generally healthy adults from the United Kingdom (including non-twins, monozygotic twins and dizygotic twins) and 100 healthy adults from the United States (non-twins; validation cohort) were enrolled and completed baseline clinic measurements. Key outcomes included postprandial metabolic responses (0–6 h; blood triglyceride, glucose and insulin concentrations) to sequential mixed-nutrient dietary challenges (containing 86 g carbohydrate and 53 g fat at 0 h; 71 g carbohydrate and 22 g fat at 4 h) administered in a tightly controlled clinical setting on day 1 (Fig. 1). A second set of outcomes was assessed over the subsequent 13 d at-home period. Lipemic and C-peptide responses (as a surrogate for insulin) to two standard meals differing in fat and carbohydrate composition were assessed at home using DBS assays collected at three postprandial time points. Glycemic responses to eight meals (seven in duplicate) of different macronutrient (fat, carbohydrate, protein and fiber) content were assessed using CGMs. In addition, participants wore physical-activity and sleep monitors for the duration of the study and provided stool samples for microbiome profiling.

We selected specific time points and increments to analyze triglyceride, glucose, insulin and C-peptide data to reflect the different pathophysiological processes for each measure. To monitor compliance, all test meals consumed by participants were logged in the Zoe study app (with an accompanying picture) and reviewed in real time by the study nutritionists. Only test meals that were consumed according to the standardized meal protocol were included in the analysis.

**Baseline clinic visit (day 1).** Participants in the United Kingdom 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<sup>42</sup>) and food-frequency questionnaire (European Prospective Investigation into Cancer and Nutrition (EPIC) Food-Frequency Questionnaire (FFQ)<sup>43</sup>). In the United States cohort, 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 instrument, was substituted for the EPIC FFQ. Stool collection and questionnaires were completed at home and returned to study staff at the baseline visit. Participants were asked to refrain from exercise and to limit fat, fiber and alcohol intake for 24 h beforehand, and to abstain from caffeine from 18:00 the night before the baseline visit. Participants arrived at 8:30 for their visit, having fasted from 21:00 the night before, and were cannulated in the forearm (antecubital vein) to collect a fasted blood sample before they were fitted with wearable devices (CGM (Freestyle Libre Pro) and wrist-based triaxial accelerometer (AX3, Axivity)). Heart rate and blood pressure were measured (in triplicate, with the mean of the second and third measurements recorded) using an automated blood-pressure monitor while participants were fasted. Participant weight, height and hip and waist circumference were measured using standard clinical techniques. Fasting blood glucose level was checked using HemoCue Glucose 201 + System (Radiometer) or Stat Strip (Nova Biomedical) in the United Kingdom and United States, respectively.

Following the baseline blood draw, participants consumed a breakfast (muffins and milkshake at 0 min) and lunch (muffins at 240 min) test meal (Supplementary

Table 2); each was to be consumed within 10 min. Additional venous blood was collected via cannula at 15, 30, 60, 120, 180, 240, 270, 300 and 360 min. Participants had access to water to sip throughout the visit. Between blood sampling, participants were trained in how to complete the study at home, including when and how to consume standardized test meals, perform DBS and use the Zoe study app. Upon completing their baseline visit, participants received all the components necessary to complete the home phase.

**Home phase (days 2–14).** During the home phase of the study, participants consumed multiple standardized test meals for breakfast and lunch over a 9- to 11-d period, while wearing the CGM and accelerometer. Meals differed in macronutrient composition (carbohydrate, fat, protein and fiber). Participants recorded all of their dietary intake and exercise on the Zoe study app throughout the study. DBS tests were completed 4 d before and after test meals, as outlined in the online protocol<sup>41</sup>. Following completion of the home phase, participants returned all study samples and devices to study staff via standard mail.

**Test-meal preparation, nutrient composition and timing, and standardized participant test-meal instructions.** Upon completion of their baseline visit, participants received a home-phase meal pack containing test-meal components (for nutrient composition, see Supplementary Table 2), which they consumed according to standardized instructions for breakfast and, on some days, lunch. Test meals consisted of either an OGTT (on 2 d) or muffins, which were consumed on their own or paired with chocolate milk, a protein shake or commercial fiber bars and were consumed in a different order depending on which protocol group (1–3) they were assigned to, as described in Supplementary Table 2. Meal order for the three protocol groups was randomized using Microsoft Access for each participant, using a two-block randomization and one non-randomized block.

Participants were instructed to fast for a minimum of 8 h prior to consuming a test breakfast meal, and to fast for 3 or 4 h after meal consumption (depending on the test meal; in protocol 1, the fasting period was 3 h for meal 5, and 4 h for all other meals; in protocols 2 and 3, the fasting period was 3 h for all breakfast meals, excluding combinations of breakfast and lunch, for which fasting periods were 4 h and 2 h, respectively). They were advised to limit exercise and drink only plain, still water during fasting periods. When fasting was completed, participants could eat, drink and exercise as they liked for the rest of the day. Participants were asked to consume all muffin-based meals within 10 min and the OGTT within 5 min, and to notify study staff if this was not achieved, in which case the data were excluded from analysis. If the participant chose to accompany their home-phase muffin-based test meals with a tea or coffee (with up to 40 ml of 0.1% fat cow's milk, but no sugar or sweeteners), they were instructed to consume this drink consistently, in the same strength and amount, alongside all muffin-based test meals throughout the study. Participants were instructed to not consume any food or drink other than water alongside the OGTT, and to avoid physical activity during the 3-h fasting period that followed it.

Participants recorded test meals and any dietary intake consumed within fasting periods, including accompanying drinks in the Zoe study app with the exact time at consumption and ingredient quantities so that study staff could monitor compliance. Only test meals that were completed according to instructions were included in analysis.

Test meals were prepared and packaged in the Dietetics Kitchen (Department of Nutritional Sciences, King's College London) 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) and commercially available fiber bars (Chocolate Fudge Brownie, Fiber One, General Mills; Goodness Bar Apple & Walnut, The Food Doctor). Test meals were shipped frozen, under temperature-controlled conditions, to the United States to limit variability in the intervention. Participants were instructed to freeze their muffins at home and defrost each set of muffins in the refrigerator the night before they were consumed. Test-meal drinks were prepared by the participant at home by mixing pre-portioned powder sachets with long-life milk provided to them (meal 1, 220 ml 0.1% fat milk; meal 8, 200 ml 1.6% fat milk). Powder sachets and fiber bars were stored at room temperature until consumption. The OGTT (meal 5) consisted of a pre-portioned powdered glucose sachet, which participants mixed with 300 ml water in the United Kingdom. US participants were provided with pre-mixed OGTTs ready for consumption (cat. no. 82028-512, VRW).

**Zoe study app and dietary-assessment methodology.** 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 assessments were due, and asked participants to report their hunger and alertness levels on visual analog scales truncated from Flint et al.<sup>44</sup>. Participants were asked to log in the app any exercise which would not be well captured by a wrist-affixed accelerometer, such as cycling. Participants logged their full dietary intake using the app over the 14-day study period, including all standardized test meals and free-living (i.e., consumed during their free time) foods, beverages (including water) and medications. Data logged in the app were uploaded onto a digital

dashboard in real time and were 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 by using photographs, product barcodes, product-specific portion sizes and digital scales. Study nutritionists also reviewed food-logging data by comparing the photographs uploaded by subjects with the items they logged on the app. Any uncertainties were clarified actively with the participant through the app's messaging system or via phone while the participant was on the study.

**Protocol versions and amendments.** Protocol amendments for the PREDICT study, following commencement of the study and participant enrollment, were as follows: the first amendment (approved by UK IRAS on 1 August 2018) allowed additional test meals to be included in the home phase and participants' logging of transit time through the gut by using a metabolic-challenge breakfast (meal 1) on the clinic day that was dyed blue with food coloring. The DBS protocol was also changed according to physiological peaks in biomarkers (triglyceride or C-peptide). Starting on 28 August 2018, triglyceride was measured on days 2–3 during fasting, 300 and 360 min postprandially, and C-peptide was quantified on days 4–5 during fasting, 30 and 120 min postprandially, as described for protocol group 2. A second saliva sample collection was added on the clinic day, at 30 min after the metabolic-challenge breakfast, to measure salivary amylase production postprandially and to provide a comparison to fasted amylase levels. The second amendment (approved by UK IRAS on 2 September 2018) was a change to the lower body-mass index limit for eligibility to 16.5 kg per m<sup>2</sup> (originally 20 kg per m<sup>2</sup>). Minor meal changes were made, not requiring ethical approval, which resulted in protocol group 3 (implemented in January 2019). In the US cohort, on 3 January 2019, the IRB approved an amendment (PREDICT-US v2.0) to address meal changes introduced in the United Kingdom 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 United States.

**Outcome variables and sample collection, handling and analysis.** *DBS collection, method validation and analysis.* DBS collection. Triglyceride and C-peptide were quantified from DBS tests completed by participants at the baseline visit (at fasted baseline and 300 min post-breakfast; for method validation) and on the first 4 d of the home phase while consuming test meals (test timings and associated meals are outlined in the online protocol<sup>41</sup>).

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

*DBS method validation.* DBS C-peptide and triglyceride concentrations were validated during PREDICT, against venous serum concentrations collected during the baseline clinic visit at 0 and 300 min after breakfast test meals. Correlations between the two methods were found to be high: for triglyceride (1,772 pairs), Pearson's  $r = 0.94$ ; for C-peptide (1,679 pairs), Pearson's  $r = 0.91$ .

*Quantification of total triglyceride from DBS.* From the DBS sample, 2 punches were taken and transferred into a high-performance liquid chromatography (HPLC) vial, and lipids were extracted with methanol at 600 r.p.m. and 25 °C for 3 h. The resulting extract was processed with a triglyceride kit (FUJIFILM Wako Chemicals) at 600 rpm and 37 °C for 2.5 h, and the reaction products were subsequently analyzed by HPLC–ultraviolet. HPLC was performed with a HP 1260/1290 infinity liquid chromatograph (Agilent Technologies) using UV detection. The analyte was separated from matrix components on a 4.6 mm × 100 mm reversed-phase column at 40 °C. A one-point calibration curve was made from analysis of triglyceride standard after enzymatic reaction with the kit. The analytical method is linear from 0.5–6 mmol l<sup>-1</sup> with a quantification limit of 0.3 mmol l<sup>-1</sup>.

*Quantification of C-peptide from DBS.* C-peptide in DBSs was assayed using a Mercodia solid-phase two-site enzyme immunoassay (enzyme-linked immunosorbent assay (ELISA); Mercodia AB). Three spots were punched into the kit plate with anti-C-peptide antibodies bound to the well. Assay buffers were added, and C-peptide was 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 the addition of acid to give a colorimetric endpoint that was read spectrophotometrically at 450 nm.

*Stool-sample collection, method validation and microbial analysis.* *Stool-sample collection.* Participants collected a stool sample at home prior to their clinical visit. Samples were collected using the EasySampler collection kit (ALPCO), and went

into fecal collection tubes containing DNA/RNA Shield buffer (Zymo Research). Upon receipt at the laboratory, samples were homogenized, aliquoted and stored at −80 °C in Qiagen PowerBeads 1.5-ml tubes (Qiagen). 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) (data not shown).

*Microbiome 16S rRNA gene sequencing and analysis.* DNA was isolated by Qiagen Genomic Services using DNeasy 96 PowerSoil Pro. Optical density measurement was performed using Spectrophotometer Quantification (Tecan Infinite 200). The V4 hypervariable region of the 16S rRNA gene was then amplified at Genomescan. 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 360,000 reads per sample. Rarefied reads were analyzed using the DADA2 pipeline<sup>45</sup>. Quality control of the reads was performed using the 'filterAndTrim' function from the DADA2 package, truncating eight nucleotides from each read to remove barcodes, discarding all reads with quality less than 20, discarding all reads with at least one N and removing the phiX Illumina spike-in. Only paired-end reads with at least 120 bp and with an expected DADA2 error less than 4 were retained for downstream analyses. Error rates were inferred from the cleaned set of reads ('learnErrors' function) and used in the DADA2 algorithm ('mergePairs' function) for merging the reads, after dereplication ('derepFastq' function). Merged reads were further processed, and only reads within 280 and 290 bp were retained, representing the majority of the distribution of the lengths. Reads were further processed to remove chimeras using the 'removeBimeraDenovo' function with a consensus method. Finally, taxonomy was assigned using the SILVA database (version 132) using the 'assignTaxonomy' function and requiring a minimum bootstrap value of 80, to obtain a table of relative abundances of operational taxonomic units. To address the issue of compositionality in the microbiome dataset<sup>46</sup>, the relative abundance values were normalized using the arcsine square-root transformation as described in ref. <sup>47</sup>. Measures of alpha diversity were computed<sup>47</sup>. The distributions of the Simpson and Shannon indices of alpha diversity on the transformed 16S abundance data are presented in Supplementary Table 4.

*Collection of venous blood samples.* Participants came into the clinical research facilities at 8:30 and were cannulated in the forearm antecubital vein. Venous blood was collected at 0 min (prior to a test breakfast) and at 9 time points postprandially (15, 30, 60, 120, 180, 240, 270, 300 and 360 min). Plasma glucose was analyzed from blood samples collected into fluoride oxalate tubes and centrifuged at 1,900g for 10 min at 4 °C. Serum C-peptide, insulin, triglyceride, fasting lipid profile, thyroid-stimulating hormone, alanine aminotransferase and liver-function panel were analyzed in blood samples collected into gel serum-separator collection tubes and were allowed to stand at room temperature before being centrifuged at 1,900g for 10 min at 4 °C. Samples were aliquoted and stored at −80 °C. Blood, for complete blood count (CBC) analysis, was collected into EDTA tubes, kept at 4 °C and analyzed within 12 h of collection.

*Serum biomarkers.* In the United Kingdom, insulin, glucose, triglyceride and C-peptide analysis was conducted by Affinity Biomarkers Labs. Glucose and triglyceride analyses were conducted on a Siemens ADVIA 1800 using Siemens assay kits (Siemens Healthcare Diagnostics). Triglyceride was analyzed using the ADVIA chemistry triglyceride method based on the Fossati three-step enzymatic reaction with a Trinder endpoint. Glucose was analyzed using the ADVIA chemistry glucose oxidase 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. CBC was measured by Viapath using standard automated clinical chemistry techniques. The inter-assay coefficient of variation for PREDICT samples analyzed by Affinity were as follows: insulin, 3.4%; C-peptide, 7.9%; triglyceride, 3.7%; and glucose, 2.6%.

In the United States, CBC was established using fresh blood samples in the MGH Core Laboratory. Hb1AC tests were performed by the MGH Diabetes A1c lab. Glucose, insulin, triglyceride and C-peptide were conducted by Quest Diagnostics using standard automated clinical chemistry techniques.

Upon completion of the US study, frozen serum and plasma samples were sent from the United States to the United Kingdom, and the entire cohort had measurements of a liver-function panel, full lipids (total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol and triglyceride), thyroid-stimulating hormone and alanine aminotransferase, which were performed by Affinity Biomarkers Labs. Details are described elsewhere<sup>48</sup>.

*Glucose using CGM.* Interstitial glucose was measured every 15 min using Freestyle Libre Pro CGMs (Abbott). Monitors were fitted by trained nurses on the upper, non-dominant arm at participants' baseline visit and were covered with Opsite Flexifix adhesive film (Smith & Nephew Medical) for improved durability, and were worn for the entire study duration (14 d). Data were collected 12 h and onwards after the device was activated for analysis. For a subgroup of participants ( $n = 377$ ), we fitted two monitors on their arms and calculated the coefficient of



variation (CV = 11.75%) and correlation ( $r = 0.97$ ) of their iAUC responses to standardized meals (Extended Data Fig. 2b).

**Time points for analyses.** Glucose. The 2-h glucose iAUC was used for both clinical and at-home analyses.

**Insulin and C-peptide.** C-peptide was measured at home as a surrogate for insulin secretion, because the reliability of C-peptide measured from DBS is higher than that of insulin (see ref. <sup>49</sup>), and C-peptide remains stable on paper filters for up to 6 months<sup>49</sup>. C-peptide was measured at 60 min postprandially to coincide with the peak in C-peptide seen in healthy individuals in the clinic, and again at 120 min to coincide with the strong decline in insulin level (Extended Data Fig. 2c). However, because previous genetic studies have tested the heritability of postprandial insulin at 120 min, this time point was included for our own heritability analyses (Fig. 2b,c). All other analyses refer to the 1-h rise for C-peptide.

**Triglyceride.** The rise in triglyceride at 6 h postprandially (triglyceride<sub>6h-rise</sub>) was selected to represent postprandial lipemic response from serum collected at clinic and in home-based DBS tests. This is a measure of lipemia that is most closely correlated with atherogenic lipoproteins, as compared with iAUC<sub>0-6h</sub>, C<sub>max</sub> (maximum serum concentration of triglycerides at 0–6 h) and 4-h triglyceride concentration (see refs. <sup>50–52</sup>).

**Activity and sleep.** Energy expenditure was measured using a triaxial accelerometer (AX3, Axivity) fitted by nurses at the baseline visit on the non-dominant wrist, which was worn for the duration of the study (except during water-based activities, including showers and swimming). Accelerometers were programmed to measure acceleration at 50 Hz with a dynamic range of  $\pm 8$  g (where g refers to local gravitational force equal to 9.8 ms<sup>2</sup>). 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<sup>53</sup>.

**Genotyping.** Whole-genome genotyping was available for 241 individuals from the UK cohort from previous TwinsUK studies. Genotyping was performed with the Illumina Infinium HumanHap610. Normalized GWAS intensity data were pooled and genotypes called on the basis of the Illuminus algorithm. No calls were assigned if the most likely call had a posterior probability of less than 0.95. Validation of pooling was done by visual inspection of 100 random, shared SNPs for overt batch effects (none were observed). SNPs that had a low call rate ( $\leq 90\%$ ), Hardy-Weinberg  $P < 1 \times 10^{-6}$  and minor allele frequencies  $< 1\%$  were excluded, and samples with call rates  $< 95\%$  were removed. Genotype imputations were performed to increase the coverage. Imputation of genotypes for all polymorphic SNPs that passed the quality-control stage were performed on the Michigan Imputation Server (<https://imputationserver.sph.umich.edu>) using the 1000G Phase3 v5 reference panel<sup>54</sup>. SNPs previously reported to be associated with postprandial glycemia, triglyceride or insulin in a GWAS<sup>17–20</sup> were extracted from the full set of genome-wide genotypes using PLINK, and were tested for association with postprandial measures using linear regression methods.

**Processing of habitual diet information.** UK nutrient intakes were determined using FETA software to calculate macro- and micronutrient data<sup>43</sup>. Submitted FFQs were excluded if more than 10 food items were left unanswered, or if the total energy intake estimate derived from FFQ as a ratio of the subject's estimated basal metabolic rate (determined by the Harris-Benedict equation)<sup>43</sup> was more than 2 s.d. outside the mean of this ratio ( $< 0.52$  or  $> 2.58$ ).

**Statistical analysis.** *Basic analyses.* The descriptive characteristics of study participants are summarized in Supplementary Table 1.

In order to reduce the dimension of the data, principal-component analyses (PCAs) with orthogonal transformation (varimax procedure) were applied to derive principal components (PCs) representative of individual characteristics (20 PCAs), microbiome (40 PCAs), meal composition (1 PCA), habitual diet (5 PCAs) and meal context (5 PCAs) (see Supplementary Table 3 for the full list of input variables). All the necessary prerequisites of PCA, including linearity, Kaiser-Meyer-Olkin measure of 0.88 and the significant Bartlett's test of sphericity ( $P < 0.001$ ), were met. Each participant received a score for each category mentioned above. To investigate the association between each outcome (iAUC, triglyceride<sub>6h-rise</sub>, C-peptide<sub>1h-rise</sub>) and our exposures (individual baseline characteristics, microbiome (16S), meal content, habitual diet and meal context), multivariable regressions were applied and  $R^2$  values were reported. Further, we derived PCAs for anthropometrics, biochemical/clinical factors, physical activity and sleep features separately to investigate their roles. Multi-collinearity for the multiple linear regressions was assessed with variance inflation factors (VIF) at each step<sup>55</sup>. Multi-collinearity was considered high when the VIF was  $> 1 \times 10^{38}$ . ROC curves were constructed and the AUC was calculated to assess the discriminatory power of fasting blood glucose versus 2 h glucose iAUC, fasting triglyceride versus triglyceride<sub>6h-rise</sub> and fasting C-peptide versus C-peptide<sub>1h-rise</sub> to detect IGT, and ASCVD 10-year risk (70% applied as a cut-off point). Values of

AUC range from 0.5 to 1, with 0.5 indicating no discrimination, and 1 indicating perfect discrimination of either IGT or ASCVD 10-year risk.  $P \leq 0.05$  was considered statistically significant. All analyses were performed using R (version 3.4.2, R Core Team).

**Meal composition.** To estimate macronutrient effects on glycemic response, we fitted a multivariate regression model with carbohydrates, fats, fiber and protein as predictors on meals 1, 2, 4, 5, 6, 7 and 8. Multi-collinearity was assessed for these predictors through VIF; we concluded that it was non-existent, VIF  $< 10$ . The regression coefficients were all significant ( $P < 0.001$ ) with values  $-79.23$  mmol per l per s,  $-142.41$  mmol per l per s and  $-185.49$  mmol per l per s for fat, fiber and protein, respectively, after adjustment by carbohydrates.

**Heritability and ACE model.** To estimate the heritability, we analyzed the data according to the classical ACE model. In this model, heritability is an approximation of the relative importance of additive genetic differences for variance of postprandial responses in the population<sup>56</sup>. Shared or familial environmental influences reflect experiences that contribute to twin similarity. Non-shared or individual-specific environmental influences refer to the contribution of environmental experiences not shared by family members. Information concerning shared genetic and environmental influences is best estimated by structural equation modeling techniques that fit models of twins by zygosity in order to describe the causes of the variance in postprandial responses. Therefore, the total variance in the trait can be partitioned into genetic variance (A), shared (familial) environmental variance (C) and individual-specific environmental variance (E). The level of statistical significance was set at  $P < 0.05$  in all analyses, and R software (version 3.0.2) together with the 'mets' (multivariate event times) package (<https://rdrr.io/cran/mets/src/R/methodstwinlm.R>) was used for all statistical analyses.

**Meal ranking.** Six different types of meal were ranked for each individual: the one with the highest 2 h glucose iAUC for that person was rank 6, the one with the second highest 2 h glucose iAUC was rank 5, and so on, down the one with the lowest 2 h glucose iAUC (rank 1). The distribution of these 'in-person rankings' is presented in Extended Data Fig. 3.

**Multilinear ANOVA to assess the role of individualized responses to meals.** The different sources of variation in glycemic response for meals 2, 3, 4, 6 and 8 (described in Supplementary Table 3) were analyzed using the multilevel linear ANOVA<sup>40</sup> model and were analyzed using a multilevel (hierarchical) linear Bayesian ANOVA model as described by Gelman and Hill<sup>57</sup>.

The different sources of variation in glycemic response for meals 2, 3, 4 and 6 and were analyzed using a multilevel (hierarchical) linear Bayesian ANOVA model as described by Gelman and Hill<sup>57</sup>.

Hierarchical Bayes models can accommodate non-normal dependent variables that are difficult to incorporate in classical ANOVA and multilevel linear models. The approach consists of sub-models at two levels: at level 1, the parameters of individuals, meals and person-meal interactions; and at level 2, the moments of the distributions from which level 1 parameters are drawn. Level 2 imposes some homogeneity on level 1 parameters, for example:

$$\alpha_m \sim N(0, \sigma_\alpha^2)$$

that is, the meal terms are distributed normally with the same s.d. ( $\sigma_\alpha$ ), ensuring homogeneity.

$$\sigma_\alpha \sim \text{half-Cauchy}(5)$$

that is, the standard deviation of the above distribution has a particular prior (a half-Cauchy distribution with a scale factor of 5).

The other terms ( $\beta_p$ ,  $\gamma_{m,p}$ ,  $\epsilon_{m,p,k}$ ,  $\epsilon_{m,p,k,n}$ ) have similar hierarchical distributions (though the s.d. of  $\epsilon_{m,p,k}$ ,  $\epsilon_{m,p,k,n}$  have a uniform prior distribution as opposed to a half-Cauchy distribution).

The parameters at both levels (that is, all the  $\alpha_m$  values and  $\sigma_\alpha$  and analogously for the other parameters) were sampled using a Markov chain Monte Carlo routine in pymc3 (ref. <sup>58</sup>). We plotted the sampled values of  $\sigma_\alpha$ ,  $\sigma_\beta$ ,  $\sigma_\gamma$ ,  $\sigma_\epsilon$  and  $\sigma_{\epsilon_n}$  in Fig. 6b.

$$\log(\text{iAUC}) = y_{m,p,k,n} = \alpha_m + \beta_p + \gamma_{m,p} + \epsilon_{m,p,k} + \epsilon_{m,p,k,n}$$

where:

$\log(\text{iAUC}) = y_{m,p,k,n}$  is the 2-h iAUC for person  $p$ , eating meal  $m$ , for the  $k$ th time measured on CGM  $n$  (given the availability of data with 2 CGMs for a subset as described in below).

$\alpha_m$  is the meal content (across all people) for meal  $m$ , for example high- or low-carbohydrate meals.

$\beta_p$  is individual glucose scaling (across all meals) for person  $p$ , for example overall high- or low-responding people.

$\gamma_{m,p}$  is the meal-specific response for individual  $p$  to meal  $m$ , for example a specific person responds particularly strongly to a specific meal.

$\epsilon_{m,p,k,n}$  is error stemming from the CGM (participants selected for this analysis wore two CGM devices, so  $n$  indexes the device providing the measurement).

$\epsilon_{m,p,k}$  is other sources of variation, including meal timing, exercise, sleep and circadian rhythm.

This Bayesian ANOVA model is a Bayesian hierarchical model that attempts to explain the observed log (iAUC) of a meal as a sum of categorical terms, that is, individuals are not classified according to any characteristics, but are included as unique individuals with log (2 h glucose iAUC) for various different meals. If this was an extended glycemic-index model, it would correspond to expressing the log (iAUC) as the sum of a meal term (analogous to the glycemic load of the meal) and an individualized term. This ‘individual glucose scaling’ is not a linear function of a person’s characteristics (such as age, sex or body-mass index), but rather is how each individual ranks overall given the log (iAUC) values for the various meals. This allowed us to test whether there was an interaction term between meals and persons, that is, an individualized response component to particular meals that was not merely due to a person being a high, average or low responder and to a meal having on average a higher glycemic response (for example, OGTT) than another meal (for example, a high-fat muffin). Given the availability of data concerning repeated occurrences of a person eating a particular meal and measurements from multiple CGMs for the same meal, we were able to extend the model to include a person–meal interaction and a CGM error and, analogously, infer the error due to the CGMs and the degree to which a person’s response to a particular meal is consistently higher or lower than expected from the glycemic index model, that is, a personalized glycemic load. The person–meal interaction effects allow different people to have different ordering of glycemic responses to meals, so one person might respond more strongly to meal A than to meal B, whereas another person might respond more strongly to meal B than meal A. Figure 6c shows 50% and 95% intervals on s.d. of the effects in the model. These can be approximately interpreted as percentage increase (or decrease) in iAUC contributed by the various effects in the model.

**CGM repeatability.** A subset of participants ( $n=377$ ) wore two CGMs simultaneously, providing duplicate measurements for the meals they consumed and therefore allowing us to distinguish CGM error from unexplained sources of variation. Postprandial glucose measurements for 3,280 meals eaten collectively by 377 participants in the UK cohort were used in this analysis. (Extended Data Fig. 2b).

**Computation of clinical indices.** Atherosclerotic cardiovascular disease risk (ASCVD) 10-year risk. (American Heart Association/Journal of the American College of Cardiology ASCVD 10-year risk) The 10-year ASCVD<sup>59</sup> risk score is a gender- and race-specific single multivariable risk-assessment tool used to estimate the 10-year CVD risk of an individual, and has clinically replaced the Framingham 10-year cardiovascular risk score. It is based on the age, sex, ethnicity, total and HDL cholesterol, systolic blood pressure, smoking status, use of blood-pressure-lowering medications and the presence of type 2 diabetes.

**Impaired glucose tolerance.** We used the standard definition from the American Diabetes Association<sup>60</sup> (fasting plasma glucose  $< 7.0$  mmol l<sup>-1</sup> and OGTT 2-h value  $\geq 7.8$  mmol l<sup>-1</sup> but  $< 11.1$  mmol l<sup>-1</sup>).

**Validation of machine-learning model cross-validation and difference (Bland–Altman plots).** To further illustrate the reliability of the machine-learning predictions, we conducted a leave-one-out cross-validation procedure and generated Bland–Altman plots to analyze the agreement between two. To generate the Bland–Altman plots we used the Predict UK and US data showing predicted versus measured postprandial responses. We generated Bland–Altman plots for predicted and measured postprandial responses for each biomarker (triglycerides, C-peptide and glucose). (Extended Data Fig. 4a).

**Leave-one-out cross-validated Pearson’s  $r$  in Predict UK.** To perform  $k$ -fold cross-validation, the entire dataset was split into  $k$  groups. Treating each group as a test set and the remaining groups as the training set, the model is fitted  $k$  times. The Pearson’s  $r$  between the values predicted by the fitted models and the measured values in the test sets is used as the metric for model evaluation, which we refer to as the cross-validated Pearson  $r$ .

The special case, in which  $k$  is the size of the dataset, is referred to as leave-one-out cross-validation, and we refer to the corresponding evaluation metric as leave-one-out cross-validated Pearson’s  $r$ . The machine-learning models for the three biomarkers of interest were evaluated using the aforementioned metric and are reported in the Extended Data Fig. 4b. These scores are similar to the cross-validated fivefold scores in the main text.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The data used for analysis in this study are held by the department of Twin Research at Kings College London. The data can be released to bona fide researchers using our normal procedures overseen by the Wellcome Trust and its guidelines as part of our core funding. We receive around 100 requests per year for our datasets and have a meeting 3 times per month with independent members to assess proposals. The application is at <https://twinsuk.ac.uk/resources-for-researchers/access-our-data/>. This means that the data need to be

anonymized and conform to GDPR standards. Specifically for this paper, all the variables used in the models can be requested as well as the summary outcome measures for each person. The 16S microbiome data will be uploaded onto the EBI website (<https://www.ebi.ac.uk/>) with unlimited access. Source data for Figs. 2–6 and Extended Data Figs. 2–4 are presented with the paper

## Code availability

The scripts for statistical analysis are freely available upon request to the Department of Twins Research and Genetic Epidemiology at King’s College London. Application is via <https://twinsuk.ac.uk/resources-for-researchers/access-our-data/>. The scripts used for machine-learning analyses can be found in the Supplementary Software.

## References

- Berry, S. D. D. et al. Personalised REsponses to Dietary Composition Trial (PREDICT): an intervention study to determine inter-individual differences in postprandial response to foods. *Protoc. Exch.* <https://doi.org/10.21203/rs.2.20798/v1> (2020).
- Verdi, S. et al. TwinsUK: the UK adult twin registry update. *Twin Res. Hum. Genet.* **22**, 523–529 (2019).
- Mulligan, A. A. et al. A new tool for converting food frequency questionnaire data into nutrient and food group values: FETA research methods and availability. *BMJ Open.* **4**, e004503 (2014).
- 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).
- Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
- Gloor, G. B., Wu, J. R., Pawlowsky-Glahn, V. & Egozcue, J. J. It’s all relative: analyzing microbiome data as compositions. *Ann. Epidemiol.* **26**, 322–329 (2016).
- Thomas, A. M. et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat. Med.* **25**, 667–678 (2019).
- 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).
- Johansson, J., Becker, C., Persson, N. G., Fex, M. & Torn, C. C-peptide in dried blood spots. *Scand. J. Clin. Lab. Invest.* **70**, 404–409 (2010).
- Samson, C. E., Galia, A. L., Llave, K. I., Zacarias, M. B. & Mercado-Asis, L. B. Postprandial peaking and plateauing of triglycerides and VLDL in patients with underlying cardiovascular diseases despite treatment. *Int. J. Endocrinol. Metab.* **10**, 587–593 (2012).
- Madhu, S., Sinha, B., Aslam, M., Mehrotra, G. & Dwivedi, S. Postprandial triglyceride responses and endothelial function in prediabetic first-degree relatives of patients with diabetes. *J. Clin. Lipidol.* **11**, 1415–1420 (2017).
- Nakamura, A. et al. Different postprandial lipid metabolism and insulin resistance between non-diabetic patients with and without coronary artery disease. *J. Cardiol.* **66**, 435–444 (2015).
- van Hees, V. T. et al. Estimating sleep parameters using an accelerometer without sleep diary. *Sci. Rep.* **8**, 12975 (2018).
- Eitel, I. et al. Genome-wide association study in takotsubo syndrome—preliminary results and future directions. *Int. J. Cardiol.* **236**, 335–339 (2017).
- Slinker, B. K. & Glantz, S. A. Multiple regression for physiological data analysis: the problem of multicollinearity. *Am. J. Physiol.* **249**, R1–R12 (1985).
- Purcell, S. Variance components models for gene-environment interaction in twin analysis. *Twin Res.* **5**, 554–571 (2002).
- Gelman, A. & Hill, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models* (Cambridge University Press, 2007).
- Salvatier, J., Wiecki, T. V. & Fonnesbeck, C. Probabilistic programming in Python using PyMC3. *PeerJ Comput. Sci.* **2**, e55 (2016).
- Goff, D. C. Jr. et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* **129**, S49–S73 (2014).
- American Diabetes Association Classification and Diagnosis of Diabetes. *Diabetes Care* **39**, S13–S22 (2016).

## Acknowledgements

We express our sincere thanks to the participants of the PREDICT 1 study. We thank M. McCarthy (University of Oxford) and L. Hodson (University of Oxford) for valuable feedback on the manuscript. We thank the staff of Z. Global, the Department of Twin Research and Massachusetts General Hospital, for their tireless work in contributing to the running of the study and data collection. We thank Abbott for their support with their CGMs. This work was supported by Zoe Global and also received support from grants from the Wellcome Trust (212904/Z/18/Z) and the Medical Research Council (MRC)/British Heart Foundation Ancestry and Biological Informative Markers for Stratification of Hypertension (AIMHY; MR/M016560/1). S.E.B. was supported in

part by a grant funded by the BBSRC (BB/NO12739/1). P.W.F. was supported in part by grants from the European Research Council (CoG-2015\_681742\_NASCENT), Swedish Research Council, Novo Nordisk Foundation and the Swedish Foundation for Strategic Research (IRC award). A.M.V. was supported by the National Institute for Health Research Nottingham Biomedical Research Centre. TwinsUK is funded by the Wellcome Trust, Medical Research Council, European Union, Chronic Disease Research Foundation (CDRF), Zoe Global and the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. J.M. was supported by a fellowship funded by the European Commission Horizon 2020 program (H2020-MSCA-IF-2015-703787) and by National Institutes of Health grant P30 DK40561. N.S. received support from the European Research Council (ERC-STG project MetaPG), the European H2020 program (ONCOBIOME-825410 project and MASTER-818368 project), and the National Cancer Institute of the National Institutes of Health (1U01CA230551).

### Author contributions

Obtained funding: J.W., G.H., T.D.S. Study design and developed concept: S.E.B., A.M.V., J.W., G.H., R.D., A.T.C., N.S., P.W., P.W.F., T.D.S. Data collection: S.E.B., D.A.D., H.A.K., M. Mangino, J.M., I.L., D.H. Data analysis: A.M.V., M. Mangino., F.A., R.D., J.C., C.B.,

S.G., E.B., M. Mazidi, P.W., N.S. Study coordination: S.E.B., H.A.K., D.A.D., G.H., J.W. Writing the manuscript: S.E.B., A.M.V., D.A.D., M. Mazidi, J.W., J.C., I.L., J.M.O., C.D.G., L.M.D., A.T.C., N.S., P.W.F., T.D.S. All authors reviewed and revised the final manuscript.

### Competing interests

T.D.S., S.E.B., A.M.V., F.A., P.W.F., L.M.D. and N.S. are consultants to Zoe Global Ltd ('Zoe'). J.W., G.H., R.D., H.A.K., J.C., C.B., S.G., E.B., P.W. and I.L. are or have been employees of Zoe. Other authors have no conflict of interest to declare.

### Additional information

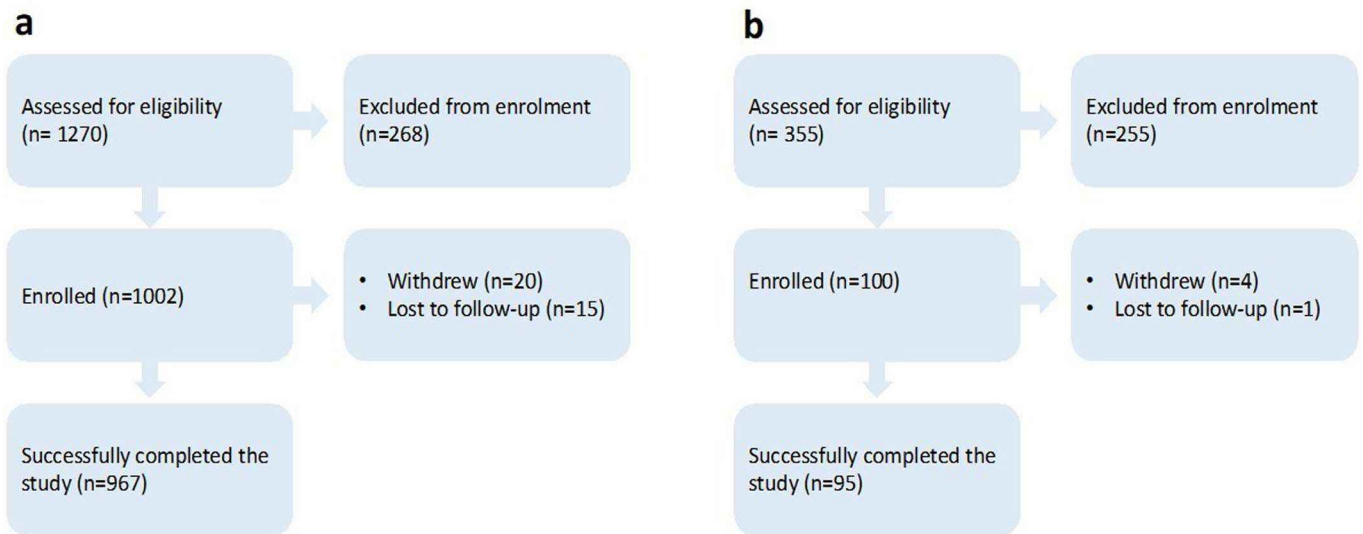
**Extended data** is available for this paper at <https://doi.org/10.1038/s41591-020-0934-0>.

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41591-020-0934-0>.

**Correspondence and requests for materials** should be addressed to A.M.V. or T.D.S.

**Peer review information** Jennifer Sargent was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).



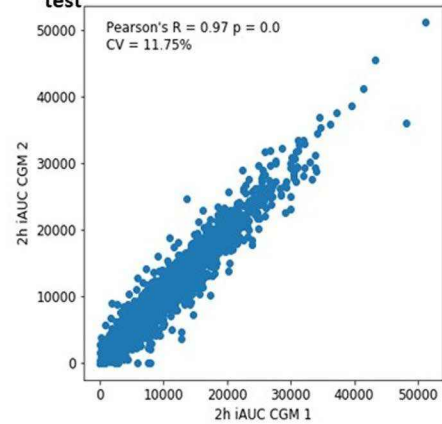
**Extended Data Fig. 1 | Consort Diagrams for UK and US populations in the PREDICT 1 study.** Consort Diagrams for **a**, UK and **b**, US populations in the PREDICT 1 study.



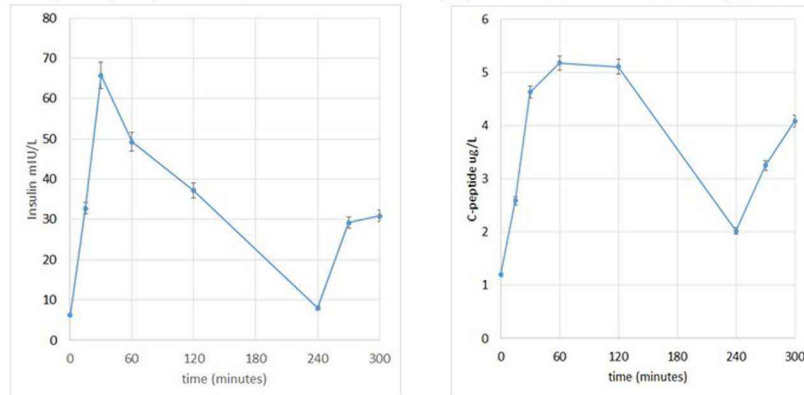
**a. Intraclass correlations**

Metabolic outcome	num meals	Timepoint/ trait ICC	95% CI
TG	648	fasting	0.670.60 - 0.72
		5h	0.660.60 - 0.72
		6h	0.730.68 - 0.78
		fasting-5h rise	0.410.32 - 0.49
		fasting-6h rise	0.460.37 - 0.54
glucose	8038	2h-iAUC	0.74 0.72 - 0.75
c-peptide	626	fasting	0.720.66 - 0.77
		0.5h	0.560.47 - 0.64
		2h	0.710.64 - 0.76
		fasting-0.5h rise	0.470.37 - 0.56
		fasting-2h rise	0.620.54 - 0.69

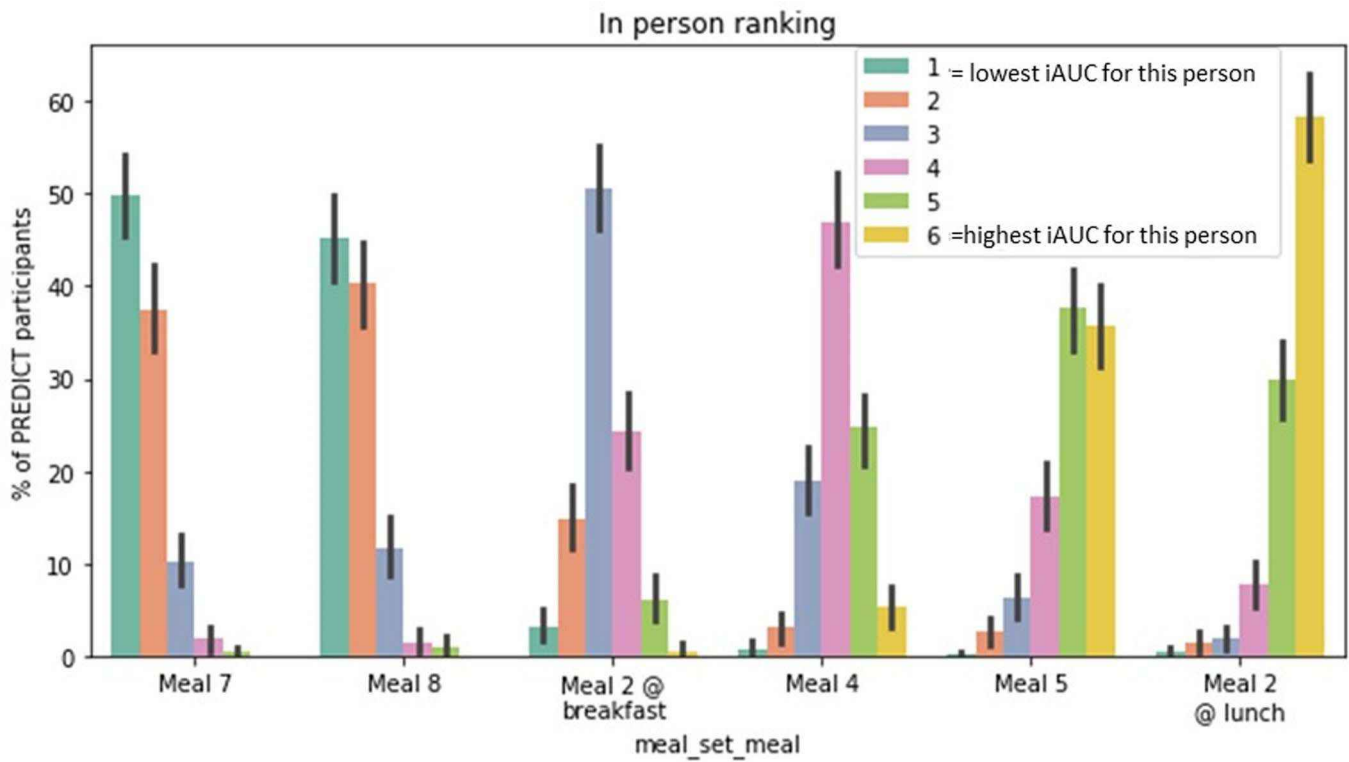
**b. Pearson's correlation and CV of 2h-iAUCs measured with two monitors worn by the same participant (n=377). P-value from two-sided t-test**



**c. Mean and standard error of fasting and postprandial serum insulin and C-peptide concentrations during the clinic visit in the PREDICT 1 study, n=1,036.**

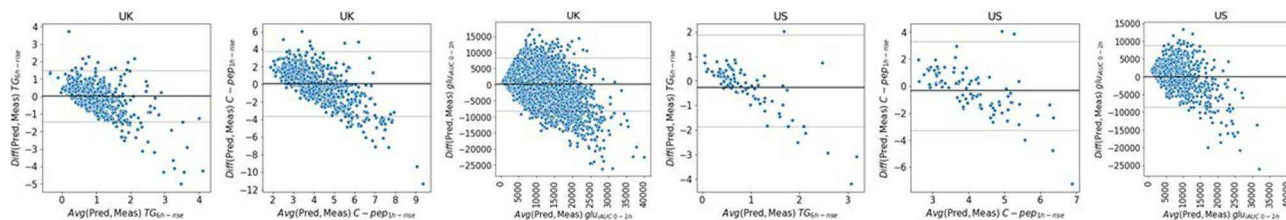


**Extended Data Fig. 2 | Repeatability in the PREDICT 1 study.** Repeatability in the PREDICT 1 study. **a.** Intraclass correlations. **b.** Pearson's correlation and CV of 2h-iAUCs measured with two monitors worn by the same participant (n=377). P-value from two-sided t-test **c.** Mean and standard error of fasting and postprandial serum insulin and C-peptide concentrations during the clinic visit in the PREDICT 1 study, n=1,036.



**Extended Data Fig. 3 | Frequency distribution of in-person ranking for 6 of meals shown in Fig. 6a.** Frequency distribution of in-person ranking for 6 of meals shown in Fig. 6a. (High fat 40 g = meal 7, High protein = meal 8, UK average = meal 2, High carb = meal 4, OGTT = meal 5, UK average at lunch = meal 2). n=1102 participants.

a. Bland-Altman plots comparing predicted and measured postprandial responses in TG, glucose and C-peptide using UK and US data. The sample sizes used n=number of meals: triglyceride UK: n=958 US: n=91; C-peptide UK: n=957 US: n=93; Glucose UK:n=11550 US: n1200

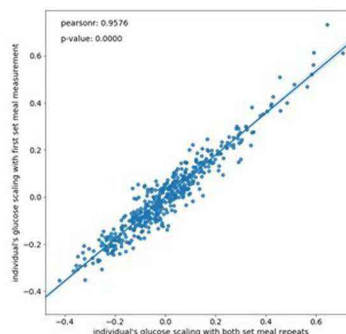


b. Leave-one-out cross-validated Pearson R scores in PREDICT UK. 5-fold cross validation for Triglyceride 6 hour rise on n=958 meals, for Glucose 2h iAUC on n=11,550 meals, p-values shown for two-sided t-test

UK data set TG <sub>6h-rise</sub>	Glucose <sub>iAUC0-2h</sub>
Leave-one-out 0.49, p = 2.03e-56	0.77, p = 0.0

c. comparison of models using repeat meals vs not using them

individual glucose scaling estimate using all repeat meals :  
 18.7% change in glucose 2h iAUC 95%CI [17.9-19.5%]  
 individual using only a single meal from each set (one OGTT, one high carb, one high fat etc).  
 18.2 % change in glucose 2h iAUC 95%CI [17.2-19.2%].



**Extended Data Fig. 4 | Machine learning comparisons, cross validation and repeatability.** Machine learning comparisons, cross validation and repeatability. **a**, Bland-Altman plots comparing predicted and measured postprandial responses in TG, glucose and C-peptide using UK and US data. Sample sizes used were (n = number of meals) triglyceride UK: n = 958 US: n = 91; C-peptide UK: n = 957 US: n = 93; Glucose UK: n = 11550 US: n = 1200. **b**, Leave-one-out cross-validated Pearson R scores in PREDICT UK. 5-fold cross validation for triglyceride<sub>6h-rise</sub> on n = 958 meals, for Glucose<sub>iAUC0-2h</sub> on n = 11,550 meals, p-values shown for two-sided t-test **c**, comparison of models using repeat meals vs not using them.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

Data from questionnaires, clinical visits and laboratory data was entered using comma delimited files, Excel spreadsheets and Microsoft access (Microsoft Office 365 2019). CGM and accelerometer data was imported from text files into the analysis pipeline.

Data analysis

Analyses were carried out using version 3.4.2 R Core Team ,the “mets” (Multivariate Event Times) package in R, and the DADA2 pipeline

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data used for analyzing this study are held by the department of Twin Research at Kings College London. The data can be released to bona fide researchers using our normal procedures overseen by the Wellcome Trust and its guidelines as part of our core funding. We receive around 100 requests per year for our datasets and have a meeting three times a month with independent members to assess proposals. Application is via <https://twinsuk.ac.uk/resources-for-researchers/access-our-data/>. This means that the data needs to be anonymized and conform to GDPR standards. Specifically for this paper, all the variables used in the models can be requested as well as the summary outcome measures for each person. The 16S microbiome data used here will be uploaded onto the EBI site with unlimited access.



## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	1002 individuals were recruited for the UK cohort (allowing for 80% power to detect correlations $r=0.15$ with $p<0.0001$ to allow adjustment for 500 test), 100 for the US cohort to enable replication of the larger effects (80% for $r=0.28$ with $p<0.05$ ), the final sample size excludes individuals who dropped out (didn't finish the clinic visit)
Data exclusions	pre-established exclusion criteria were T2D. At the analysis level, exclusions were performed wherever c-peptide, glucose, triglyceride measures were missing for any of the postprandial responses.
Replication	the US cohort has been used as replication for findings in the larger UK cohort. Prediction models for triglyceride and glucose were successfully replicated but not so for c-peptide
Randomization	given the single arm nature of the study, randomization was not applicable
Blinding	there was no control or placebo arm therefore blinding was not applicable

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	Study participants were healthy individuals aged between 18-65 years and able to provide written informed consent
Recruitment	Participants enrolled already in the TwinsUK cohort were recruited as part of studies which are included in the cohort's annual newsletter and also mentioned on our website: <a href="http://www.twinsuk.ac.uk/">http://www.twinsuk.ac.uk/</a> Non-twins were recruited via independent recruitment agencies, project specific website and online advertising including the use of social media platforms and social media
Ethics oversight	Ethics was granted by the London - Hampstead Research Ethics Committee (REC approval 18/LO/0663) and by the Partners Healthcare IRB (2018P002078)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	NCT03479866
Study protocol	<a href="https://clinicaltrials.gov/ct2/show/record/NCT03479866">https://clinicaltrials.gov/ct2/show/record/NCT03479866</a>

Data collection

Data was collected at St Thomas' Hospital, London and Massachusetts General Hospital, Boston

Outcomes

- Gut microbiome profile, triglyceride blood concentration, glucose blood concentration, record of sleep pattern using a wearable device (i.e. fitness watch), record of physical activity using a wearable device (i.e. fitness watch)