

# The association between body mass index and metabolite response to a liquid mixed meal challenge.

## Authors

David A. Hughes<sup>1,2</sup>, Ruifang Li-Gao<sup>3</sup>, Caroline J. Bull<sup>1,2</sup>, Renée de Mutsert<sup>3</sup>, Frits R. Rosendaal<sup>3</sup>, Dennis O. Mook-Kanamori<sup>3,4</sup>, Ko Willems van Dijk<sup>5,6</sup>, Nicholas J. Timpson<sup>1,2</sup>.

## Affiliations

<sup>1</sup> MRC Integrative Epidemiology Unit at the University of Bristol, UK  
<sup>2</sup> Population Health Science, Bristol Medical School, University of Bristol, UK  
<sup>3</sup> Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands  
<sup>4</sup> Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands  
<sup>5</sup> Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands  
<sup>6</sup> Department of Internal Medicine, Division of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands

## Keywords

BMI, metabolomics, metabolite, Mendelian Randomization, MR, meal challenge, response, NEO

## Abstract

### 18 Background

Metabolite abundance is a dynamic trait that is not only variable in a fasting state, but also varies in response to environmental stimuli, such as food consumption. Postprandial abundance and response to a meal are emergent traits in studies of disease and which themselves may be subject to specific risk factors. We investigated body mass index (BMI) as a recognized risk factor for numerous health outcomes that may influence metabolite response to feeding. Here we use the Netherlands Epidemiology of Obesity (NEO) study to examine associations between BMI and metabolite response to a liquid meal and extend this by using Mendelian randomization (MR) to estimate potential causal effects.

### 27 Methods and findings

The NEO study conducted a liquid meal challenge and collected metabolite profiles using the Nightingale metabolomics platform in 5744 study participants. Observational and one-sample MR analysis were conducted to estimate the effect of BMI on metabolites and ratios of metabolites (n = 229) in the fasting, postprandial and response (or change in abundance) states. After an appropriate multiple testing correction, we observed 473 associations with BMI (175 fasting, 188 postprandial, 110 response) in observational analyses. In MR analyses, we observed 20 metabolite traits (5 fasting, 12 postprandial, 3 response) to be associated with BMI. In both the fasting and postprandial state, this included citrate and the ratios of linoleic acid, omega-6 fatty acid and polyunsaturated fatty acids to total fatty acids. In addition, the glucogenic amino acid alanine was inversely associated with BMI in the response state, suggesting that as alanine increased in postprandial abundance, that increase was attenuated with increasing BMI.

### Conclusions

Overall, MR estimates were strongly correlated with observational effect estimates suggesting that the broad associations seen between BMI and metabolite variation in fasting, postprandial and response states have a causal underpinning. Specific effects in previously unassessed postprandial and response states were detected and these may likely mark novel life course risk exposures driven by regular nutrition.

## 48 Introduction

51 The excess accumulation of body fat and obesity are an established global health burden, the prevalence of which is increasing. The incidence and cost in life and dollars was estimated to be 1.9 billion cases in 2016, 4 million attributed deaths in 2015 and two trillion US dollars in 2014, respectively (1–3). Body mass index (BMI), or the ratio of one’s weight and height squared ( $\text{kg}/\text{m}^2$ ), is a common metric for measuring excess weight or fat and BMI thresholds of greater than or equal to a value of 25 and 30 are used to designate the conditions of overweight and obesity, respectively (4,5). While the use of BMI as an obesity metric – one that varies with age, sex, and ethnicity/genetic ancestry - may underestimate the true global burden of excess body fat, it has been an important and easily derived anthropometric and diagnostic metric in the study of obesity (6–9). Increases in BMI have been associated with 60 decreased life expectancy (10–13), some cancers (14–16) and with cardiometabolic diseases including cardiovascular disease, and type 2 diabetes (17).

63 Despite the scale of the BMI related health burden, there remains a need to elucidate how and by what intermediate physiological traits excess body fat increases disease risk. Possible intermediates are metabolites and lipoproteins, some of which – like low density lipoprotein (LDL) or its’ crucial protein apolipoprotein B - have been previously associated with both BMI and disease risk (18–26). In a recall-by-genotype framework, previous work supports a causal role of body mass index on metabolite variation (27,28). Other work demonstrated a 69 broad observational and consistent causal association between BMI and metabolites or specifically lipoprotein subclasses, branch-chain amino acids, and inflammatory markers in a young adult cohort (18).

72 In 1950, John Moreton illustrated the presence of intra-individual variability in chylomicrons after a high fat food (cream) challenge (29). More recent work has illustrated additional 75 metabolite variation postprandially that could prove informative in evaluating disease risk and informing precision medicine (30–35). Yet, Schutte et al, building on the work of others (18), has illustrated that observational effect estimates between BMI and metabolite abundances are 78 strongly, but not uniformly, consistent between fasted and postprandial metabolite abundance measurements (36). This observation suggests that for some metabolites fasting versus postprandial status may not be critically important for meta-analyses that aim to estimate the

81 association between BMI and metabolite abundance. It also suggests that for other metabolites  
there is novelty in postprandial measurements and their association with BMI.

84 In this context and with most individuals spending much of their waking hours in a non-fasted  
or postprandial state, it is important to understand both the variation in postprandial  
87 metabolite abundance and the association and influence BMI may have on dynamic ranges in  
metabolomic response. Similar to an oral glucose tolerance test (OGTT), whose postprandial  
abundance is informative about an individuals' insulin sensitivity and diabetic state,  
postprandial metabolite abundances may prove informative in the study of disease traits (37).  
90 Consequently, a comprehensive assessment of postprandial variation and response is essential  
in the study of adiposity, metabolites, and disease.

93 Response, defined as the change in metabolite abundance between fasted and postprandial  
states, has received little to no attention in disease risk research. In contrast to fasting status,  
response can be assessed either through a basic measurement of postprandial metabolite  
96 profile after a standardized feed (assuming that baseline variability is essentially random in a  
large enough sample), or through the explicitly modelling of response accounting  
appropriately for pre-feed variation. Either way, there is evidence that this dynamic state is  
99 different and potentially informative versus classically measured fasting levels. For example,  
a recent genome-wide association analysis identified genetic variants uniquely associated with  
metabolite response traits that did not share associations with fasting or postprandial  
102 abundances (38). Moreover, estimates of genotype heritability for some response traits such as  
the (branched-chain) amino acids, glucose and extremely large VLDL are as large or larger  
than those observed for fasting and postprandial estimates. Taken together, results suggest that  
105 such state specific traits are likely to be qualitatively different and may prove informative in  
explaining the aetiology of disease.

108 Currently missing from the literature are studies explicitly assessing the potentially causal  
relationship between variation in the commonly recognized risk factor BMI and the dynamic  
response of the metabolome to feeding. This is potentially difficult to examine given the  
111 strongly confounded nature of BMI as a risk factor, however Mendelian randomization (MR)  
is a framework using instrumental variable (IV) analysis to deploy genetic variants as  
exposure proxies in efforts to obtain evidence of causal relationships (39,40). It has been  
114 increasingly possible to exploit an increasing collection of confirmed genetic associations

with BMI for the purpose of estimating likely downstream consequences in variation (41–49) and aggregate genetic scores have become useful in their capacity to capture heritable  
117 variance in exposures of interest (like BMI) despite potential complexities (50). There is, therefore, an opportunity to apply this approach to the examination of potential relationships between BMI and metabolite profiles.

120

We aimed to bring together data measuring the metabolic response to a liquid mixed meal and measures of the exposure, BMI, in observational and MR frameworks. These analyses,  
123 undertaken in the Netherlands Epidemiology of Obesity (NEO) study, sought to estimate relationships between variation in BMI and metabolite abundance in both a fasted and postprandial states, and to assess metabolite response – or change in abundance between a  
126 fasted and postprandial states.

## Materials and Methods

### Study Sample

129 The sample population was 6,671 participants of The Netherlands Epidemiology of Obesity  
(NEO) study, a prospective cohort study (51). The NEO study was conducted at the Leiden  
University Medical Center (LUMC) and approved by the LUMC Medical Ethical Committee.  
132 All participants gave written informed consent.

The NEO sample population is derived from two populations. First, all individuals between  
135 the ages of 45 and 65 years, living in the greater Leiden area, with a self-reported BMI of 27  
kg/m<sup>2</sup> or higher. Second, all individuals in the municipality of Leiderdorp of the same age  
range, regardless of BMI (**S1 Fig.**). We will refer to these two sample populations as Leiden  
138 and Leiderdorp respectively, and the entire study sample as NEO. Here our study sample  
population has a size of 5,744 defined by those participants who have genotype data,  
metabolite data, and other phenotype or covariable data available. Further details of the NEO  
141 study and study design are previously published (52).

### Meal challenge

An overnight fast was requested of all participants who subsequently visited the LUMC NEO  
144 study center and given a food challenge (38). During the visit an initial baseline, fasted blood  
sample was drawn. Second, within 5 minutes of the initial blood draw participants were given  
a liquid mixed meal. The meal was 400mL in volume with 600 kcal of energy. Sixteen  
147 percent of that energy (En%) was derived from protein, 50 En% from carbohydrates, and 34  
En% was from fat. Third, following completion of the food challenge a blood sample was  
drawn at 30 minutes and again at 150 minutes. The NEO meal challenge occurred between the  
150 months of September 2008 and September 2012, with participants from Leiden sampled  
throughout this period of time. Those participants from Leiderdorp were sampled, non-  
exclusively, during the months of June 2011 and September 2012.

153  
The meal challenge sampling of Leiderdorp and Leiden population participants was structured  
by sampling date and as a product BMI was structured by sampling date. Specifically, BMI  
156 was associated with the date of measurement (univariable linear model F-test p-value =  
4.7x10<sup>-229</sup>), explaining 20.3% of the total variation in BMI (**S2 Fig.**). Despite this, the BMI-  
PGS had a smaller proportion (2.0%) of its variation explained by sampling date (univariate

159 linear model F-test p-value =  $3.9 \times 10^{-7}$ ). As a product of this structure, sampling date is included as a covariate in all observational and MR analyses.

### Metabolite data

162 NEO participants had plasma derived metabolite profiles measured using the Nightingale Health (Helsinki, Finland)  $^1\text{H}$  nuclear magnetic resonance (NMR) platform (53). For each participant both their fasting and 150 minutes postprandial samples were assayed. We note  
165 here that metabolites are commonly defined as biological molecules less than 1.5 kilodaltons in size and that many of the molecules (lipids) assayed here are larger than this threshold. For simplicity we will refer to them all here as metabolites. At the time of sampling this  
168 metabolomics platform provided 229 measurements for 149 metabolites including 80 derived ratio measurements (**Table S1 in S1 File**) from 14 substance classes: amino acids (n=8), apolipoproteins (n=3), cholesterol (n=9), fatty acids (n=11), fatty acids ratios (n=8), fluid  
171 balance (n=2), glycerides and phospholipids (n=9), glycerides and phospholipid ratios (n=2), glycolysis related metabolites (n=3), inflammation (n=1), ketone bodies (n=2), lipoprotein particle size (n=3), lipoprotein subclasses (n=98), lipoprotein subclass ratios (n=70).

### 174 Data quality control

Prior to data analysis, metabolites and covariable quality control and data filtering steps were implemented. First, a single individual who self-described themselves as “other” rather than  
177 “white” was excluded from the analysis. Second, any non-metabolite covariable with less than 1000 observations (n=1), with no variation (n=1), or binary covariables with fewer than 10 observations in either of the two binary classes (n=1) were removed. A complete list of  
180 evaluated non-metabolite covariables is available in **Table S2 in S1 File**.

The initial metabolite data set consisted of 5,744 individuals, 229 fasting metabolites, 229  
183 postprandial metabolites, and 148 previously derived (38) orthogonal nonlinear least squares (ornls) response metabolite traits. Data quality control of this data set was performed with the R package *metaboprep* (54). A complete description of the procedure and parameters used  
186 with *metaboprep* can be found in **S1 Text**. After data QC 226 samples and 3 metabolites (all ornls response traits) were filtered from the data set (**S1 Log**). The log file (**S1 Log**) and report (**S1 Report**) generated by *metaboprep* are provided. After running *metaboprep* we also  
189 performed the following steps on the fasted and postprandial data. First, all zero values were converted into NAs. Second, for each metabolite (in the fasting and postprandial state, individually) all values 10 interquartile distances from the median were also turned into NAs.

192 Third, we used the expected correlated nature of the fasting and postprandial data to identify  
outliers of that relationship, and turned them into NAs, in both dietary states. Further details  
on this procedure can be found in **S1 Text** and pre- and post-QC scatter plots for each  
195 metabolite can be found in **S2 File**.

### Response trait

A response trait, or a measure of change between the postprandial and fasting dietary state,  
198 was derived for each of the 229 metabolites traits by univariable Deming regression. A  
Deming regression was used here as it allows for error in both independent (fasting) and  
dependent (postprandial) variables of the model and as such is a type of total least squares  
201 regression. Following the QC steps described above we fit a Deming regression of  
postprandial on fasting data, for each metabolite. After model fitting the residuals were  
extracted and used as the response trait. The function `deming()` from the `deming (v1.4)` R  
204 package was used in the analysis. The intercept and slope of Deming regression were recorded  
and are provided in **Table S1** in **S1 File**, along with trait annotations and population summary  
statistics for all metabolite traits.

207  
The residuals of the Deming regressions derived here were used as metabolite traits of  
response. These response traits effectively mirror those of a simple delta, in that positive  
210 values indicate an increase in metabolite abundance in the postprandial state and negative  
values would indicate a decrease. Effect estimates from an association analysis with BMI,  
described below, therefore suggests whether response is associated with BMI. Positive effect  
213 estimates suggest that as BMI increases, so too does the change in metabolite abundance.  
Negative effect estimates suggest that as BMI increases the change in metabolite abundance  
decreases.

216  
Deming regressions residuals were taken to be the most appropriate way to summarize  
response for second stage analyses - observational and MR. However, in addition to the  
219 Deming regression, we also used two alternative methods to derive response. The first was a  
simple delta estimate (postprandial minus fasting abundance) and the second was an  
orthogonal nonlinear least squares (ornls) regression. We note that simple delta values and  
222 Deming residuals correlate with a mean Pearson's  $r$  of 0.94 (95% CI 0.65-0.99). The ornls  
response traits were derived and used previously in a genome wide association study of  
metabolite response and was used in our metabolite quality control steps described above



225 (38). Our result and discussion will focus exclusively on the Deming response traits, but all  
effect estimates for the delta and ornls response traits are provided in **S3 File**.

### Effective number of tested metabolites

228 Whilst there are 687 metabolite traits (fasting = 229, postprandial = 229, response = 229)  
these are not independent (**S3 Fig.**). We used the R package *iPVs*  
(<https://github.com/hughesevoanth/iPVs>) to estimate the effective number of metabolites in  
231 the data set (55–57). This allows us to ensure that we are not over correcting for the number  
tests we are performing in the study. We estimated 43 representative variables in the NEO  
metabolite data set. The data reduced, study-wide Bonferroni (BF) corrected p-value was set  
234 to 0.05/43 or  $1.163 \times 10^{-3}$ . A full description of the iPVs procedure and parameters can be  
found in **S1 Text**.

### Metabolite data description

237 Two analyses were carried out to describe this Nightingale Health metabolomics data set.  
First, given the abundance of lipoproteins and their lipids in this data set we estimated the  
mean abundance of each lipid for each lipoprotein in both the fasting and postprandial dietary  
240 states. Second, a paired Student's t-test was performed to determine if mean abundances  
differed between the fasting and postprandial states. Our threshold for declaring a change in  
mean abundance was 0.05/229, where 229 is the number of metabolites tested. We also  
243 estimated the change as a delta value between the postprandial and fasting states and then  
derived estimates for average change and the 95% confidence interval of the change  
distribution.

### Rank normal transformation of metabolite traits

246 Each metabolite trait distribution was tested for normality with a Shapiro-Wilk test. All W-  
statistics (untransformed distributions and model residuals) are reported in **S3 File**. In total  
249 55.56% of all metabolite traits had a W-statistics less than 0.95, a threshold used here to  
define an inconsistency with normality. A total of 43.94% of all log-transformed metabolite  
traits had a W-statistics less than 0.95. Consequently, for the purpose of signal discovery and  
252 to allow for parametric analysis, each metabolite was rank normal transformed (tied values  
randomly ranked), prior to linear modelling. These steps resulted in the residuals from all  
observational and MR models to have a W-statistic greater than 0.9 and 0.99, respectively.

255

## Covariables

Variables from 14 categories were compiled and available to identify possible confounders by  
258 evaluating associations with BMI, BMI polygenic score (BMI-PGS), and metabolite traits  
(**Table S2** in **S1 File**). These 14 categories are: anthropomorphic (n = 8), education (n = 5),  
income (n = 2), smoking (n = 3), diet including alcohol intake in grams (n = 24), medication  
261 including glucose and lipid lowering (n = 10), systolic and diastolic blood pressure (n = 2),  
health including diabetes, glucose, hypertension and cancer status (n = 4), physical exercise (n  
= 18), basal metabolic rate (n = 1), indirect calorimetry (n = 5), genotype principal  
264 components (n = 4). In addition, seven metrics of biological sample quality (n = 7) and the  
sampling information (visit date and sub-population, n = 2) were included in the study to  
assess influence on metabolites. During quality control, described above, 3 variables were  
267 filtered. As performed with metabolites above we estimated the effect number of covariables  
in the data set. After running an iPVs analysis with a tree cut height of 0.5 we estimated that  
there are 54 representative or independent covariables. As a product any test performed across  
270 all covariables was corrected for multiple testing at a p-value of 0.05/54 or  $9.29 \times 10^{-4}$ .

## Genotype data

The NEO participants were genotyped on the Illumina HumanCoreExome-24 BeadChip  
273 (Illumina Inc., San Diego, California, United States of America), at the Centre National de  
Génotypage (Evry Cedex, France). Following genotype quality control steps, as previously  
described (58), sample were imputed to the Haplotype Reference Consortium (HRC) release  
276 1.1 (59), with further details in Li-Gao et al (38).

## Construction of the BMI polygenic score

A polygenic score for BMI (BMI-PGS) was derived for each individual using variants  
279 previously and independently associated with BMI (n=656) at a p-value less than  $1 \times 10^{-8}$  from  
the study by Yengo et al (60). To ensure the integrity of the PGS only those genetic variants,  
in the NEO genotype dosage data set, where the effect allele, the alternative allele, and the  
282 minor allele frequency ( $\pm 0.1$ ) matched the data from Yengo et al were retained. If a match  
could not be made, then that SNP was dropped. All effect estimates were aligned to be  
positive, such that the effect allele was always that which increased an individual's genotype  
285 predicted BMI. Finally, a weighted PGS was constructed by weighting the number of BMI  
increasing alleles, at each SNP (n=646, after data harmonization), by the effect estimate for  
that SNP and then summing across all values (60). The R script "01\_generate\_bmi\_grs.R" for

288 deriving the PGS can be found in the github repository  
[https://github.com/hughesevoanth/NEO\\_BMI\\_Metabolite\\_MR](https://github.com/hughesevoanth/NEO_BMI_Metabolite_MR).

### Observational and Mendelian randomization analysis

291 Two observational association analyses were undertaken in cross-section. The first was an  
association analysis between individual metabolite traits and BMI. The second uses a PGS as  
a proxy instrument for the exposure in a Mendelian randomization framework to provide an  
294 estimate of the causal effect of the exposure on outcome (here dietary state specific metabolite  
values). We will refer to the first as an observational analysis and the later as an MR analysis.

297 Observational analyses were performed by a multivariable generalized linear regression with  
sampling data, sub-population, age, and sex as additional covariables - taking the form of  
 $\text{glm}(\text{metabolite} \sim \text{visit date} + \text{sub-population} + \text{age} + \text{sex} + \text{BMI})$ . For each generalized linear  
300 model we also perform a Breusch-Pagan test of homoskedasticity using the `bptest()` function  
from the `lmtest` R package (61). In addition, for each association analysis we performed a  
Type I analysis of variance or ANOVA producing a table of deviances and estimated an eta-  
303 squared statistic ( $\eta^2$ ), providing an estimate of the variance explained, for the model and for  
the primary exposure, BMI (**S3 File**).

306 One-sample Mendelian randomization analyses were performed using the `ivreg` R package  
(<https://cran.r-project.org/web/packages/ivreg/>), which implements a two-stage least squares  
instrumental variable analysis. In all instances and given the relative viability of BMI  
309 instruments versus reciprocal metabolite instrumentation, BMI was defined as our exposure,  
the PGS described above is our instrument for BMI and each metabolite trait was iteratively  
defined as the outcome. The same model and covariables as described in the association  
312 analysis paragraph above was used here. Along with MR effect estimates, we also report the  
Breusch-Pagan test of homoskedasticity, the F-statistic and F-test p-value testing for weak  
instruments, and summary statistics for the Durbin-Wu-Hausman endogeneity test (**S3 File**).

315  
Unless stated otherwise - because our outcomes are rank normal transformed (zero centered  
with a standard deviation of one) - the effect estimates (beta) and standard errors (se) are  
318 reported as rank normalized standard deviation units change per unit increase in BMI in  
kilograms per meter-squared ( $\text{kg}/\text{m}^2$ ).

## 321 **Study structure**

In all observational and MR analyses, four (sub-)population analyses were performed. The primary analysis, in all instances, was a weighted NEO (wNEO) analysis using data from all available samples. The wNEO analyses included weights for each sample that were previously derived to have the Leiden sub-population BMI distribution emulate that of the randomly sampled Leiderdorp sub-population (62). Sensitivity analyses include (1) a Leiderdorp sub-population analysis, (2) an un-weighted NEO analysis, and (3) a Leiden sub-population analysis. In addition, sex specific analyses were carried out using the primary weighted NEO framework and the Leiderdorp sub-population, as neither have a sampling bias in their BMI distributions.

## **Code availability**

All statistical analysis were conducted in the R language (v 4.0.2, Taking Off Again), and all bespoke functions and analytical code for the study can be found in the github repository [https://github.com/hughesevoanth/NEO\\_BMI\\_Metabolite\\_MR](https://github.com/hughesevoanth/NEO_BMI_Metabolite_MR).

## **Results**

### 336 **NEO cohort description**

Following quality control and derivation of the response trait the data set consisted of 5517 individuals (51.6% female), 687 metabolite traits (229 fasting, 229 postprandial, 229 response) and 85 covariables. The average age of participants was 56 years, with an average BMI of 29.98 kg/m<sup>2</sup> (**Table 1**).

The NEO cohort is made up of two distinct sub-samples, (1: “Leiderdorp”) a randomly sampled population of 1,406 individuals from the municipality of Leiderdorp (mean BMI = 26.24 kg/m<sup>2</sup>, 95% quantile interval (QI)19.94-37.2), and (2: “Leiden”) a sample population of 4,111 participant from the city of Leiden who were oversampled for individuals with a BMI greater than 27 kg/m<sup>2</sup> (mean BMI = 31.26 kg/m<sup>2</sup>, 95% QI 26.02-42.18). BMI distributions differed between the two sub-populations (generalized linear model (glm) Student’s t test p-value = 3.66x10<sup>-280</sup>, **S1 Fig.**). In addition, the two sub-populations also differ in sex ratio, alcohol intake (g/day), smoking (packyears), educational attainment and BMI polygenic score (BMI-PGS; generalized linear model Student’s t test p-value < 1.95x10<sup>-2</sup>), but not in age or height (**Table 1**). Whilst the variance explained by sub-populations is large for BMI (20.7%) and closely related traits (weight, hip and waist circumference), it is an order smaller for other

354 variables like BMI-PGS, packyears, and educational attainment (variance explained by sub-  
population < 2%, **Table 1**).

### Description of metabolite data

357 Of the 229 assayed metabolites 98 are lipoproteins and their particle concentrations and an  
additional 70 are lipoprotein ratios. The expected lipid concentrations were observed across  
lipoprotein classes (**S4 Fig.**). Namely, triglycerides dominate VLDLs, cholesterol dominates  
IDL and LDLs, and phospholipids and cholesterol dominate HDLs, on average. This remains  
360 true in both the fasting and postprandial data. All but 14 metabolites differed in mean  
concentration or ratio between the fasted and postprandial states (paired t-test,  $P < 0.05/229$ ,  
**Table S3** in **S1 File**). These 14 include the concentration of particles: total lipids, total  
363 cholesterol, cholesterol esters, and free cholesterol in medium LDL. It also includes total  
cholesterol in LDL, large LDL, and small LDL, and cholesterol esters in small LDL. Overall,  
126 metabolites increased in abundance, and 89 decreased postprandially (**Table S3** in **S1**  
366 **File**). Those with the largest scaled and centred, mean increase were the amino acids tyrosine,  
leucine, valine, isoleucine, and phenylalanine (**S5 Fig.**). Those with the largest mean decrease  
were the ratio of saturated fatty acids to total fatty acids, the ratio of triglycerides to total  
369 lipids in very large VLDL, the ratio of phospholipids to total lipids in small VLDL and the  
ketone body 3-hydroxybutyrate (**Table S3** in **S1 File**). For all metabolites, 150-minute  
postprandial change was not in the same direction for all participants (**S5 Fig.**).

372

### The observational effect of BMI on metabolite trait variation

wNEO observational analyses suggested broad association between BMI and metabolite  
375 traits, with 473 metabolite traits (across fasting, postprandial, and response states) showing  
evidence of association with BMI ( $P < 1.163 \times 10^{-3}$ , **Fig. 2**, **S6 Fig**, and **S3 File**). Effect  
estimates correlated between sexes (Pearson's  $r = 0.965$ ), but more associations were  
378 observed in women (426 total: 162 fasting, 175 postprandial, 89 response) than men (369  
total: 161 fasting, 168 postprandial, 40 response, **Fig. 3A**). A total of 333 associations were  
shared between men and women, with 93 specific to women, 36 specific to men, and 225  
381 metabolite traits having no association with BMI in either sex (**S3 File**). Among the  
associated traits all but 10 are directionally consistent among men and women (**Fig. 3A**).  
Larger effect sizes were observed in men (mean absolute beta = 0.0356) than women (mean  
384 beta = 0.0291), across all traits on average (**Fig. 3A**), but the variance explained by BMI,

across all traits on average was similar among men and women ( $\eta^2$ : men = 1.737%, women = 1.734%).

387

#### Observational results - fasting

A total of 175 or 76% of fasting traits showed evidence of association with BMI. The  
390 strongest association observed was with the inflammation marker glycoprotein acetyls (beta = 0.081, se = 0.0037,  $P = 2.89 \times 10^{-101}$ ). This was followed closely by the branched chain amino acids isoleucine (beta = 0.072, se = 0.0033,  $P = 4.18 \times 10^{-101}$ ), leucine (beta = 0.066, se =  
393 0.0031,  $P = 1.16 \times 10^{-94}$ ) and valine (beta = 0.067, se = 0.0032,  $P = 8.70 \times 10^{-92}$ ). The strongest inverse associations were observed for average diameter of HDL particles (beta = -0.071, se = 0.0035,  $P = 6.74 \times 10^{-95}$ ), the ratio of free cholesterol to total lipids in IDL (beta = -0.078, se =  
396 0.0036,  $P = 1.04 \times 10^{-100}$ ) and large HDL (beta = -0.077, se = 0.0037,  $P = 8.09 \times 10^{-95}$ ). If data is summarized by metabolite annotation class (**Table S1** in **S1 File**), we find that the classes with the largest average absolute BMI effect are inflammation (average absolute beta =  
399 0.081), amino acids (0.051), lipoprotein particle size (0.049) and glycolysis related metabolites (0.049). Those with the smallest average absolute effect are ketone bodies (0.021), fluid balance (0.019) and fatty acids (0.018) (**S6 Fig.**).

402

#### Observational results - postprandial

A total of 188 or 82% of postprandial traits showed evidence of association with BMI. Fasting  
405 and postprandial point estimates correlate strongly (Pearson's  $r = 0.976$ ) with each other, with no mean difference in their distributions (t-test  $P = 0.83$ ) and 166 shared associations. However, 31 associations are unique to one of the two dietary states, with nine associations  
408 specific to fasting and 22 specific to postprandial data (**Fig. 3C**). We tested for a difference in effect estimates between the dietary states and observed 14 estimates to differ (z-test,  $P < 0.05/229$ , **Fig. 3C**). Eight of those 14 are amino acids (tyrosine, alanine, phenylalanine,  
411 histidine, valine, leucine, glutamine, and isoleucine). The other six that differ between dietary states include four very large HDL ratios, the ketone beta-hydroxybutyrate (bohbut) and the glycolysis related metabolite lactate (lac).

414

#### Observational results - response

A total of 70 response traits showed evidence of a positive association with BMI and 40 traits  
417 had an inverse association. Response and fasting (Pearson's  $r = 0.34$ ,  $P = 1.39 \times 10^{-7}$ ), and response and postprandial (Pearson's  $r = 0.514$ ,  $P = 7.96 \times 10^{-17}$ ) point estimates modestly

correlate with each other, indicative of the relative independence of the response trait. The  
420 strongest associations were all amino acids, which (ordered by association P-value: tyrosine,  
alanine, phenylalanine, leucine valine, glutamine, histidine, and leucine) were all inversely  
associated with BMI (**S6 Fig.**). For example, while alanine levels (fasting mean = 0.359  
423 mmol/L) were elevated after a liquid mixed meal (postprandial mean = 0.376 mmol/L, paired  
t-test  $P = 2.53 \times 10^{-181}$ ) the effect attenuated as BMI increased (beta = -0.056, se = 0.004,  $P =$   
 $5.72 \times 10^{-51}$ , **Fig. 4a**) such that the effect of BMI on alanine response was negative (**Fig. 4b**).  
426 This relationship was observed for all amino acids in the data set (**Fig. 2**).

The strongest positive response association was observed for triglycerides in very large HDL  
429 (xlhdltg: beta = 0.037, se = 0.004,  $P = 5.88 \times 10^{-21}$ ). This was followed closely by the ratio of  
(1) cholesterol esters, (2) triglycerides, and (3) total cholesterol to total lipids in very large  
HDL. Conversely, the response trait ratio of phospholipids to total lipids in very large HDL  
432 decreased as BMI increased. This suggests that following a liquid mixed meal the increase of  
phospholipids relative to total lipids in very large HDL is attenuated by BMI and conversely,  
the ratio of cholesterol and triglycerides to total lipids in very large HDL increased as BMI  
435 increased (**Fig. 2**).

Outside of amino acids and lipoproteins, other metabolite response traits that had a positive  
438 association with BMI included 3-hydroxybutyrate (ketone body), creatinine (fluid balance),  
linoleic acid, and total fatty acids (fatty acids). Others with an inverse association with BMI  
included glucose and lactate (glycolysis), the ratio of linoleic acid to total fatty acids, and the  
441 ratio of omega-6 fatty acids to total fatty acids (fatty acid ratios; **S3 File**).

### Possible confounders

444 The instrumental variable used, BMI-PGS, explains 4.7% of the variation in BMI, and for  
each unit increase in BMI-PGS there is a 3.26 (se = 0.20) unit ( $\text{kg/m}^2$ ) increase in BMI ( $P =$   
 $3.61 \times 10^{-60}$ , wNEO analysis, **Table 1**). An assessment of the correlation between BMI and  
447 BMI-PGS with study covariables was performed to identify possible confounders, specifically  
those that would violate the second MR assumption (exchangeability). Of the 91 tested  
covariables, 60 were associated with BMI (wNEO,  $P < 9.26 \times 10^{-4}$ ; **Table S4** in **S1 File** and **S7**  
450 **Fig.**), and 17 were associated with BMI-PGS (**S8 Fig.**). Those 17 that associated with BMI-  
PGS included other adiposity traits ( $n = 5$ ), sub-population, visit date, genetic principal

component 3, (imputed) smoking packyears, on a weight loss diet, type of diet, on a diet last  
453 month, a food rule, mean CO<sub>2</sub> production, and a peroxidation sample quality flag. The 17  
covariables associated the BMI-PGS were tested for an association with each metabolite trait  
(the outcomes) in a univariable linear model. Results illustrated broad association across all  
456 traits and covariables, defining each covariable as a confounder in MR analysis leading to  
sensitivity analyses discussed below (**S9 Fig.**). A more comprehensive description of these  
analyses is in **S1 Text**.

459

### Estimating the causal effect of BMI on metabolite trait variation

MR analyses (wNEO) support broad associations between BMI and metabolites as suggested  
462 by the overall strong correlation between observational and MR estimates (Pearson's  $r =$   
 $0.852$ ,  $P = 1.02 \times 10^{-194}$ , intercept =  $1.54 \times 10^{-03}$ , slope =  $0.911$ ). The correlation was strongest  
for postprandial traits (Pearson's  $r = 0.921$ ,  $P = 4.44 \times 10^{-95}$ , intercept =  $3.59 \times 10^{-03}$ , slope =  
465  $0.922$ ) and weaker for fasting (Pearson's  $r = 0.871$ ,  $P = 7.71 \times 10^{-72}$ , intercept =  $7.62 \times 10^{-03}$ ,  
slope =  $0.855$ ) and response traits (Pearson's  $r = 0.62$ ,  $P = 1.11 \times 10^{-25}$ , intercept =  $-6.36 \times 10^{-03}$ ,  
slope =  $1.08$ , **Figure 3D**). Forty-two observational and MR effect estimates showed evidence  
468 of being different from each other (z-test  $P < 0.05$ ), yet none remain so after correcting for  
multiple tests ( $P < 0.05/687$ ). Overall, these observations indicate that cross-sectional  
estimates have reasonable power at predicting MR estimates for these exposure-outcome  
471 analyses, consistent with previous work (18).

A total of 201 metabolite traits (across fasting, postprandial, and response states; **Figure 5, S3**  
474 **File**) showed nominal evidence of association ( $P < 0.05$ ) with BMI and 20 (**Table 2, Figure**  
**6**) showed evidence of association with BMI ( $P < 1.16 \times 10^{-3}$ ; **S10 Fig.**). MR effect estimates  
varied more between sexes (Pearson's  $r = 0.368$ ,  $p = 1.93 \times 10^{-23}$ ) than observed in  
477 observational results, with 86 showing evidence of being different from each other (z-test  $P <$   
 $0.05$ ), yet none remaining so after correcting for multiple tests ( $P < 0.05/687$ ). Nine metabolite  
traits (7 fasting, 2 response) showed evidence of association with BMI in females, eight of  
480 which were unique to females (**Figure 6**). This included fasting small HDL triglycerides  
(shdltg\_f) and fasting very small VLDL triglycerides (xsvldltg\_f), both of which increased  
with increases in BMI. In contrast, five metabolites were associated with BMI in men, one of  
483 which was unique to men – postprandial glutamine, which had an inverse relationship with  
BMI (gln\_p; **Figure 6**). The other four metabolites shared an association with the general



population. They are, fasting and postprandial citrate (cit\_f, cit\_p), and the ratio of linoleic  
486 acid to total fatty acids in both the fasting and postprandial state (lafa\_f, lafa\_p). Each of  
which were also inversely associated with BMI. On average MR effect estimates were larger  
in females (absolute mean beta = 0.047) than males (absolute mean beta = 0.026, paired t-test  
489  $P = 1.43 \times 10^{-45}$ ).

#### MR results - fasting

492 A total of 87 fasting traits exhibited nominal evidence of an association with BMI in the  
general population, five remained after correction for multiple testing (**Table 2, Figure 6, S3  
File**). The strongest association observed was with the ratio of linoleic acid to total fatty acids  
495 (lafa; beta = -0.106, se = 0.025,  $P = 1.80 \times 10^{-5}$ ). This was followed closely by 4 other inverse  
associations: the ratio polyunsaturated fatty acids (pufafa; beta = -0.081, se = 0.024,  $P =$   
6.80x10<sup>-4</sup>) to total fatty acids, the ratio omega-6 to total fatty acids (faw6fa; beta = -0.081, se  
498 = 0.024,  $P = 6.97 \times 10^{-4}$ ) and the abundance of citrate (cit; beta = -0.079, se = 0.024,  $P =$   
8.63x10<sup>-4</sup>). In addition, one fasting metabolite showed evidence of a positive association with  
BMI, the branched-chain amino acid leucine (leu; beta = 0.068, se = 0.020,  $P = 5.62 \times 10^{-4}$ ). If  
501 data is summarized by metabolite annotation class (**Table S1 in S1 File**), we find that the  
classes with the largest average absolute BMI effect are ketone bodies (average beta = -0.047,  
n = 2), amino acids (0.036, n = 8), inflammation (0.029, n = 1), glycerides and phospholipids  
504 (0.027, n = 9) and lipoprotein subclasses (0.023, n = 98). Those with the smallest average  
absolute effect are lipoprotein particle size (0.009, n = 3), lipoprotein subclass ratios (0.003, n  
= 70), and cholesterol (0.002, n = 9) (**S10 Fig.**).

507

#### MR results - postprandial

A total of 96 postprandial metabolites exhibited nominal evidence of an association with BMI,  
510 12 remained after correcting for multiple testing (**Table 2, Figure 6, S3 File**). Like the fasting  
metabolite traits the fatty acid ratios linoleic acid (lafa; beta = -0.092, se = 0.025,  $P = 2.0 \times 10^{-4}$ ),  
omega-6 (faw6fa; beta = -0.080, se = 0.024,  $P = 7.83 \times 10^{-4}$ ), polyunsaturated fatty acids  
513 (pufafa; beta = -0.077, se = 0.024,  $P = 1.10 \times 10^{-3}$ ) and citrate (cit; beta = -0.105, se = 0.022,  $P =$   
2.83x10<sup>-6</sup>), who had the largest effect, each also had an inverse relationship with BMI  
(**Figure 5-6**). In addition, the amino acid histidine (his), large HDL cholesterol (l.hdl.c) and  
516 free cholesterol (l.hdl.fc) decreased with increases in BMI. Further, the ratio of cholesterol  
(l.hdl.c.pct), cholesterol esters (l.hdl.ce.pct), and free cholesterol (l.hdl.fc.pct) to total lipids in  
HDL decreased with increases in BMI. Finally, the ratio of free cholesterol to total lipids in

519 IDL (idl.fc.pct) and the ratio of phospholipids to total lipids in medium VLDL (m.vldl.pl.pct) decreased with increases in BMI. If data is summarized by metabolite annotation class (**Table S1** in **S1 File**), we find that the classes with the largest average absolute BMI effect are  
522 glycolysis related metabolites (average beta = -0.046, n = 3), inflammation (0.40, n = 1) and lipoprotein subclasses (0.02, n = 98). The average effect among amino acids is not negative (-0.019, n = 8) in opposition to that seen in fasting data (**S10 Fig.**).

525

### MR results - response

A total of 18 response metabolites exhibited nominal evidence of an association with BMI,  
528 three remained after correcting for multiple testing (**Table 2, Figure 6, S3 File**). The largest effect was observed for the amino acid alanine (ala: beta = -0.081, se = 0.023, P = 5.91x10<sup>-04</sup>). Data suggested that increases in BMI attenuates the increase in abundance of the amino acid  
531 alanine after a liquid mixed meal (**Figure 5-6**) – consistent with observational analyses above (**Figure 4**). The other two response traits that associated with BMI are the ratio of phospholipids in medium VLDL to total lipids in medium VLDL (m.vldl.pl.pct; beta = 0.091, se = 0.027, P = 6.75x10<sup>-4</sup>) and the ratio of free cholesterol in small VLDL to total lipids in  
534 small VLDL (s.vldl.fc.pct; beta = 0.082, se = 0.025, P = 1.11x10<sup>-3</sup>). As a class, amino acids (average beta = -0.059, n=8) had the largest average BMI response effect estimate (**S10 Fig.**).  
537 This was followed by the ketone bodies acetate and beta-hydroxybutyrate (mean beta = 0.048, n = 2) and the glycolysis related metabolites citrate, glucose, and lactate (mean beta = -0.042, n = 3).

540

### Sub-population and sensitivity analyses

543 All observational and MR analyses were repeated in each of the two sub-populations and in the NEO cohort without the inclusion of weights. These analyses allowed us to evaluate the variability in effect estimates measured in a random (Leiderdorp) and biased (Leiden)  
546 population sample as well as the effectiveness of the weights in our primary analyses (wNEO). We found that effect estimates from the Leiderdorp and the wNEO frameworks are strongly correlated (observational Pearson's r = 0.983, MR = 0.855) and correlation  
549 coefficients weakened when compared to the un-weighted NEO and the Leiden frameworks (**S11 Figure**). Decreases in agreement between analyses as the sample population mean BMI shifts would indicate that there are either un-accounted confounders influencing the results or  
552 that the relationship between BMI and metabolite trait variation is not always linear.

To evaluate the influence of confounders on MR effect estimates we reran the association  
555 analysis in the wNEO data set. Three additional covariates were included in the model  
(smoking, on a weight loss diet, and PC3) and 101 sample with a peroxidation sample quality  
flag, a sample quality metric identified by Metabolon, were removed. Overall primary  
558 (wNEO) observational (Pearson's  $r = 0.998$ ) and MR (Pearson's  $r = 0.963$ ) effect estimates  
correlate strongly with those from the sensitivity analysis (**S7 File**). A more comprehensive  
description of these analyses is in **S1 Text**.

## 561 **Discussion**

This study provides effect estimates for observational and MR associations between BMI and  
564 metabolites in a (1) fasted, (2) postprandial, and (3) response state, using individual level data  
from a middle-aged cohort of 5517 individuals of Northern European ancestry including a  
liquid mixed meal challenge. Each dietary state provides a unique examination at metabolite  
567 variation, the value of which are actively being explored. Broad observational associations  
were observed between BMI and metabolites (69% of tested traits). In addition, MR estimates  
were largely concordant with the observational estimates (**Figure 3D**). This is consistent with  
570 previous research and suggests that, at least for this lipidomic platform, observational  
estimates are reasonably well aligned to MR effect estimates (18). This generalization holds  
less true for response traits than for postprandial and fasting trait (**Figure 3D**). After  
573 correcting for multiple testing, 20 metabolite traits maintain evidence of a causal BMI effect  
(**Figure 6**). This includes the inverse association with fasting and postprandial ratios of  
linoleic acid (lafa), omega-6 (faw6fa), and polyunsaturated fatty acids (pufafa) to total fatty  
576 acids and fasting and postprandial citrate (cit) abundance. In addition, the amino acid leucine  
has a positive MR association with BMI in the fasting state, histidine has an inverse MR  
association with BMI in the postprandial state and alanine has an inverse MR association with  
579 BMI in the response state.

We compared our observational effect estimates to fifty-seven matching fasting metabolites  
582 measured using the same platform, in a young adult cohort (18). Overall effect estimates  
correlated well (Pearson's  $r = 0.82$ ), with the caveat that estimates cannot be directly  
compared given differences in data transformations between the studies. However, there are

585 37 shared associations, 6 unique to this study, 12 unique to the other and 2 yielding no  
association in either study (**Figure 3B**). Five of the shared associations were different across  
studies (total cholesterol, sphingomyelins, phosphatidylcholines, phosphoglycerides, and  
588 linoleic acid). Among the 12 metabolites unique to existing work before that presented here,  
are 10 metabolites that previously exhibited a strong positive effect but showed no evidence  
for an association or negative point estimates here (total lipids in IDL, large LDL, medium  
591 LDL and small LDL, total cholesterol in IDL and LDL, and four fatty acid metabolites  
omega-3, omega-6, polyunsaturated fatty acids, and docosahexaenoic acid). These differences  
may be the product of non-linear associations between these metabolites and BMI in  
594 combination with age effects and differences in the distribution of BMI between these studies.  
Despite these difference, 57 fasting MR estimates that did overlap with previous work showed  
strong correlation between the two studies (Person's  $r = 0.862$ ,  $P = 7.99 \times 10^{-18}$ ; **S12 Fig.**), with  
597 one shared fasting MR association – the amino acid leucine.

Given the prevalence of lipoproteins and their lipids on this platform we looked for the  
600 expected atherogenic lipoprotein profile – that is decreases in HDL and increases in non-HDL  
lipoproteins with increase in BMI (63–66). Both observational and MR associations between  
BMI and fasting and postprandial dietary states do suggest a worsening atherogenic  
603 lipoprotein profile with BMI (**Figure 7**, **S13 Fig.**, and **S14 Fig.**). As illustrated with the fasting  
observational effect estimates in **Figure 7**, the atherogenic VLDL lipoproteins are higher with  
increases in BMI while non-atherogenic lipoproteins – or those involved in the reverse  
606 cholesterol transport system (HDL), were lower. In contrast to expectations, other atherogenic  
lipoproteins IDL and LDL, largely do not associate with BMI. In addition, triglycerides in  
non-atherogenic medium HDL and small HDL increase with increases in BMI. Further, as  
609 HDL lipoprotein density increases the inverse relationship with BMI weakens and as seen  
with small HDL the association becomes positive (**Figure 7**). This is consistent with the  
observation above that the strongest inverse association with BMI is that with average  
612 diameter of HDL particles, suggesting that as BMI increases there is a associated shift toward  
smaller HDL particle size. Along with these changes, other observations were consistent with  
obesity influencing lipid profiles that are associated with cardiometabolic disease (17,25,67).  
615 These included increases in apolipoprotein B, decreases in apolipoprotein A-1 and increases  
in the ratio of B to A-1 with increases in BMI, in both the fasting and postprandial states  
(65,68).

618

The MR association of BMI across branched chain (BCAA) and aromatic (AAA) amino acids were similar with positive effects in the fasting state, no effect in the postprandial state, and an inverse association in the response state (**Figure 2 and 5**). Meanwhile the dietary profile of (glucogenic) amino acids suggest no effect in the fasting state and inverse effects in both the postprandial and response states. These observations may suggest that BMI may have a causal influence on the synthesis or metabolism of amino acids in general, particularly regarding their inverse association with BMI in the response state (**Figure 2 and 5**). BCAAs and AAAs have long been associated with obesity, glucose, insulin (resistance), and type 2 diabetes (69–75). In support of early observations, numerous recent studies that have shown associations between amino acids and BMI, visceral adipose tissue and weight change (18,36,76–78), between BCAAs and AAAs with insulin resistance and type 2 diabetes (79–82), used MR to show that insulin resistance increase BCAAs concentration (83) and used a randomized control trail to show that restricting BCAAs can improve glucose tolerance and reduce fat accumulation (84). Indeed evidence now suggests a potential causal pathway from BMI to type 2 diabetes via the intermediate traits of insulin resistance and BCAAs, respectively (83). Observations here reinforce the causal association between BMI and amino acids, with the additional observations that postprandial abundance of the glucogenic amino acids decreases with BMI and that the relative change in amino acid abundance in response to a liquid mixed meal decreases with increases in BMI.

As the study was centered in the Netherlands, study samples were limited to individuals of Northern European ancestry and as such inferences made from results should be limited to populations of similar ancestry and environment. Sampling was limited to individuals of middle age (range 44 to 66 years and a mean of 56, Table 1). Whilst this provides specificity to a middle age population it also limits the results to this population age range as well. The NEO population sampling was performed in two distinct batches, one focused on oversampling individuals of high BMI (Leiden) and a second a random and representative sampling of the population (Leiderdorp). The product of this led to the use of weighted linear regression analyses to maintain sample size and power, while also providing estimates that are representative of the study population at large. The validity of this assumption rests in the credibility of the weights, which were specifically designed to make the BMI distribution of the Leiden sample mirror that of the Leiderdorp sample (85). We have illustrated through comparison with the Leiderdorp sample that effect estimates are consistent across the two analyses, imparting support to the weighted analysis (**S11 Figure**). With these factors, there

was just a single postprandial time point evaluated here limiting the inferences that can be  
654 drawn. It is reasonable to speculate that an earlier time point, or a complete time series would  
provide additional information about how BMI might influence postprandial variation.  
Finally, while a single complex meal standardizes the complexities in evaluating metabolite  
657 response and postprandial abundances alternative meals could provide novel insights in the  
interplay between BMI, nutrition, and metabolite variation.

660 In conclusion, using a middle-aged cohort of 5517 individuals we have demonstrated that  
effect estimates for BMI on metabolite traits in the fasting, postprandial and response dietary  
states. We have been able to show that results are broadly correlated between observational  
663 and one-sample Mendelian randomization analyses and in the context of the numerous  
documented associations between BMI, metabolites and disease, this gives support to a  
conclusion that metabolites may act as intermediates – or at the least biomarkers of underlying  
666 driver physiology – between adiposity accumulation and disease. Furthermore, this work  
suggests that the dynamic metabolome may potentially flag common biological events which  
will systematically vary by BMI, which may be linked to the aetiology of disease and which  
669 are linked to life course events such as feeding and metabolic response.

## Acknowledgements

The authors thank all individuals who participated in the Netherlands Epidemiology of  
672 Obesity study, and all participating general practitioners for inviting eligible participants. We  
also thank the NEO study group, Pat van Beelen and all research nurses for data collection,  
Petra Noordijk and her team for laboratory management, and Ingeborg de Jonge for data  
675 management of the NEO study. We acknowledge the support from the Netherlands  
Cardiovascular Research Initiative: an initiative with support of the Dutch Heart Foundation  
(CVON2014-02 ENERGISE). We thank Nutricia Research, Utrecht, The Netherlands, for  
678 providing the liquid mixed meal. Nicholas J. Timpson is a Wellcome Trust Investigator  
(202802/Z/16/Z), is the PI of the Avon Longitudinal Study of Parents and Children (MRC &  
WT 102215/2/13/2), is supported by the University of Bristol NIHR Biomedical Research  
681 Centre (BRC-1215-20011), the MRC Integrative Epidemiology Unit (MC\_UU\_12013/3) and  
works within the CRUK Integrative Cancer Epidemiology Programme (C18281/A19169).  
David A. Hughes is supported by N.J.Timpson's Wellcome Investigator Award  
684 (202802/Z/16/Z), and along with Caroline Bull is also supported by the University of Bristol

and UK Medical Research Council (MC\_UU\_00011/1 and MC\_UU\_00011/6). Dennis Mook-Kanamori is supported by Dutch Science Organization (ZonMW-VENI Grant 916.14.023)

## References

- 690 1.     Bentham J, Di Cesare M, Bilano V, Bixby H, Zhou B, Stevens GA, et al. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. *Lancet*. 2017;390(10113):2627–42.
- 695 2.     Yanovski JA. Obesity: Trends in underweight and obesity - scale of the problem. *Nat Rev Endocrinol* [Internet]. 2018;14(1):5–6. Available from: <http://dx.doi.org/10.1038/nrendo.2017.157>
3.     World Health Organization. World Health Organization. Obesity and Overweight Fact Sheet. [Internet]. 2021 [cited 2022 Mar 1]. Available from: <https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight>
- 700 4.     Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults--The Evidence Report. National Institutes of Health. *Obes Res*. 1998 Sep;6 Suppl 2:51S-209S.
- 705 5.     Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *J Am Coll Cardiol*. 2014 Jul;63(25 Pt B):2985–3023.
6.     Okorodudu DO, Jumean MF, Montori VM, Romero-Corral A, Somers VK, Erwin PJ, et al. Diagnostic performance of body mass index to identify obesity as defined by body adiposity: a systematic review and meta-analysis. *Int J Obes (Lond)*. 2010 May;34(5):791–9.
- 710 7.     Reilly JJ, Kelly J, Wilson DC. Accuracy of simple clinical and epidemiological definitions of childhood obesity: systematic review and evidence appraisal. *Obes Rev an Off J Int Assoc Study Obes*. 2010 Sep;11(9):645–55.
8.     Javed A, Jumean M, Murad MH, Okorodudu D, Kumar S, Somers VK, et al. Diagnostic performance of body mass index to identify obesity as defined by body adiposity in children and adolescents: A systematic review and meta-analysis. *Pediatr Obes*. 2015;10(3):234–44.
- 715 9.     Reilly JJ, El-Hamdouchi A, Diouf A, Monyeki A, Somda SA. Determining the worldwide prevalence of obesity. *Lancet* [Internet]. 2018;391(10132):1773–4. Available from: [http://dx.doi.org/10.1016/S0140-6736\(18\)30794-3](http://dx.doi.org/10.1016/S0140-6736(18)30794-3)
- 720 10.    Wade KH, Carlslake D, Tynelius P, Smith GD, Martin RM. Variation of all-cause and cause-specific mortality with body mass index in one million Swedish parent-son pairs: An instrumental variable analysis. *PLoS Med*. 2019;16(8):1–15.
11.    Jenkins DA, Wade KH, Carlslake D, Bowden J, Sattar N, Loos RJF, et al. Estimating the causal effect of BMI on mortality risk in people with heart disease, diabetes and cancer using Mendelian randomization. *Int J Cardiol* [Internet]. 2021;330:214–20. Available from: <https://doi.org/10.1016/j.ijcard.2021.02.027>
- 725 12.    Solomon CG, Manson JE. Obesity and mortality: a review of the epidemiologic data. *Am J Clin Nutr*. 1997 Oct;66(4 Suppl):1044S-1050S.



13. Reis JP, Macera CA, Araneta MR, Lindsay SP, Marshall SJ, Wingard DL. Comparison of overall obesity and body fat distribution in predicting risk of mortality. *Obesity (Silver Spring)*. 2009 Jun;17(6):1232–9.
- 730 14. Pearson-Stuttard J, Zhou B, Kontis V, Bentham J, Gunter MJ, Ezzati M. Worldwide burden of cancer attributable to diabetes and high body-mass index: a comparative risk assessment. *lancet Diabetes Endocrinol*. 2018 Jun;6(6):e6–15.
15. Lauby-Secretan B, Scocciati C, Loomis D, Grosse Y, Bianchini F, Straif K. Body Fatness and Cancer--Viewpoint of the IARC Working Group. *N Engl J Med*. 2016 Aug;375(8):794–8.
- 735 16. Sung H, Siegel RL, Torre LA, Pearson-Stuttard J, Islami F, Fedewa SA, et al. Global patterns in excess body weight and the associated cancer burden. *CA Cancer J Clin*. 2018;88–112.
17. Powell-Wiley TM, Poirier P, Burke LE, Després J-PP, Gordon-Larsen P, Lavie CJ, et al. Obesity and Cardiovascular Disease: A Scientific Statement from the American Heart Association. *Circulation*. 2021 May;143(21):E984–1010.
- 740 18. Würtz P, Wang Q, Kangas AJ, Richmond RC, Skarp J, Tiainen M, et al. Metabolic Signatures of Adiposity in Young Adults: Mendelian Randomization Analysis and Effects of Weight Change. Sheehan NA, editor. *PLoS Med* [Internet]. 2014 Dec 9;11(12):e1001765. Available from: <https://dx.plos.org/10.1371/journal.pmed.1001765>
- 745 19. Würtz P, Havulinna AS, Soininen P, Tynkkynen T, Prieto-Merino D, Tillin T, et al. Metabolite Profiling and Cardiovascular Event Risk: A Prospective Study of 3 Population-Based Cohorts. *Physiol Behav*. 2016 Mar;176(1):139–48.
20. von Krüchten R, Lorbeer R, Müller-Peltzer K, Rospleszcz S, Storz C, Askani E, et al. Association between Adipose Tissue Depots and Dyslipidemia: The KORA-MRI Population-Based Study. *Nutrients*. 2022;14(4).
- 750 21. Ference BA, Yoo W, Alesh I, Mahajan N, Mirowska KK, Mewada A, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: A Mendelian randomization analysis. *Ration Pharmacother Cardiol* [Internet]. 2012/10/23. 2013;9(1):90–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23083789>
- 755 22. Krittanawong C, Khawaja M, Rosenson RS, Amos CI, Nambi V, Lavie CJ, et al. Association of PCSK9 Variants With the Risk of Atherosclerotic Cardiovascular Disease and Variable Responses to PCSK9 Inhibitor Therapy. *Curr Probl Cardiol*. 2021 Nov;101043.
23. Wilt TJ, Bloomfield HE, MacDonald R, Nelson D, Rutks I, Ho M, et al. Effectiveness of statin therapy in adults with coronary heart disease. *Arch Intern Med*. 2004 Jul;164(13):1427–36.
- 760 24. Berliner JA, Navab M, Fogelman AM, Frank JS, Demer LL, Edwards PA, et al. Atherosclerosis: basic mechanisms. Oxidation, inflammation, and genetics. *Circulation*. 1995 May;91(9):2488–96.
- 765 25. Richardson TG, Sanderson E, Palmerid TM, Korpelaid MA, Ference BA, Smith GD, et al. Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS Med* [Internet]. 2020;17(3):1–22. Available from: <http://dx.doi.org/10.1371/journal.pmed.1003062>
26. Ala-Korpela M. The culprit is the carrier, not the loads: Cholesterol, triglycerides and apolipoprotein B in atherosclerosis and coronary heart disease. *Int J Epidemiol*.

- 2019;48(5):1389–92.
- 770 27. Corbin LJ, Tan VY, Hughes DA, Wade KH, Paul DS, Tansey KE, et al. Formalising recall by genotype as an efficient approach to detailed phenotyping and causal inference. *Nat Commun.* 2018;9(1).
28. Fang S, Wade KH, Hughes DA, Fitzgibbon S, Yip V, Timpson NJ, et al. A multivariant recall-by-genotype study of the metabolomic signature of BMI. *Obesity [Internet].* 2022 Jun 22 [cited 2023 Feb 27];30(6):1298–310. Available from: <https://pubmed.ncbi.nlm.nih.gov/35598895/>
- 775 29. Moreton JR. Chylomicronemia, fat tolerance, and atherosclerosis. *Int J Epidemiol.* 2016;45(5):1372–9.
30. Karimpour M, Surowiec I, Wu J, Gouveia-Figueira S, Pinto R, Trygg J, et al. Postprandial metabolomics: A pilot mass spectrometry and NMR study of the human plasma metabolome in response to a challenge meal. *Anal Chim Acta [Internet].* 2016;908:121–31. Available from: <http://dx.doi.org/10.1016/j.aca.2015.12.009>
- 780 31. Berry SE, Valdes AM, Drew DA, Asnicar F, Mazidi M, Wolf J, et al. Human postprandial responses to food and potential for precision nutrition. *Nat Med [Internet].* 2020;26(6):964–73. Available from: <https://doi.org/10.1038/s41591-020-0934-0>
- 785 32. Wang Q, Jokelainen J, Auvinen J, Puukka K, Keinänen-Kiukaanniemi S, Järvelin MR, et al. Insulin resistance and systemic metabolic changes in oral glucose tolerance test in 5340 individuals: An interventional study. *BMC Med.* 2019;17(1):1–12.
33. Bütikofer U, Burnand D, Portmann R, Blaser C, Schwander F, Kopf-Bolanz KA, et al. Serum metabolites responding in a dose-dependent manner to the intake of a high-fat meal in normal weight healthy men are associated with obesity. *Metabolites.* 2021;11(6).
- 790 34. Shrestha A, Müllner E, Poutanen K, Mykkänen H, Moazzami AA. Metabolic changes in serum metabolome in response to a meal. *Eur J Nutr.* 2017;56(2):671–81.
35. Labarre JL, Hirschfeld E, Soni T, Kachman M, Wigginton J, Duren W, et al. Comparing the fasting and random-fed metabolome response to an oral glucose tolerance test in children and adolescents: Implications of sex, obesity, and insulin resistance. *Nutrients.* 2021;13(10):1–17.
- 795 36. Schutte BAM, Van Den Akker EB, Deelen J, Van De Rest O, Van Heemst D, Feskens EJM, et al. The effect of standardized food intake on the association between BMI and 1 H-NMR metabolites. *Sci Rep [Internet].* 2016;6(December):1–6. Available from: <http://dx.doi.org/10.1038/srep38980>
- 800 37. Stumvoll M, Mitrakou A, Pimenta W, Jenssen T, Yki-Järvinen H, Van Haeflten T, et al. Use of the oral glucose tolerance test to assess insulin release and insulin sensitivity. *Diabetes Care [Internet].* 2000 [cited 2023 Jan 16];23(3):295–301. Available from: <https://pubmed.ncbi.nlm.nih.gov/10868854/>
- 805 38. Li-Gao R, Hughes DA, van Klinken JB, de Mutsert R, Rosendaal FR, Mook-Kanamori DO, et al. Genetic Studies of Metabolomics Change After a Liquid Meal Illuminate Novel Pathways for Glucose and Lipid Metabolism. *Diabetes [Internet].* 2021 Dec 1;70(12):2932–46. Available from: <https://diabetesjournals.org/diabetes/article/70/12/2932/139064/Genetic-Studies-of-Metabolomics-Change-After-a>
- 810 39. Wade KH, Lawlor D, Borges MC, Palmer T, Hartwig FP, Hemani G, et al. MR Dictionary

- [Internet]. 2021 [cited 2022 Mar 1]. Available from: <https://mr-dictionary.mrcieu.ac.uk>
40. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23(R1):R89-98.
- 815 41. Bull CJ, Bell JA, Murphy N, Sanderson E, Davey Smith G, Timpson NJ, et al. Adiposity, metabolites, and colorectal cancer risk: Mendelian randomization study. *BMC Med* [Internet]. 2020 Dec 1 [cited 2023 Jul 10];18(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/33327948/>
- 820 42. Kalaoja M, Corbin LJ, Tan VY, Ahola-Olli A V., Havulinna AS, Santalahti K, et al. The Role of Inflammatory Cytokines as Intermediates in the Pathway from Increased Adiposity to Disease. *Obesity (Silver Spring)* [Internet]. 2021 Feb 1 [cited 2023 Jul 10];29(2):428–37. Available from: <https://pubmed.ncbi.nlm.nih.gov/33491305/>
- 825 43. Fang S, Wade KH, Hughes DA, Fitzgibbon S, Yip V, Timpson NJ, et al. A multivariant recall-by-genotype study of the metabolomic signature of BMI. *Obesity (Silver Spring)* [Internet]. 2022 Jun 1 [cited 2023 Jul 10];30(6):1298–310. Available from: <https://pubmed.ncbi.nlm.nih.gov/35598895/>
44. Bell JA, Richardson TG, Wang Q, Sanderson E, Palmer T, Walker V, et al. Effects of general and central adiposity on circulating lipoprotein, lipid, and metabolite levels in UK Biobank: A multivariable Mendelian randomization study. *Lancet Reg Heal Eur* [Internet]. 2022 Oct 1 [cited 2023 Jul 10];21. Available from: <https://pubmed.ncbi.nlm.nih.gov/35832062/>
- 830 45. Bell JA, Carlslake D, O’Keeffe LM, Frysz M, Howe LD, Hamer M, et al. Associations of Body Mass and Fat Indexes With Cardiometabolic Traits. *J Am Coll Cardiol*. 2018 Dec 18;72(24):3142–54.
46. Wade KH, Carlslake D, Sattar N, Davey Smith G, Timpson NJ. BMI and Mortality in UK Biobank: Revised Estimates Using Mendelian Randomization. *Obesity*. 2018 Nov 1;26(11):1796–806.
- 835 47. Goudswaard LJ, Bell JA, Hughes DA, Corbin LJ, Walter K, Davey Smith G, et al. Effects of adiposity on the human plasma proteome: observational and Mendelian randomisation estimates. *Int J Obes* [Internet]. 2021 Oct 1 [cited 2023 Jul 10];45(10):2221–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/34226637/>
- 840 48. Wade KH, Chiesa ST, Hughes AD, Chaturvedi N, Charakida M, Rapala A, et al. Assessing the Causal Role of Body Mass Index on Cardiovascular Health in Young Adults: Mendelian Randomization and Recall-by-Genotype Analyses. *Circulation* [Internet]. 2018 Nov 13 [cited 2023 Jul 10];138(20):2187–201. Available from: <https://pubmed.ncbi.nlm.nih.gov/30524135/>
- 845 49. Timpson NJ, Nordestgaard BG, Harbord RM, Zacho J, Frayling TM, Tybjærg-Hansen A, et al. C-reactive protein levels and body mass index: Elucidating direction of causation through reciprocal Mendelian randomization. *Int J Obes* [Internet]. 2011 Feb [cited 2023 Jul 10];35(2):300–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/20714329/>
- 850 50. Tan VY, Timpson NJ. The UK Biobank: A Shining Example of Genome-Wide Association Study Science with the Power to Detect the Murky Complications of Real-World Epidemiology [Internet]. Vol. 23, Annual Review of Genomics and Human Genetics. *Annu Rev Genomics Hum Genet*; 2022 [cited 2023 Feb 20]. p. 569–89. Available from: <https://pubmed.ncbi.nlm.nih.gov/35508184/>
51. de Mutsert R, den Heijer M, Rabelink TJ, Smit JW, Romijn JA, Jukema JW, et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. *Eur J Epidemiol* [Internet]. 2013;28(6):513–23. Available from:

<https://www.ncbi.nlm.nih.gov/pubmed/23576214>

- 855 52. de Mutsert R, den Heijer M, Rabelink TJ, Smit JWA, Romijn JA, Jukema JW, et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. *Eur J Epidemiol* [Internet]. 2013 Jun;28(6):513–23. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/23576214>
- 860 53. Soininen P, Kangas AJ, Würtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ Cardiovasc Genet*. 2015 Feb;8(1):192–206.
54. Hughes DA, Taylor K, McBride N, Lee MA, Mason D, Lawlor DA, et al. Metaboprep: an R package for pre-analysis data description and processing. *Bioinformatics*. 2022 Feb;38(7):1980–7.
- 865 55. Cheverud JM. A simple correction for multiple comparisons in interval mapping genome scans. *Heredity (Edinb)*. 2001;87(1):52–8.
56. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)* [Internet]. 2005 Sep 3 [cited 2019 Jun 24];95(3):221–7. Available from: <http://www.nature.com/articles/6800717>
- 870 57. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol* [Internet]. 2008 May 1 [cited 2020 Feb 19];32(4):361–9. Available from: <http://doi.wiley.com/10.1002/gepi.20310>
- 875 58. Blauw LL, Li-Gao R, Noordam R, de Mutsert R, Trompet S, Berbee JFP, et al. CETP (Cholesteryl Ester Transfer Protein) Concentration: A Genome-Wide Association Study Followed by Mendelian Randomization on Coronary Artery Disease. *Circ Genomic Precis Med*. 2018 May;11(5):e002034.
- 880 59. Consortium the HR, McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* [Internet]. 2016 Oct 22 [cited 2019 Jan 25];48(10):1279–83. Available from: <http://www.nature.com/articles/ng.3643>
- 885 60. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of genome-wide association studies for height and body mass index in approximately 700000 individuals of European ancestry. *Hum Mol Genet* [Internet]. 2018/08/21. 2018;27(20):3641–9. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/30124842>
61. Zeileis A, Hothorn T. Diagnostic Checking in Regression Relationships. *R News* [Internet]. 2002;2(3):7–10. Available from: <https://cran.r-project.org/doc/Rnews/>
- 890 62. Dekkers IA, de Mutsert R, de Vries APJ, Rosendaal FR, Cannegieter SC, Jukema JW, et al. Determinants of impaired renal and vascular function are associated with elevated levels of procoagulant factors in the general population. *J Thromb Haemost* [Internet]. 2018 Mar 1 [cited 2022 Aug 23];16(3):519–28. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/jth.13935>
63. Nordestgaard BG. A Test in Context: Lipid Profile, Fasting Versus Nonfasting. Vol. 70, *Journal of the American College of Cardiology*. Elsevier; 2017. p. 1637–46.
- 895 64. Arsenault BJ, Rana JS, Stroes ESG, Després J-P, Shah PK, Kastelein JJP, et al. Beyond low-

density lipoprotein cholesterol: respective contributions of non-high-density lipoprotein cholesterol levels, triglycerides, and the total cholesterol/high-density lipoprotein cholesterol ratio to coronary heart disease risk in apparently healthy. *J Am Coll Cardiol*. 2009 Dec;55(1):35–41.

- 900 65. Carr SS, Hooper AJ, Sullivan DR, Burnett JR. Non-HDL-cholesterol and apolipoprotein B compared with LDL-cholesterol in atherosclerotic cardiovascular disease risk assessment. *Pathology*. 2019 Feb;51(2):148–54.
66. Sniderman AD, Islam S, McQueen M, Pencina M, Furberg CD, Thanassoulis G, et al. Age and Cardiovascular Risk Attributable to Apolipoprotein B, Low-Density Lipoprotein Cholesterol or Non-High-Density Lipoprotein Cholesterol. *J Am Heart Assoc*. 2016 Oct;5(10).
- 905 67. Holmes MV, Lange LA, Palmer T, Lanktree MB, North KE, Almoguera B, et al. Causal Effects of Body Mass Index on Cardiometabolic Traits and Events: A Mendelian Randomization Analysis. *Am J Hum Genet* [Internet]. 2014;94(2):198–208. Available from: <http://www.sciencedirect.com/science/article/pii/S0002929713005843>
- 910 68. Arsenault BJ, Després J-P, Stroes ESG, Wareham NJ, Kastelein JJP, Khaw K-T, et al. Lipid assessment, metabolic syndrome and coronary heart disease risk. *Eur J Clin Invest*. 2010 Dec;40(12):1081–93.
69. Felig P, Marliss E, Cahill GF. Plasma Amino Acid Levels and Insulin Secretion in Obesity. *N Engl J Med* [Internet]. 1969 Oct 9;281(15):811–6. Available from: <http://www.nejm.org/doi/abs/10.1056/NEJM196910092811503>
- 915 70. Adibi SA. Influence of dietary deprivations on plasma concentration of free amino acids of man. *J Appl Physiol*. 1968 Jul;25(1):52–7.
71. Luck JM, Morse SW. The effects of insulin and adrenaline on the amino-acid content of blood. *Biochem J*. 1933;27(5):1648–54.
- 920 72. Floyd JCJ, Fajans SS, Conn JW, Knopf RF, Rull J. Stimulation of insulin secretion by amino acids. *J Clin Invest*. 1966 Sep;45(9):1487–502.
73. IVY JH, SVEC M, FREEMAN S. Free plasma levels and urinary excretion of eighteen amino acids in normal and diabetic dogs. *Am J Physiol*. 1951 Oct;167(1):182–92.
74. Carlsten A, Hallgren B, Jagenburg R, Svanborg A, Werkö L. Amino acids and free fatty acids in plasma in diabetes. I. The effect of insulin on the arterial levels. *Acta Med Scand*. 1966 Mar;179(3):361–70.
- 925 75. White PJ, Newgard CB. Branched-chain amino acids in disease. *Science (80- )* [Internet]. 2019 Feb 8;363(6427):582–3. Available from: <https://www.science.org/doi/10.1126/science.aav0558>
- 930 76. Neeland IJ, Boone SC, Mook-Kanamori DO, Ayers C, Smit RAJ, Tzoulaki I, et al. Metabolomics Profiling of Visceral Adipose Tissue: Results From MESA and the NEO Study. *J Am Heart Assoc*. 2019;8(9).
77. Geidenstam N, Al-Majdoub M, Ekman M, Spégel P, Ridderstråle M. Metabolite profiling of obese individuals before and after a one year weight loss program. *Int J Obes*. 2017;41(9):1369–78.
- 935 78. Geidenstam N, Danielsson APH, Spégel P, Ridderstråle M. Changes in glucose-elicited blood

- metabolite responses following weight loss and long term weight maintenance in obese individuals with impaired glucose tolerance. *Diabetes Res Clin Pract.* 2016;113:187–97.
- 940 79. Weng L, Quinlivan E, Gong Y, Beitelshes AL, Shahin MH, Turner ST, et al. Association of Branched and Aromatic Amino Acids Levels with Metabolic Syndrome and Impaired Fasting Glucose in Hypertensive Patients. *Metab Syndr Relat Disord* [Internet]. 2015 Jun;13(5):195–202. Available from: <http://www.liebertpub.com/doi/10.1089/met.2014.0132>
80. Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, et al. Metabolite profiles and the risk of developing diabetes. *Nat Med.* 2011;17(4):448–53.
- 945 81. Porcu E, Gilardi F, Darrous L, Yengo L, Bararpour N, Gasser M, et al. Triangulating evidence from longitudinal and Mendelian randomization studies of metabolomic biomarkers for type 2 diabetes. *Sci Rep* [Internet]. 2021;11(1):1–10. Available from: <https://doi.org/10.1038/s41598-021-85684-7>
- 950 82. Lotta LA, Scott RA, Sharp SJ, Burgess S, Luan J, Tillin T, et al. Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. *PLoS Med* [Internet]. 2016 Nov 1 [cited 2020 Oct 14];13(11). Available from: <https://pubmed.ncbi.nlm.nih.gov/27898682/>
- 955 83. Wang Q, Holmes M V, Davey Smith G, Ala-Korpela M. Genetic Support for a Causal Role of Insulin Resistance on Circulating Branched-Chain Amino Acids and Inflammation. *Diabetes Care* [Internet]. 2017 Dec;40(12):1779–86. Available from: <http://care.diabetesjournals.org/lookup/doi/10.2337/dc17-1642>
84. Fontana L, Cummings NE, Arriola Apelo SI, Neuman JC, Kasza I, Schmidt BA, et al. Decreased Consumption of Branched-Chain Amino Acids Improves Metabolic Health. *Cell Rep* [Internet]. 2016;16(2):520–30. Available from: <http://dx.doi.org/10.1016/j.celrep.2016.05.092>
- 960 85. Dekkers IA, de Mutsert R, de Vries APJ, Rosendaal FR, Cannegieter SC, Jukema JW, et al. Determinants of impaired renal and vascular function are associated with elevated levels of procoagulant factors in the general population. *J Thromb Haemost.* 2018 Mar;16(3):519–28.

## Tables

**Table 1. Population summary statistics**

Trait	NEO	Men	Women	Leiderdorp	Leiden	var exp by sub-pop	p-value
sex	M:2671 F:2846	M:2671 F:0	M:0 F:2846	M:623 F:783	M:2048 F:2063	0.17%	3.65E-04
age	55.99 (45.00-65.00)	56.13 (45.00-65.00)	55.86 (45.00-65.00)	56.13 (46.00-65.00)	55.95 (45.00-65.00)	0.02%	3.26E-01
height	1.74 (1.57-1.92)	1.81 (1.68-1.95)	1.67 (1.55-1.79)	1.73 (1.57-1.92)	1.74 (1.57-1.92)	0.03%	2.03E-01
weight	90.58 (60.00-126.80)	97.45 (72.95-131.60)	84.14 (57.20-121.52)	79.16 (54.00-116.15)	94.49 (69.60-128.85)	16.12%	9.63E-213
waist circ.	102.01 (74.00-129.00)	106.32 (85.79-131.00)	97.96 (71.00-126.00)	91.22 (70.00-119.00)	105.70 (86.00-131.00)	22.87%	2.85e-313
hip circ.	110.23 (94.00-133.00)	108.57 (96.00-125.00)	111.79 (92.00-138.00)	103.40 (90.00-124.17)	112.57 (99.00-135.00)	16.22%	3.33E-214
pack years	11.21 (0.00-52.38)	13.71 (0.00-61.69)	8.88 (0.00-43.83)	8.69 (0.00-43.32)	12.09 (0.00-55.20)	0.83%	1.01E-10
alcohol	15.52 (0.00-60.52)	21.40 (0.00-73.24)	10.00 (0.00-41.72)	14.58 (0.00-54.06)	15.84 (0.00-62.57)	0.10%	1.95E-02
smokin g (N F C) higher ed %	34.11   49.85   16.03	30.32   50.49   19.19	37.67   49.26   13.07	39.29   46.76   13.95	32.34   50.91   16.74	0.38%	4.11E-06
BMI	29.98 (21.40-41.35)	29.74 (22.72-38.99)	30.20 (20.64-42.90)	26.24 (19.94-37.20)	31.26 (26.02-42.19)	20.70%	3.66E-280
BMI-PGS	10.20 (9.64-10.78)	10.20 (9.63-10.78)	10.20 (9.65-10.77)	10.14 (9.56-10.71)	10.22 (9.67-10.80)	1.63%	1.53E-21
R <sup>2</sup>	0.0431, 0.0473**	0.0439	0.0453	0.0427	0.0243	-	-
beta (se)	3.38 (0.21), 3.26 (0.20)**	2.77 (0.25)	4.00 (0.34)	3.14 (0.40)	2.25 (0.22)	-	-
P	8.10x10 <sup>-55</sup> 3.61x10 <sup>-60**</sup>	7.23x10 <sup>-28</sup>	1.57x10 <sup>-30</sup>	5.10x10 <sup>-15</sup>	8.22x10 <sup>-24</sup>	-	-

40

**Table 1:** Population summary statistics for the (1) complete NEO cohort, for (2) men and (3) women of the NEO cohort and the two NEO sub-populations (4) Leiderdorp and (5) Leiden, respectively. For each variable: age, height, weight, waist circumference, hip circumference, waist-hip ratio, smoking in pack years, alcohol consumption in grams per day, the percent of the population with higher education, BMI, and BMI-PGS the mean (95% confidence intervals) is provided. In addition, the number of men and women in each sample (sub-)population is provided (sex). The last three rows of data provide summary statistics describing the relationship between the instrumental variable (BMI-PGS) on the exposure (BMI). Specifically, the variance explained ( $R^2$ ), the effect estimate (beta) and standard error (se), and the P-value (P) are provided. Estimates suffixed with \*\* are derived from linear models that include sample weights. The last two columns of data provide an estimate of the proportion of variation (in NEO) of the trait (in the row) explained by variation between the sub-population samples, as estimated by a linear regression, and the model P-value.

50

**Table 2 Metabolite traits causally associated with BMI**

55

trait	metabolite	class	dietary state	beta	se	P
<b>lafa_f</b>	Ratio of 18:2 linoleic acid to total fatty acids (%)	Fatty acids ratios	fasting	-0.106	0.025	1.80E-05
<b>faw6fa_f</b>	Ratio of omega-6 fatty acids to total fatty acids (%)	Fatty acids ratios	fasting	-0.081	0.024	6.97E-04
<b>pufafa_f</b>	Ratio of polyunsaturated fatty acids to total fatty acids (%)	Fatty acids ratios	fasting	-0.081	0.024	6.80E-04
<b>cit_f</b>	Citrate (mmol/l)	Glycolysis related metabolites	fasting	-0.079	0.024	8.63E-04
<b>leu_f</b>	Leucine (mmol/l)	Amino acids	fasting	0.068	0.020	5.62E-04
<b>lhdlc_p</b>	Total cholesterol in large HDL (mmol/l)	Lipoprotein subclasses	postprandial	-0.076	0.023	1.12E-03
<b>lhdlfc_p</b>	Free cholesterol in large HDL (mmol/l)	Lipoprotein subclasses	postprandial	-0.077	0.023	9.73E-04
<b>mvdldlpct_p</b>	Phospholipids in medium VLDL to total lipids in medium VLDL ratio (%)	Lipoprotein subclasses ratios	postprandial	-0.073	0.022	9.98E-04
<b>idlfcpt_p</b>	Free cholesterol in IDL to total lipids in IDL ratio (%)	Lipoprotein subclasses ratios	postprandial	-0.079	0.023	5.08E-04
<b>lhdlcpt_p</b>	Total cholesterol in large HDL to total lipids in large HDL ratio (%)	Lipoprotein subclasses ratios	postprandial	-0.085	0.025	7.90E-04
<b>lhdlcept_p</b>	Cholesterol esters in large HDL to total lipids in large HDL ratio (%)	Lipoprotein subclasses ratios	postprandial	-0.086	0.026	1.03E-03
<b>lhdlfcpt_p</b>	Free cholesterol in large HDL to total lipids in large HDL ratio (%)	Lipoprotein subclasses ratios	postprandial	-0.095	0.025	1.20E-04
<b>lafa_p</b>	Ratio of 18:2 linoleic acid to total fatty acids (%)	Fatty acids ratios	postprandial	-0.092	0.025	2.00E-04
<b>faw6fa_p</b>	Ratio of omega-6 fatty acids to total fatty acids (%)	Fatty acids ratios	postprandial	-0.080	0.024	7.83E-04
<b>pufafa_p</b>	Ratio of polyunsaturated fatty acids to total fatty acids (%)	Fatty acids ratios	postprandial	-0.077	0.024	1.10E-03
<b>cit_p</b>	Citrate (mmol/l)	Glycolysis related metabolites	postprandial	-0.105	0.022	2.83E-06



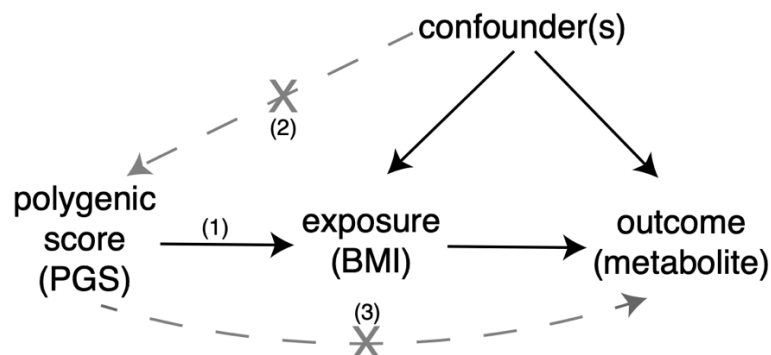
<b>his_p</b>	Histidine (mmol/l)	Amino acids	postprandial	-0.085	0.024	3.96E-04
<b>mvlldlplpct_r</b>	Phospholipids in medium VLDL to total lipids in medium VLDL ratio (%)	Lipoprotein subclasses ratios	response	-0.091	0.027	6.75E-04
<b>svldlfcpc_r</b>	Free cholesterol in small VLDL to total lipids in small VLDL ratio (%)	Lipoprotein subclasses ratios	response	-0.082	0.025	1.11E-03
<b>ala_r</b>	Alanine (mmol/l)	Amino acids	response	-0.081	0.023	5.91E-04

**Table 2:** MR association summary statistics for the 20 metabolite traits associated with BMI. Provided are the trait name, metabolite description and units, metabolite class as defined by Nightingale Heath, the dietary state, and the effect estimate, standard error, and P-value. The effect estimates are in normalized standard deviation units of change per 1 unit change of BMI (kg/m<sup>2</sup>).

## Figures

**Figure 1 Mendelian Randomization schematic**

65

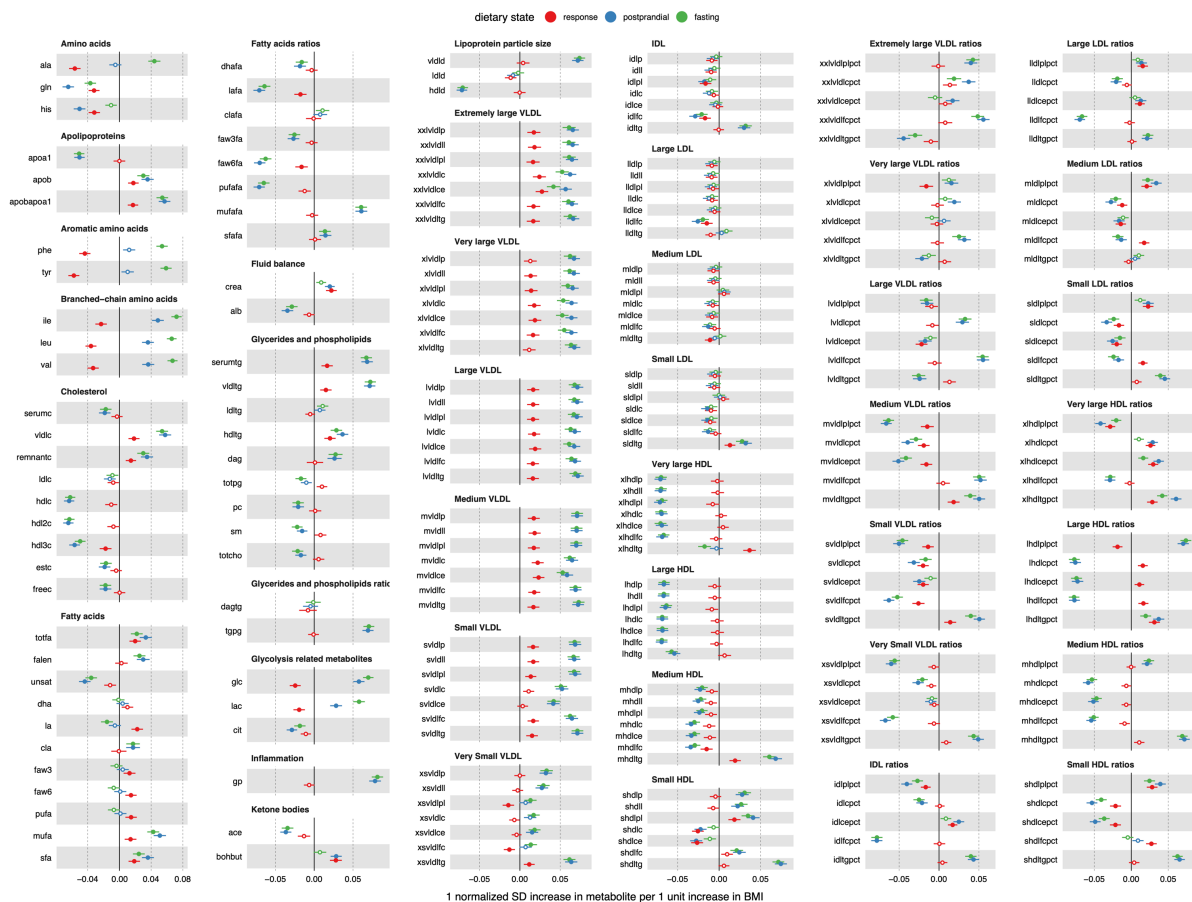


70

**Figure 1:** A schematic or directed acyclic graph of the Mendelian randomization framework. The MR framework in this study assumes (1) that the instrument is robustly associated with the outcome, (2) that the instrument is not associated with variables that are associated to the outcome and (3) that the instrument acts on the outcome only through the exposure and not through a pleiotropic pathway.

75

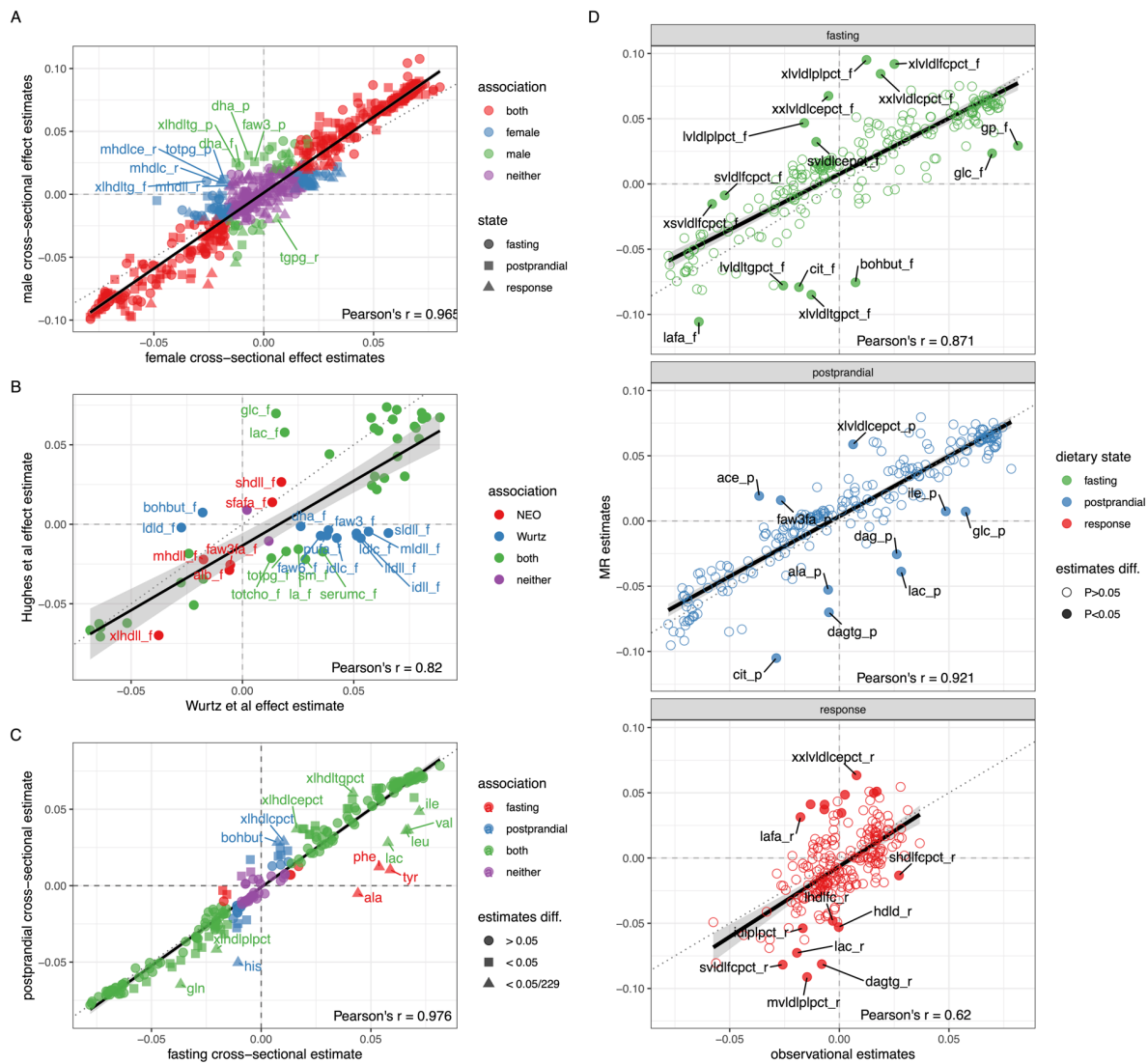
**Figure 2: Observational effect estimate forest plot**



**Figure 2: Observational BMI – metabolite abundance and response effect estimates.** Point estimates and 95% confidence intervals for BMI on fasting metabolite abundance (green), postprandial metabolite abundance (blue), and metabolite response (red). Metabolites are classified and organized by Nightingale Health sub-class assignments. Metabolites that are associated with BMI, after multiple test correction ( $P < 0.05/43$ ), are indicated as solid color point estimates.

80

85 **Figure 3: Scatter plots**



**Figure 3: Scatter plots.** In all plots the solid black line is a best fit regression line, the dotted black line is an equivalency line (intercept = 0, slope = 1), the dashed vertical and horizontal line indicate the zero values, and a Pearson's correlation coefficient is in the bottom right corner. (A) Scatter plot of

90 BMI-metabolite cross-sectional effect estimates for females (x-axis) and males (y-axis). Point estimates are colored red if there is an association between the metabolite and BMI in both sexes, blue if an association was only observed in females, green if an association was only observed in males, and purple if no association was observed in either sex. Point estimates are shaped as circles, squares, and triangles to represent the three dietary states fasting, postprandial, and response, respectively. (B) Scatter plot of BMI-metabolite cross-sectional effect estimates from Wurtz et al (x-axis) and NEO (this study, y-axis). Point estimates are colored red if there is an association between the metabolite and BMI in NEO, blue if an association was only observed in Wurtz et al, green if an association was observed in both studies, and purple if no association was observed in either study. (C) Scatter plot of BMI-metabolite cross-sectional effect estimates for fasting (x-axis) and postprandial (y-axis) dietary states. Point estimates are colored red if there is an association between

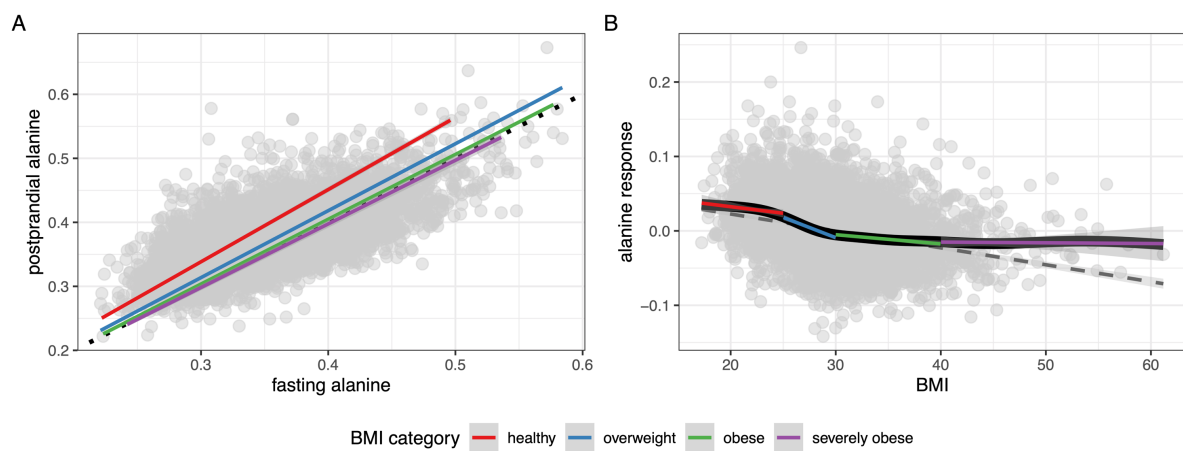
95

100

the metabolite and BMI in the fasting state, blue if an association was only observed in the postprandial state, green an association was observed in both dietary states, and purple if no association was observed in either. Point estimates are shaped as circles if they are not different between the two dietary states – as determined by a z-test – shaped as squares if they are nominally different ( $P < 0.05$ ) and triangles and labeled if they are different after correcting for multiple tests ( $P < 0.05/229$ ). (D) Scatter plot of BMI-metabolite cross-sectional (x-axis) and MR (y-axis) effect in the fasting (top and green), postprandial (middle and blue), and response (bottom and red) dietary states. Metabolites with effect estimates that differ between the cross-sectional and MR analyses are solid circles and labelled with the metabolite name.

110

### Figure 4 Alanine response

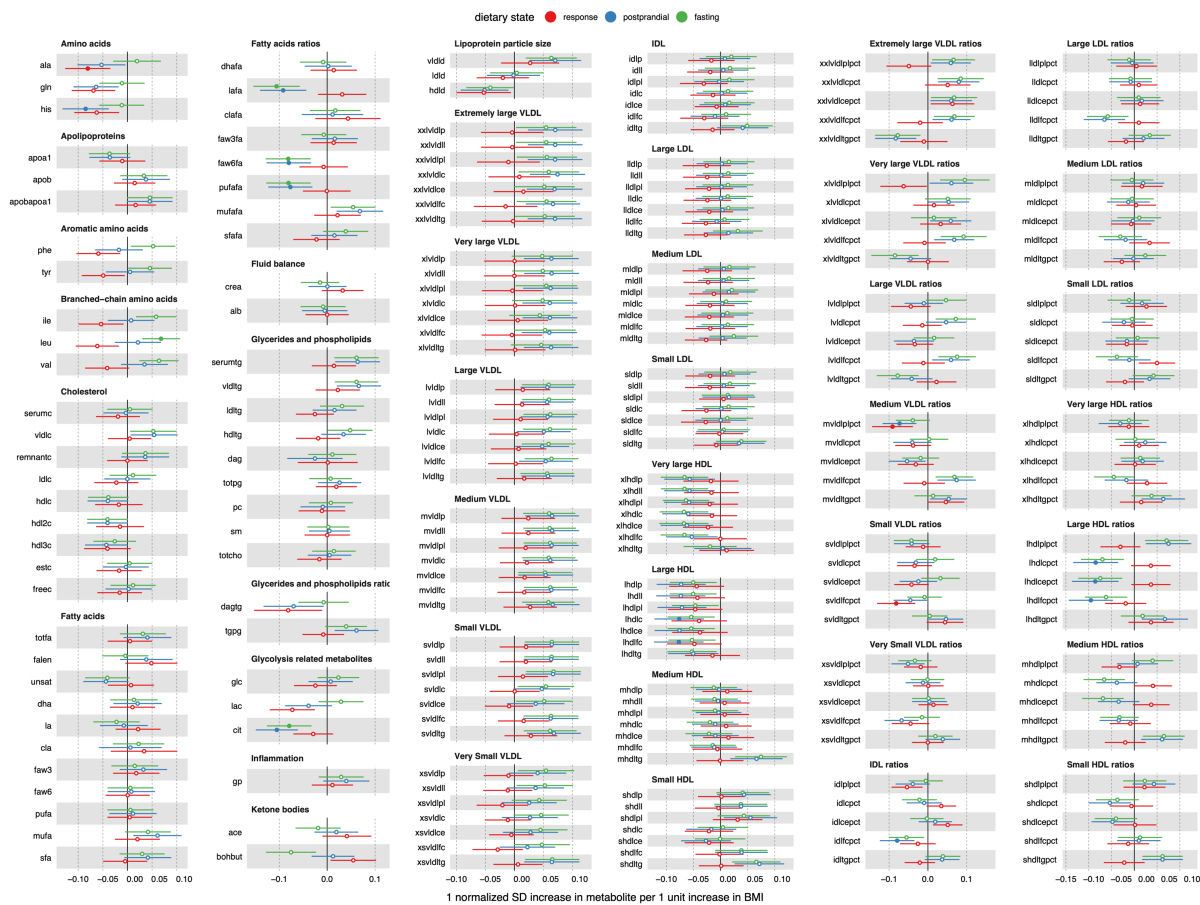


**Figure 4: Association of BMI with alanine response.** Plot on the left illustrates the correlation between fasting (x-axis) and postprandial (y-axis) tyrosine abundance, with a best fit line for 4 BMI classes of individuals: healthy ( $BMI < 25$ ), overweight ( $25 \leq BMI < 30$ ), obese ( $30 \leq BMI < 40$ ), and severely obese ( $BMI > 40$ ). All best fit lines were forced through an intercept of zero. Note that the slope of the best fit line is largest for individuals who have a healthy BMI and smallest for those who have severe obesity. The plot on the right illustrates the inverse relationship between BMI (x-axis) and tyrosine response (y-axis). The blue line is a best fit linear regression line, the black line is a best fit generalized additive model (GAM) smooth (non-linear regression), and the red, blue, green, and purple line intervals are best fit regression for each clinically defined portion of the BMI distribution, which mirror the non-linear regression.

115

120

125 **Figure 5: Forest plot of MR effect estimates**

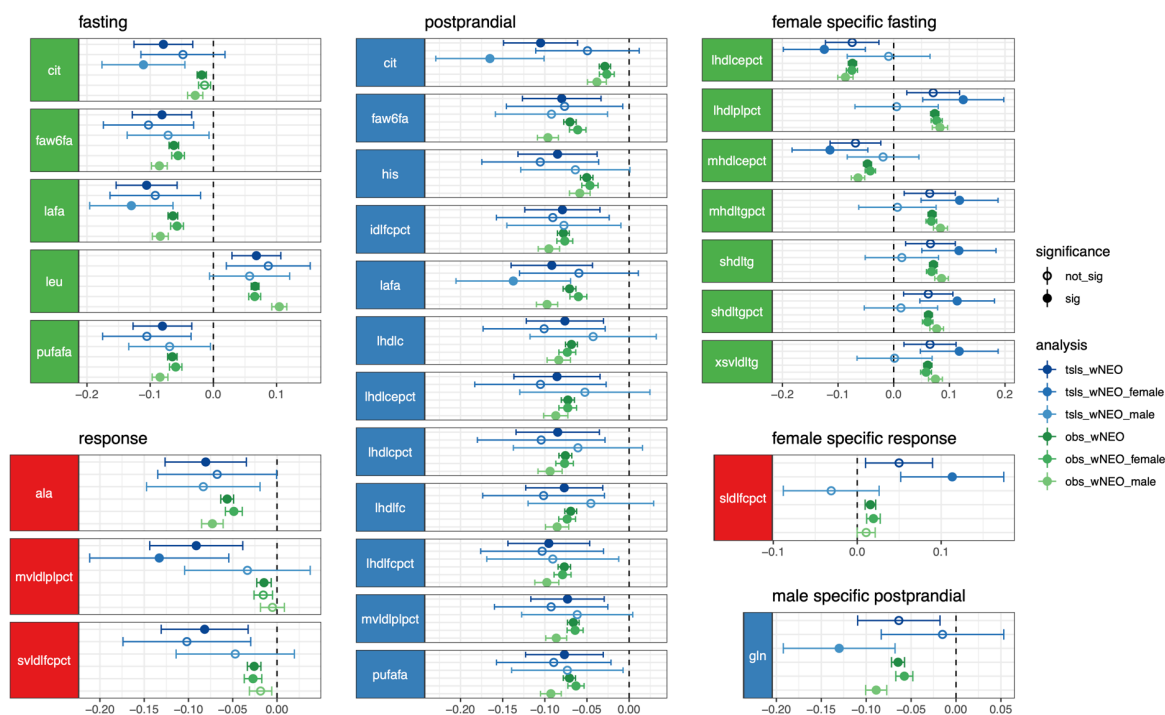


**Figure 5: Forest Plot of MR effect estimates.** Point estimates and 95% confidence intervals for BMI on fasting metabolite abundance (green), postprandial metabolite abundance (blue), and metabolite response (red). Metabolites are classified and organized by Nightingale Health sub-class assignments. Metabolites that are associated with BMI, after multiple test correction ( $P < 0.05/43$ ), are indicated as solid color point estimates.

130

135

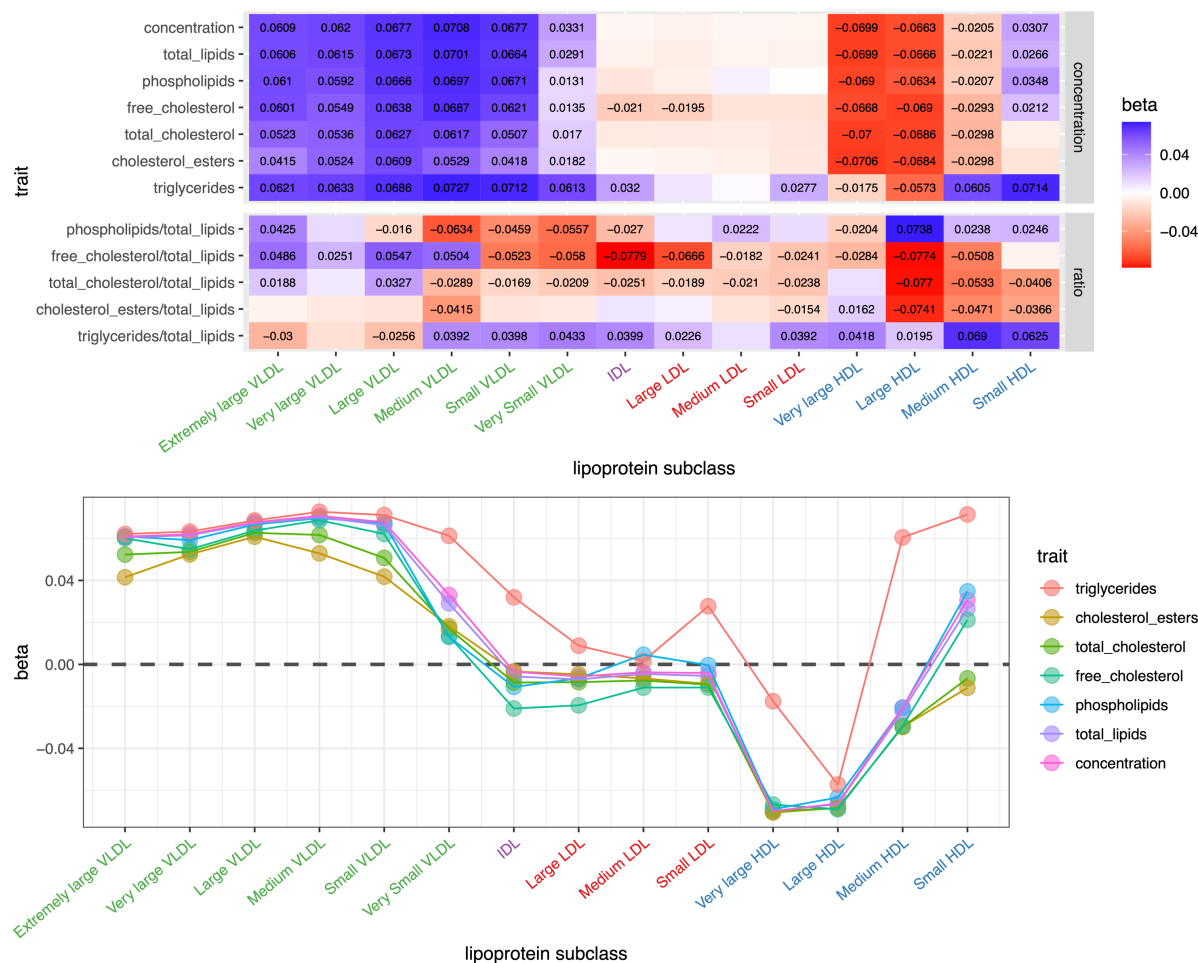
**Figure 6: Metabolites associated with BMI in MR analyses**



**Figure 6: Metabolites associated with BMI in MR analyses.** A forest plot of effect estimates (points) and 95% confidence intervals (whiskers) for MR (blue, “tsls” prefix) and observational (green, “obs” prefix) effect estimates for the weighted NEO (wNEO), weighted NEO female (wNEO\_female), and weighted NEO male (wNEO\_male) (sub-)populations. Fasting, postprandial, and response association observed in the general population as well as those association only seen in females and males.

140

**Figure 7: Fasting lipoprotein cross-sectional profile**



145

**Figure 7: Fasting lipoprotein cross-sectional profile.** (Upper) A tile plot of cross-sectional effect estimates for lipoproteins and lipoprotein ratios in the fasting state. Tiles with an effect estimate provided in text are those associated with BMI ( $p < 1.16 \times 10^{-3}$ ). The lipoproteins are organized by size or density along the x-axis, and the component or ratio being measured is along the y-axis.

150

(Lower) A dot plot or profile of cross-sectional effect estimates for lipoproteins (x-axis) ordered by lipoprotein size or density with effect estimates (y-axis). The component or measurement of each lipoprotein are defined by the color as described in the key.