Title: From menarche to menopause: the impact of reproductive factors on the metabolic profile of over 65,000 women

# Rationale

Markers of women’s reproductive health, such as age at menarche, parity and age at menopause, have been associated with several common chronic conditions, cardiovascular and cardiometabolic diseases (1). Some attempts have been made to explore the extent to which these associations are causal, as opposed to explained by residual confounding, using approaches such as Mendelian randomization (MR) and negative control designs, which are less prone to bias by key confounders from conventional observational studies.

Metabolites could act as mediators of the relationship of reproductive markers, and related hormonal changes, with chronic diseases (5-7). Determining the effect of women’s reproductive markers on multiple metabolites would be the first step to exploring this and could provide crucial insights into mechanisms underlying women’s long term health.

# Aim

The aim of this paper is to explore the extent to which women’s reproductive markers (age at menarche, parity/number of children, and age at natural menopause) have a causal effect on 249 metabolic measures (covering lipids, fatty acids, amino acids, glycolysis, ketone bodies and an inflammatory marker). We do this by triangulating evidence (8) across multivariable regression, a negative control design (for parity only), and MR. Given each of these approaches has unique strengths and limitations, results that agree across them are less likely to be spurious (8).

# Sample

UK biobank including both men and women.

# Exposures

* Age at menarche
* Parity (in women) and number of children in men as a paternal negative control
* Age at menopause

# Outcomes

Metabolic traits were measured using a targeted high-throughput NMR metabolomics (Nightingale Health Ltd; biomarker quantification version 2020). This platform provides simultaneous quantification of 249 metabolic measures, consisting of concentrations of 165 metabolic measures and 84 derived ratios, encompassing routine lipids, lipoprotein subclass profiling (including lipid composition within 14 subclasses), fatty acid composition, and various low-molecular weight metabolites such as amino acids, ketone bodies and glycolysis metabolites. Technical details and epidemiological applications have been previously reviewed (**Table 1**)

# Quality control

Pre-release data from a random subset of 126,846 non-fasting plasma samples collected at baseline or first repeat assessment were made available to early access analysts. 121,577 samples were retained for analyses after removing duplicates and observations not passing quality control (QC) (i.e. sample QC flag “Low protein”, biomarker QC flag “Technical error”, or samples with insufficient material)

All metabolites will be normalised using the rank-based inverse normal transformation (INT) to account for their non-normal distributions. In an INT transformation the distribution of the variable is ordered from smallest to largest such as 1, 2, 5, 10 in which the same distance is assumed between each value.

# Interpretation of association/effect sizes

Difference in mean [metabolite] SD per 1 [unit of reproductive trait] [reproductive trait]. For example, difference in mean glycine SD per 1 year older age at menarche: xx (95%CI: xx, xx)

# Confounders

By definition, a confounder should be known to, or plausibly, influence both the exposure (reproductive traits) and outcome (metabolites). Directed acyclic graphs (DAGs), informed by knowledge from the literature are used to decide what to include as potential confounders. Where possible, we will use measurements of confounders obtained before the exposure and outcome assessments. However, given we are using UKB, data on confounders are likely obtained from the baseline assessment which occurs after, say age at menarche. This will be added as a discussion point in the paper.

For multivariable analyses, confounders were defined a priori based on them being known or plausible causal factors for reproductive traits and cardiovascular risk via higher/lower metabolites. A minimal set of adjustments will be made in the main multivariable regression analyses as most confounders were not assessed prior to or around when the reproductive traits occurred. Specifically, we will adjust for education as a categorical variable (University, A-levels, O levels (or equivalent) or other), age at baseline and retrospectively reported body size at age 10 (average, thinner, plumper) in all regression analyses. In additional analyses we also partially adjusted for the full set of defined confounders using baseline measurements (mostly after the occurrence of exposures) as correlates of the before exposure measures (see below in statistical analyses). Visual representation of potential causal relationships between parity (exposure) and metabolites (outcome), whilst accounting for potential:

Age at baseline

Reproductive traits (x)

Metabolites (y)

Confounders (prior to exposure)

1. Education
2. BMI
3. Those with a lower socio-economic position (SEP), proxied by education, may be associated with higher parity, earlier age at menarche and menopause. Lower SEP also associated with poorer heath which could be linked to worse metabolite outcomes.
4. BMI (highly correlated with lower SEP influences fertility and so may be associated with higher parity through link with SEP or lower parity through impact on fertility and higher BMI results in worse metabolic outcomes. Pre-puberty BMI is likely to influence age at menopause

Note: We do control for any covariate, such as age, that’s related to the outcome (i.e. worse metabolite outcomes), to reduce the variance of our measure as much as possible.

# Other variables of interest

Statin use

# Statistical analyses

# Multivariable regression of associations between women’s reproductive traits and future NMR metabolites in women in UKB

To examine the research question, multivariable linear regression will be conducted to assess if women’s reproductive health are associated with future NMR metabolites in women in UKB. The regression equation will be based on the DAG.

To facilitate comparisons between methods such as MR, we will assume age to have a linear relationship with each reproductive trait.

The assumptions of multiple regression—linearity, homoscedasticity (of residuals) and multicollinearity—will be assessed. Linearity assumes a straight-line relationship between (one or more) predictor variables (and the linear predictor) and the outcome variable and assessed with a scatterplot. Homoscedasticity or equal variance assumes that residuals are evenly distributed about the regression line. and assessed with a scatterplot between the residuals and the fitted values of the outcome. We also check equal variance occurs for each covariate. The normality of the residuals will also be checked with a histogram and qq plot.

## Multiple testing

As there are 249 metabolites and therefore 249 multiple regression analyses, we cannot assume these are all independent tests. However, within the 249 outcomes some are highly correlated, so we correct by the number of independent principal component outcomes (22) rather than all 249 outcomes.

## Missing data

Missing data will be descriptively described only and discussed in the limitations of the paper.

## Other things to think about

In UKB not everyone will have experienced a natural menopause (due to age or other factors such as surgical menopause or missing data) and some women will only have recently experienced the menopause i.e less than two years before study recruitment (when blood samples for NMR metabolomics were collected); therefore, multivariable regression estimates could be influenced by bias due to sample selection or short follow-up between menopause and assessment of metabolic measures.

# Paternal negative control for parity only

Negative control analyses aim to emulate a condition that cannot involve the hypothesized causal mechanism but is likely to have similar sources of bias that may have been present in the association of interest. We use males as negative controls to assess potential biases in the association between parity (proxied by number of live births) and metabolic measures in women. If associations between number of live births and metabolic measures in women reflect a causal effect of parity on women’s metabolic health, one would expect number of live births to be associated with metabolic measures in women but not in men given men do not experience pregnancy. Similar to the multivariable regression analyses, we test the association between number of children (men) and their measured metabolites and present three sets of models: (1) with no adjustments, (2) adjusted for education, age at baseline and retrospectively self-reported body composition at age 10 and (3) model (2) additionally adjusted for baseline variables collected at the first assessment at (mean) age 56 years (SD=8) including BMI, smoking and alcohol status.

# Two sample MR

We use two-sample MR to explore the effect of older age at menarche, higher parity, and older age at natural menopause on women’s metabolic profile. Publicly available GWAS summary data will be used for SNP-reproductive traits associations (sample 1) and UK Biobank summary GWAS data for SNP-metabolite associations (sample 2). This approach does not require all participants to have data on both exposure and outcome, and, therefore, allows us to retain the largest possible sample sizes, meaning that power to detect a causal effect is increased (10).

*Selection of genetic instruments*

Age at menarche

Genetic instruments will be selected from a GWAS of age at menarche, which includes 329,345 women of European ancestry (11). Linear regression models will be used to estimate the association between genetic variants and age at menarche (in years) adjusting for age at study visit and study-specific covariables. For our analyses, we select the 389 independent SNPs reported by the GWAS to be strongly associated with age at menarche (P-value < 5\*10-8) in the discovery metanalyses. Given the age at menarche GWAS included UK Biobank participants (maximum estimated sample overlap: ~20%), we also select an additional set of age at menarche-associated genetic variants (N = 68 SNPs) using data from a previous GWAS that did not including UK Biobank (details in ‘Sensitivity analyses’ below (12)).

Parity

Genetic instruments will be selected from a GWAS of number of children ever born, as a proxy of parity, which includes 785,604 men and women of European ancestry from 45 studies (13). Number of children ever born was treated as a continuous measure and included both parous and nulliparous women. Linear regression models were used to estimate the association between genetic variants and number of children ever born adjusting for principal components of ancestry, birth year, its square and cubic, to control for non-linear birth cohort effects. Family-based studies controlled for family structure or excluded relatives. The sex-combined metanalysis also included interactions of birth year and its polynomials with sex. For our analyses, we will use the 32 independent SNPs reported by the GWAS to be strongly associated with number of children ever born (P-value < 5\*10-8) in either the sex-combined (28 SNPs) or female-specific (4 SNPs) metanalyses and summary association data from the female-specific metanalyses. The GWAS included UK Biobank (maximum estimated sample overlap: 14%).

Age at natural menopause

Genetic instruments will be selected from a GWAS of age at natural menopause conducted in 201,323 women of European ancestry (14). Linear regression models were used to estimate the association between genetic variants and age at natural menopause (in years). For our analyses, we select 290 SNPs reported by the GWAS to be strongly associated with age at natural menopause (P-value < 5\*10-8). Where available, we will use association data from the sample combining discovery and replication stages (N = 496,151). Given the age at menarche GWAS included UK Biobank participants (maximum estimated sample overlap: 13% considering the GWAS combined discovery and replication samples), we will also select an additional set of age at natural menopause-associated genetic variants (N = 42 SNPs) using data from a previous GWAS that did not include UK Biobank (details in ‘Sensitivity analyses’ below(15).

*Main analyses*

We will use a standard two-sample MR method, the inverse variance weighted (IVW) estimator, to explore the effect of age at menarche, parity and age at natural menopause on women’s metabolic profile by combining genetic association estimates for reproductive traits (extracted from published GWASes data) with genetic association estimates for the metabolic measures (generated from UK Biobank data). Given a priori evidence of a potential bidirectional relationship between age at menarche and BMI, we will also use multivariable IVW to test the effect of age at menarche on metabolic measures accounting for adult BMI. For multivariable IVW analysis, apart from the data previously described, we will use summary genetic association data for BMI extracted from the 2015 metanalysis by the GIANT consortium (N = 339,224 individuals not including UK Biobank participants) (16).

*Sensitivity analyses*

Several sensitivity analyses will be conducted to explore the plausibility of the three core MR assumptions, which are required for the method to provide a valid test of the presence of a causal effect.

Assumption 1: the genetic instrument must be associated with the reproductive trait

We select the genetic variants reported to be strongly associated with reproductive in the largest available GWAS and estimate the proportion of phenotypic variance explained (R2) and F-statistics for the association of SNPs with reproductive traits among females as an indicator of instrument strength.

Assumption 2: the association between genetic instrument and outcome is unconfounded

One of the main motivations for using MR is to avoid unmeasured confounding. However, there is growing evidence that, in some instances, MR studies can be confounded when using data from unrelated individuals due to population stratification, assortative mating and indirect genetic effects of parents (17, 18). We use two approaches to explore whether these will be likely to bias our main results. First, we will use sex-combined data from a recent within-sibship GWAS, including up to 159,701 siblings from 17 cohorts, to test the effect of genetic susceptibility to higher age at menarche, parity and age at menopause on metabolic markers (i.e. LDL-cholesterol, triglycerides, and glycated haemoglobin) (17). Within-sibling MR designs control for variation in parental genotypes, and so should not be affected by population stratification, assortative mating and indirect genetic effects of parents (17-19). Second, we will perform IVW on negative control outcomes (i.e. skin colour and skin tanning ability) since these could not conceivably be affected by the exposures and any evidence for an association between reproductive traits and, these negative control outcomes would be indicative of residual population stratification in the exposure GWAS (20).

Assumption 3: the genetic instrument does not affect the outcome except through its possible effect on the exposure

A key violation of this assumption is known as horizontal pleiotropy, where genetic variants influence the outcome through pathways that are not mediated by the exposure (21). We explore the presence of bias due to horizontal pleiotropy by using other MR methods: the weighted median estimator and MR-Egger. These methods can provide valid tests of a causal effect under different (and weaker) assumptions about the nature of the underlying horizontal pleiotropy. The weighted median estimator requires that at least 50% of the weight in the analysis stems from valid instruments. The MR-Egger estimator assumes that the instrument strength is independent of its the direct effects on the outcome (i.e. INSIDE assumption).

In addition to the core assumptions, the two-sample MR approach assumes that genetic associations with exposure and outcome will be estimated from two comparable but non-overlapping samples. We restrict our analyses to European adult individuals to ensure that samples will be comparable. We assessed potential bias due to sample overlap by conducting MR using SNPs selected from previous GWAS of age at menarche and age at natural menopause that do not include UK Biobank

# Table 1 List of outcomes

|  |  |
| --- | --- |
|  |  |
| **Abbreviation** | **Metabolite name** |
| Ala | Alanine |
| ApoA1 | Apolipoprotein A1 |
| ApoB | Apolipoprotein B |
| ApoB\_by\_ApoA1 | Ratio of apolipoprotein B to apolipoprotein A1 |
| Cholines | Total cholines |
| Clinical\_LDL\_C | Clinical LDL cholesterol |
| DHA | Docosahexaenoic acid |
| DHA\_pct | Ratio of docosahexaenoic acid to total fatty acids |
| Gln | Glutamine |
| Gly | Glycine |
| GlycA | Glycoprotein acetyls |
| HDL\_C | HDL cholesterol |
| HDL\_CE | Cholesteryl esters in HDL |
| HDL\_FC | Free cholesterol in HDL |
| HDL\_L | Total lipids in HDL |
| HDL\_P | Concentration of HDL particles |
| HDL\_PL | Phospholipids in HDL |
| HDL\_TG | Triglycerides in HDL |
| HDL\_size | Average diameter for HDL particles |
| His | Histidine |
| IDL\_C | Cholesterol in IDL |
| IDL\_CE | Cholesteryl esters in IDL |
| IDL\_CE\_pct | Cholesteryl esters to total lipids ratio in IDL |
| IDL\_C\_pct | Cholesterol to total lipids ratio in IDL |
| IDL\_FC | Free cholesterol in IDL |
| IDL\_FC\_pct | Free cholesterol to total lipids ratio in IDL |
| IDL\_L | Total lipids in IDL |
| IDL\_P | Concentration of IDL particles |
| IDL\_PL | Phospholipids in IDL |
| IDL\_PL\_pct | Phospholipids to total lipids ratio in IDL |
| IDL\_TG | Triglycerides in IDL |
| IDL\_TG\_pct | Triglycerides to total lipids ratio in IDL |
| Ile | Isoleucine |
| LA | Linoleic acid |
| LA\_pct | Ratio of linoleic acid to total fatty acids |
| LDL\_C | LDL cholesterol |
| LDL\_CE | Cholesteryl esters in LDL |
| LDL\_FC | Free cholesterol in LDL |
| LDL\_L | Total lipids in LDL |
| LDL\_P | Concentration of LDL particles |
| LDL\_PL | Phospholipids in LDL |
| LDL\_TG | Triglycerides in LDL |
| LDL\_size | Average diameter for LDL particles |
| L\_HDL\_C | Cholesterol in large HDL |
| L\_HDL\_CE | Cholesteryl esters in large HDL |
| L\_HDL\_CE\_pct | Cholesteryl esters to total lipids ratio in large HDL |
| L\_HDL\_C\_pct | Cholesterol to total lipids ratio in large HDL |
| L\_HDL\_FC | Free cholesterol in large HDL |
| L\_HDL\_FC\_pct | Free cholesterol to total lipids ratio in large HDL |
| L\_HDL\_L | Total lipids in large HDL |
| L\_HDL\_P | Concentration of large HDL particles |
| L\_HDL\_PL | Phospholipids in large HDL |
| L\_HDL\_PL\_pct | Phospholipids to total lipids ratio in large HDL |
| L\_HDL\_TG | Triglycerides in large HDL |
| L\_HDL\_TG\_pct | Triglycerides to total lipids ratio in large HDL |
| L\_LDL\_C | Cholesterol in large LDL |
| L\_LDL\_CE | Cholesteryl esters in large LDL |
| L\_LDL\_CE\_pct | Cholesteryl esters to total lipids ratio in large LDL |
| L\_LDL\_C\_pct | Cholesterol to total lipids ratio in large LDL |
| L\_LDL\_FC | Free cholesterol in large LDL |
| L\_LDL\_FC\_pct | Free cholesterol to total lipids ratio in large LDL |
| L\_LDL\_L | Total lipids in large LDL |
| L\_LDL\_P | Concentration of large LDL particles |
| L\_LDL\_PL | Phospholipids in large LDL |
| L\_LDL\_PL\_pct | Phospholipids to total lipids ratio in large LDL |
| L\_LDL\_TG | Triglycerides in large LDL |
| L\_LDL\_TG\_pct | Triglycerides to total lipids ratio in large LDL |
| L\_VLDL\_C | Cholesterol in large VLDL |
| L\_VLDL\_CE | Cholesteryl esters in large VLDL |
| L\_VLDL\_CE\_pct | Cholesteryl esters to total lipids ratio in large VLDL |
| L\_VLDL\_C\_pct | Cholesterol to total lipids ratio in large VLDL |
| L\_VLDL\_FC | Free cholesterol in large VLDL |
| L\_VLDL\_FC\_pct | Free cholesterol to total lipids ratio in large VLDL |
| L\_VLDL\_L | Total lipids in large VLDL |
| L\_VLDL\_P | Concentration of large VLDL particles |
| L\_VLDL\_PL | Phospholipids in large VLDL |
| L\_VLDL\_PL\_pct | Phospholipids to total lipids ratio in large VLDL |
| L\_VLDL\_TG | Triglycerides in large VLDL |
| L\_VLDL\_TG\_pct | Triglycerides to total lipids ratio in large VLDL |
| Leu | Leucine |
| MUFA | Monounsaturated fatty acids |
| MUFA\_pct | Ratio of monounsaturated fatty acids to total fatty acids |
| M\_HDL\_C | Cholesterol in medium HDL |
| M\_HDL\_CE | Cholesteryl esters in medium HDL |
| M\_HDL\_CE\_pct | Cholesteryl esters to total lipids ratio in medium HDL |
| M\_HDL\_C\_pct | Cholesterol to total lipids ratio in medium HDL |
| M\_HDL\_FC | Free cholesterol in medium HDL |
| M\_HDL\_FC\_pct | Free cholesterol to total lipids ratio in medium HDL |
| M\_HDL\_L | Total lipids in medium HDL |
| M\_HDL\_P | Concentration of medium HDL particles |
| M\_HDL\_PL | Phospholipids in medium HDL |
| M\_HDL\_PL\_pct | Phospholipids to total lipids ratio in medium HDL |
| M\_HDL\_TG | Triglycerides in medium HDL |
| M\_HDL\_TG\_pct | Triglycerides to total lipids ratio in medium HDL |
| M\_LDL\_C | Cholesterol in medium LDL |
| M\_LDL\_CE | Cholesteryl esters in medium LDL |
| M\_LDL\_CE\_pct | Cholesteryl esters to total lipids ratio in medium LDL |
| M\_LDL\_C\_pct | Cholesterol to total lipids ratio in medium LDL |
| M\_LDL\_FC | Free cholesterol in medium LDL |
| M\_LDL\_FC\_pct | Free cholesterol to total lipids ratio in medium LDL |
| M\_LDL\_L | Total lipids in medium LDL |
| M\_LDL\_P | Concentration of medium LDL particles |
| M\_LDL\_PL | Phospholipids in medium LDL |
| M\_LDL\_PL\_pct | Phospholipids to total lipids ratio in medium LDL |
| M\_LDL\_TG | Triglycerides in medium LDL |
| M\_LDL\_TG\_pct | Triglycerides to total lipids ratio in medium LDL |
| M\_VLDL\_C | Cholesterol in medium VLDL |
| M\_VLDL\_CE | Cholesteryl esters in medium VLDL |
| M\_VLDL\_CE\_pct | Cholesteryl esters to total lipids ratio in medium VLDL |
| M\_VLDL\_C\_pct | Cholesterol to total lipids ratio in medium VLDL |
| M\_VLDL\_FC | Free cholesterol in medium VLDL |
| M\_VLDL\_FC\_pct | Free cholesterol to total lipids ratio in medium VLDL |
| M\_VLDL\_L | Total lipids in medium VLDL |
| M\_VLDL\_P | Concentration of medium VLDL particles |
| M\_VLDL\_PL | Phospholipids in medium VLDL |
| M\_VLDL\_PL\_pct | Phospholipids to total lipids ratio in medium VLDL |
| M\_VLDL\_TG | Triglycerides in medium VLDL |
| M\_VLDL\_TG\_pct | Triglycerides to total lipids ratio in medium VLDL |
| Omega\_3 | Omega-3 fatty acids |
| Omega\_3\_pct | Ratio of omega-3 fatty acids to total fatty acids |
| Omega\_6 | Omega-6 fatty acids |
| Omega\_6\_by\_Omega\_3 | Ratio of omega-6 fatty acids to omega-3 fatty acids |
| Omega\_6\_pct | Ratio of omega-6 fatty acids to total fatty acids |
| PUFA | Polyunsaturated fatty acids |
| PUFA\_by\_MUFA | Ratio of polyunsaturated fatty acids to monounsaturated fatty acids |
| PUFA\_pct | Ratio of polyunsaturated fatty acids to total fatty acids |
| Phe | Phenylalanine |
| Phosphatidylc | Phosphatidylcholines |
| Phosphoglyc | Phosphoglycerides |
| Remnant\_C | Remnant cholesterol (non-HDL, non-LDL -cholesterol) |
| SFA | Saturated fatty acids |
| SFA\_pct | Ratio of saturated fatty acids to total fatty acids |
| S\_HDL\_C | Cholesterol in small HDL |
| S\_HDL\_CE | Cholesteryl esters in small HDL |
| S\_HDL\_CE\_pct | Cholesteryl esters to total lipids ratio in small HDL |
| S\_HDL\_C\_pct | Cholesterol to total lipids ratio in small HDL |
| S\_HDL\_FC | Free cholesterol in small HDL |
| S\_HDL\_FC\_pct | Free cholesterol to total lipids ratio in small HDL |
| S\_HDL\_L | Total lipids in small HDL |
| S\_HDL\_P | Concentration of small HDL particles |
| S\_HDL\_PL | Phospholipids in small HDL |
| S\_HDL\_PL\_pct | Phospholipids to total lipids ratio in small HDL |
| S\_HDL\_TG | Triglycerides in small HDL |
| S\_HDL\_TG\_pct | Triglycerides to total lipids ratio in small HDL |
| S\_LDL\_C | Cholesterol in small LDL |
| S\_LDL\_CE | Cholesteryl esters in small LDL |
| S\_LDL\_CE\_pct | Cholesteryl esters to total lipids ratio in small LDL |
| S\_LDL\_C\_pct | Cholesterol to total lipids ratio in small LDL |
| S\_LDL\_FC | Free cholesterol in small LDL |
| S\_LDL\_FC\_pct | Free cholesterol to total lipids ratio in small LDL |
| S\_LDL\_L | Total lipids in small LDL |
| S\_LDL\_P | Concentration of small LDL particles |
| S\_LDL\_PL | Phospholipids in small LDL |
| S\_LDL\_PL\_pct | Phospholipids to total lipids ratio in small LDL |
| S\_LDL\_TG | Triglycerides in small LDL |
| S\_LDL\_TG\_pct | Triglycerides to total lipids ratio in small LDL |
| S\_VLDL\_C | Cholesterol in small VLDL |
| S\_VLDL\_CE | Cholesteryl esters in small VLDL |
| S\_VLDL\_CE\_pct | Cholesteryl esters to total lipids ratio in small VLDL |
| S\_VLDL\_C\_pct | Cholesterol to total lipids ratio in small VLDL |
| S\_VLDL\_FC | Free cholesterol in small VLDL |
| S\_VLDL\_FC\_pct | Free cholesterol to total lipids ratio in small VLDL |
| S\_VLDL\_L | Total lipids in small VLDL |
| S\_VLDL\_P | Concentration of small VLDL particles |
| S\_VLDL\_PL | Phospholipids in small VLDL |
| S\_VLDL\_PL\_pct | Phospholipids to total lipids ratio in small VLDL |
| S\_VLDL\_TG | Triglycerides in small VLDL |
| S\_VLDL\_TG\_pct | Triglycerides to total lipids ratio in small VLDL |
| TG\_by\_PG | Ratio of triglycerides to phosphoglycerides |
| Total\_BCAA | Total concentration of branched-chain amino acids (leucine + isoleucine + valine) |
| Total\_C | Total cholesterol |
| Total\_CE | Total esterified cholesterol |
| Total\_FA | Total fatty acids |
| Total\_FC | Total free cholesterol |
| Total\_L | Total lipids in lipoprotein particles |
| Total\_P | Total concentration of lipoprotein particles |
| Total\_PL | Total phospholipids in lipoprotein particles |
| Total\_TG | Total triglycerides |
| Tyr | Tyrosine |
| Unsaturation | Degree of unsaturation |
| VLDL\_C | VLDL cholesterol |
| VLDL\_CE | Cholesteryl esters in VLDL |
| VLDL\_FC | Free cholesterol in VLDL |
| VLDL\_L | Total lipids in VLDL |
| VLDL\_P | Concentration of VLDL particles |
| VLDL\_PL | Phospholipids in VLDL |
| VLDL\_TG | Triglycerides in VLDL |
| VLDL\_size | Average diameter for VLDL particles |
| Val | Valine |
| XL\_HDL\_C | Cholesterol in very large HDL |
| XL\_HDL\_CE | Cholesteryl esters in very large HDL |
| XL\_HDL\_CE\_pct | Cholesteryl esters to total lipids ratio in very large HDL |
| XL\_HDL\_C\_pct | Cholesterol to total lipids ratio in very large HDL |
| XL\_HDL\_FC | Free cholesterol in very large HDL |
| XL\_HDL\_FC\_pct | Free cholesterol to total lipids ratio in very large HDL |
| XL\_HDL\_L | Total lipids in very large HDL |
| XL\_HDL\_P | Concentration of very large HDL particles |
| XL\_HDL\_PL | Phospholipids in very large HDL |
| XL\_HDL\_PL\_pct | Phospholipids to total lipids ratio in very large HDL |
| XL\_HDL\_TG | Triglycerides in very large HDL |
| XL\_HDL\_TG\_pct | Triglycerides to total lipids ratio in very large HDL |
| XL\_VLDL\_C | Cholesterol in very large VLDL |
| XL\_VLDL\_CE | Cholesteryl esters in very large VLDL |
| XL\_VLDL\_CE\_pct | Cholesteryl esters to total lipids ratio in very large VLDL |
| XL\_VLDL\_C\_pct | Cholesterol to total lipids ratio in very large VLDL |
| XL\_VLDL\_FC | Free cholesterol in very large VLDL |
| XL\_VLDL\_FC\_pct | Free cholesterol to total lipids ratio in very large VLDL |
| XL\_VLDL\_L | Total lipids in very large VLDL |
| XL\_VLDL\_P | Concentration of very large VLDL particles |
| XL\_VLDL\_PL | Phospholipids in very large VLDL |
| XL\_VLDL\_PL\_pct | Phospholipids to total lipids ratio in very large VLDL |
| XL\_VLDL\_TG | Triglycerides in very large VLDL |
| XL\_VLDL\_TG\_pct | Triglycerides to total lipids ratio in very large VLDL |
| XS\_VLDL\_C | Cholesterol in very small VLDL |
| XS\_VLDL\_CE | Cholesteryl esters in very small VLDL |
| XS\_VLDL\_CE\_pct | Cholesteryl esters to total lipids ratio in very small VLDL |
| XS\_VLDL\_C\_pct | Cholesterol to total lipids ratio in very small VLDL |
| XS\_VLDL\_FC | Free cholesterol in very small VLDL |
| XS\_VLDL\_FC\_pct | Free cholesterol to total lipids ratio in very small VLDL |
| XS\_VLDL\_L | Total lipids in very small VLDL |
| XS\_VLDL\_P | Concentration of very small VLDL particles |
| XS\_VLDL\_PL | Phospholipids in very small VLDL |
| XS\_VLDL\_PL\_pct | Phospholipids to total lipids ratio in very small VLDL |
| XS\_VLDL\_TG | Triglycerides in very small VLDL |
| XS\_VLDL\_TG\_pct | Triglycerides to total lipids ratio in very small VLDL |
| XXL\_VLDL\_C | Cholesterol in chylomicrons and extremely large VLDL |
| XXL\_VLDL\_CE | Cholesteryl esters in chylomicrons and extremely large VLDL |
| XXL\_VLDL\_CE\_pct | Cholesteryl esters to total lipids ratio in chylomicrons and extremely large VLDL |
| XXL\_VLDL\_C\_pct | Cholesterol to total lipids ratio in chylomicrons and extremely large VLDL |
| XXL\_VLDL\_FC | Free cholesterol in chylomicrons and extremely large VLDL |
| XXL\_VLDL\_FC\_pct | Free cholesterol to total lipids ratio in chylomicrons and extremely large VLDL |
| XXL\_VLDL\_L | Total lipids in chylomicrons and extremely large VLDL |
| XXL\_VLDL\_P | Concentration of chylomicrons and extremely large VLDL particles |
| XXL\_VLDL\_PL | Phospholipids in chylomicrons and extremely large VLDL |
| XXL\_VLDL\_PL\_pct | Phospholipids to total lipids ratio in chylomicrons and extremely large VLDL |
| XXL\_VLDL\_TG | Triglycerides in chylomicrons and extremely large VLDL |
| XXL\_VLDL\_TG\_pct | Triglycerides to total lipids ratio in chylomicrons and extremely large VLDL |
| bOHbutyrate | 3-Hydroxybutyrate |
| non\_HDL\_C | Total cholesterol minus HDL-C |

# References

1. Rich-Edwards JW, Fraser A, Lawlor DA, Catov JM. Pregnancy Characteristics and Women's Future Cardiovascular Health: An Underused Opportunity to Improve Women's Health? Epidemiol Rev. 2014;36(1):57-70.

2. Wang Q, Wurtz P, Auro K, Makinen VP, Kangas AJ, Soininen P, et al. Metabolic profiling of pregnancy: cross-sectional and longitudinal evidence. BMC Med. 2016;14(1):205.

3. Mills HL, Patel N, White SL, Pasupathy D, Briley AL, Santos Ferreira DL, et al. The effect of a lifestyle intervention in obese pregnant women on gestational metabolic profiles: findings from the UK Pregnancies Better Eating and Activity Trial (UPBEAT) randomised controlled trial. BMC Med. 2019;17(1):15.

4. Lawlor DA, Emberson JR, Ebrahim S, Whincup PH, Wannamethee SG, Walker M, et al. Is the association between parity and coronary heart disease due to biological effects of pregnancy or adverse lifestyle risk factors associated with child-rearing? Findings from the British Women's Heart and Health Study and the British Regional Heart Study. Circulation. 2003;107(9):1260-4.

5. Wurtz P, Kangas AJ, Soininen P, Lawlor DA, Davey Smith G, Ala-Korpela M. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. Am J Epidemiol. 2017;186(9):1084-96.

6. 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. Circulation. 2015;131(9):774-85.

7. Wang Q, Würtz P, Auro K, Mäkinen V-P, Kangas AJ, Soininen P, et al. Metabolic profiling of pregnancy: cross-sectional and longitudinal evidence. BMC Medicine. 2016;14(1):205.

8. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. International Journal of Epidemiology. 2016;45(6):1866-86.

9. Catov JM, Newman AB, Sutton-Tyrrell K, Harris TB, Tylavsky F, Visser M, et al. Parity and cardiovascular disease risk among older women: how do pregnancy complications mediate the association? Ann Epidemiol. 2008;18(12):873-9.

10. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. Int J Epidemiol. 2016;45(3):908-15.

11. Day FR, Thompson DJ, Helgason H, Chasman DI, Finucane H, Sulem P, et al. Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. Nat Genet. 2017;49(6):834-41.

12. Perry JR, Day F, Elks CE, Sulem P, Thompson DJ, Ferreira T, et al. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature. 2014;514(7520):92-7.

13. Mathieson I, Day FR, Barban N, Tropf FC, Brazel DM, Consortium e, et al. Genome-wide analysis identifies genetic effects on reproductive success and ongoing natural selection at the <em>FADS</em> locus. bioRxiv. 2020:2020.05.19.104455.

14. Ruth KS, Day FR, Hussain J, Martínez-Marchal A, Aiken CE, Azad A, et al. Genetic insights into the biological mechanisms governing human ovarian ageing. medRxiv. 2021:2021.01.11.20248322.

15. Day FR, Ruth KS, Thompson DJ, Lunetta KL, Pervjakova N, Chasman DI, et al. Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. Nat Genet. 2015;47(11):1294-303.

16. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518(7538):197-206.

17. Howe LJ, Nivard MG, Morris TT, Hansen AF, Rasheed H, Cho Y, et al. Within-sibship GWAS improve estimates of direct genetic effects. bioRxiv. 2021:2021.03.05.433935.

18. Davies NM, Howe LJ, Brumpton B, Havdahl A, Evans DM, Davey Smith G. Within family Mendelian randomization studies. Hum Mol Genet. 2019;28(R2):R170-R9.

19. Morris TT, Davies NM, Hemani G, Smith GD. Population phenomena inflate genetic associations of complex social traits. Sci Adv. 2020;6(16):eaay0328.

20. Sanderson E, Richardson TG, Hemani G, Davey Smith G. The use of negative control outcomes in Mendelian randomization to detect potential population stratification. Int J Epidemiol. 2021.

21. Zheng J, Baird D, Borges MC, Bowden J, Hemani G, Haycock P, et al. Recent Developments in Mendelian Randomization Studies. Curr Epidemiol Rep. 2017;4(4):330-45.