

Integrating menopause duration and plasma metabolomics enhances cardiovascular risk stratification in aging women

Received: 19 September 2025

Accepted: 19 December 2025

Cite this article as: Wang, Q., Xie, B., Fu, C. *et al.* Integrating menopause duration and plasma metabolomics enhances cardiovascular risk stratification in aging women. *npj Aging* (2025). <https://doi.org/10.1038/s41514-025-00323-z>

Qi Wang, Bo Xie, Chunying Fu, Meiling Li, Xiaoyi Wang, Nipun Shrestha, Salim S. Virani, Shiva Raj Mishra & Dongshan Zhu

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Original Article

Integrating Menopause Duration and Plasma Metabolomics Enhances Cardiovascular Risk Stratification in Aging Women

Qi Wang^{1†}, Bo Xie^{1†}, Chunying Fu¹, Meiling Li², Xiaoyi Wang¹, Nipun Shrestha³, Salim S. Virani⁴, Shiva Raj Mishra⁵, Dongshan Zhu^{1,6}

¹ Department of Epidemiology, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China

² Medical School of Liaocheng University, Liaocheng, China

³ Health Evidence Synthesis, Recommendations and Impact (HESRI), School of Public Health, The University of Adelaide, Adelaide, Australia.

⁴ The Aga Khan University, Karachi, Pakistan

⁵ School of Medicine, Western Sydney University, Sydney, Australia

⁶ Center for Clinical Epidemiology and Evidence-Based Medicine, Shandong University, Jinan, China

[†]These authors contributed equally to this work

Correspondence to:

Dongshan Zhu

Department of Epidemiology, School of Public Health, Cheeloo College of Medicine, Shandong University, 44 Wenhuxi Road, Jinan 250012, Shandong, China.

Email: dongshan.zhu@uq.net.au

ABSTRACT

Menopause-related metabolic remodeling may contribute to the excess cardiovascular disease (CVD) burden in aging women, yet the longitudinal metabolic correlates of time since menopause (TSM) and their prognostic value are unclear. In this prospective analysis of 67,582 postmenopausal women without baseline CVD from the UK Biobank, we profiled 251 plasma metabolites by nuclear magnetic resonance and followed participants for a median 13.7 years (8313 incident CVD events). Elastic net regression identified a 95-metabolite TSM-associated metabolomic signature (Spearman r with TSM = 0.29). In multivariable Cox models, each 5-year increment in TSM (HR 1.14, 95% CI 1.11–1.16) and each 1-standard deviation increases in the metabolomic signature (HR 1.18, 95% CI 1.15–1.21) were independently associated with higher composite CVD risk, with consistent associations across myocardial infarction, ischemic heart disease, atrial fibrillation, heart failure and stroke. Mendelian randomization supported potential causal roles for 29 of the signature metabolites in CVD. Adding TSM or the metabolomic signature to SCORE2 improved 10-year risk discrimination (area under the curve 0.584 to 0.657 and 0.660, respectively) and reclassification (net reclassification improvement +0.027 and +1.043). These findings implicate cumulative postmenopausal metabolic alterations in vascular risk and support metabolomic enhancement of risk stratification in postmenopausal women.

Keywords: time since menopause; cardiovascular disease; metabolomic signature; prediction; postmenopausal women

Introduction

Cardiovascular diseases (CVD) remain the leading cause of mortality and disease burden globally, particularly among women¹. The risk of CVD increases substantially after menopause, a phenomenon attributed to declining estrogen levels, metabolic disturbances, and vascular dysfunction². While previous studies have primarily focused on traditional CVD risk factors (such as dyslipidemia, hypertension, and insulin resistance) in relation to menopause³, the complex metabolic alterations and underlying biological mechanisms associated with menopausal transition remain incompletely understood.

Recent advances in metabolomics technologies have provided new perspectives for investigating disease mechanisms. Metabolomics, through high-throughput measurement of metabolites in blood or other biological samples, offers comprehensive insights into dynamic metabolic states, reflecting the interplay between genetic and environmental factors while potentially revealing disease pathways⁴. Studies have demonstrated that specific metabolomic signatures are closely associated with CVD and its risk factors⁵. For instance, certain lipid metabolites, amino acids, and intermediates of glucose metabolism have been identified as potential biomarkers for CVD⁶. Metabolomic studies in postmenopausal women are limited, especially regarding the link between time since menopause (TSM), metabolomic signatures, and CVD risk. While metabolomic profiles are associated with cardiovascular outcomes, their ability to enhance risk prediction beyond traditional models remains unclear.

Menopause represents a critical physiological transition in women's lives, with metabolic changes potentially extending over years and having long-term health implications⁷. Age at menopause and time since menopause (defined as the duration from age at menopause to current age) may be key indicators for assessing the metabolic impact of menopause. Previous studies have largely treated menopause as a static exposure variable, overlooking the dynamic changes that occur over time post-menopause⁸. In addition, the metabolic pathways underlying the association between TSM and CVD risk remain poorly understood, and it is unknown whether TSM and its related metabolomic signature can enhance the prediction of

cardiovascular events. Identifying metabolomic signature associated with TSM and elucidating their relationship with cardiovascular risk may offer novel insights into the mechanisms underlying menopause-related CVD and inform risk stratification and intervention strategies for postmenopausal women.

Based on these considerations, this study aims to systematically investigate the relationships between time since menopause (TSM), metabolomic signature, and CVD risk using metabolomics data from the UK Biobank (**Figure 1**). Specific objectives were to: (1) identify and construct a metabolomic signature associated with time since menopause; (2) analyze the associations of time since menopause and the metabolomic signature with CVD risk; (3) investigate potential causal relationships between the metabolomic signature and CVD risk using Mendelian randomization; and (4) assess the predictive value of TSM and the metabolomic signature for CVD risk in postmenopausal women.

Results

Characteristics of the participants

The study included a total of 67,582 participants with metabolites at baseline, with a mean (SD) age of 59.8 (5.4) years. The majority of participants were White (96.35%). After a median follow-up of 13.7 years, 8313 people developed CVD. Participants with incident CVD were more likely to have lower education levels (≤ 10 years: 58.45% vs. 50.40%) and lower annual household income ($<£18,000$: 35.98% vs. 26.33%) compared to those without incident CVD. The incident CVD group also showed a higher prevalence of obesity ($\text{BMI} \geq 30 \text{ kg/m}^2$: 30.57% vs. 20.35%) and current smoking (11.12% vs. 7.48%). Poor sleep quality was more common in the incident CVD group (71.31% vs. 66.67%), as was the use of medications such as statins (19.07% vs. 10.46%) and aspirin (15.64% vs. 7.44%) (**Table 1**). The characteristics of participants at baseline was similar to those observed in the first repeat data (**Supplementary Table S8**).

Metabolomic signature in response to time since menopause (TSM)

The correlation matrix of all metabolites exhibited an obvious clustering pattern due to the strong correlation between lipids and lipoprotein lipid subclasses (**Supplementary Figure S1**). Elastic net regressions on 251 metabolites in baseline data were performed to determine the metabolomic signature in response to TSM. A total of 95 metabolites were selected from the model to calculate the total metabolomic signature, which was significantly correlated with year since menopause (baseline data: $r = 0.29$, $P < 0.001$; first repeat assessment: $r = 0.26$, $P < 0.001$; **Figure 2**). The metabolomic signature spanned various metabolic classes including lipids, lipoprotein subclass, amino acids, fatty acids, ketone bodies, fluid balance-related, glycolysis-related, and inflammation-related metabolites (**Supplementary Figure S2**). KEGG enrichment analysis revealed that these metabolites were predominantly associated with four metabolic pathways: glyoxylate and dicarboxylate metabolism; valine, leucine and isoleucine biosynthesis and degradation; alanine, aspartate and glutamate metabolism; and phenylalanine, tyrosine and tryptophan biosynthesis (**Supplementary Table S9**). As per the metabolites' coefficients (weights) in the signature (**Supplementary Figure S3**), the most pronounced contribution to the positive coefficient of the metabolomic signature came from triglycerides in large LDL, triglycerides in large HDL, and phospholipids to total lipids in Small LDL percentage. Conversely, choline, free cholesterol to total lipids in small VLDL percentage, triglycerides in small LDL, and albumin played a significant role in influencing the reverse coefficient of the metabolomic signature. Associations between TSM and the 95 metabolites comprising the metabolomic signature, as well as the relationships between these metabolites and the risk of CVD and its subtypes were illustrated in **Supplementary Figure S3**.

Associations of time since menopause, metabolomic signature with CVD risk

In the fully adjusted model, i.e., after TSM and metabolomic signature were mutually adjusted, we observed that TSM per 1 unit (representing a 5-year interval) increment associated with an increased risk of CVD (HR 1.14, 95% CI 1.11-1.16) and its subtypes of IHD (1.16, 1.12-1.21), AF (1.15, 1.11-1.19), HF (1.18, 1.12-1.24), MI (1.18, 1.12-1.25) and stroke (1.12, 1.06-1.29) (**Table 2**). The associations also behaved in a dose-response manner (**Supplementary Figure**

S4).

Metabolomic signature per 1-SD increment associated with increased risk of CVD (HR 1.18, 1.15-1.21) and its subtypes of IHD (1.21, 1.17-1.26), AF (1.15, 1.11-1.20), HF (1.30, 1.23-1.36), MI (1.21, 1.15-1.28) and stroke (1.18, 1.11-1.25) (**Table 2**). The dose–response curves depicted an increased trend of CVD and its subtypes' risk with increment of the metabolomic signature (**Supplementary Figure S5**). In addition, a log-rank test indicated a significant difference among individuals with low (Q1), intermediate (Q2-Q4), and high (Q5) metabolomic signature in the probability of CVD and its subtypes at any time point in Kaplan–Meier analysis (**Figure 3**). Pathway-specific analyses also revealed consistent associations with CVD risk across all five group-based metabolomic signatures (**Supplementary Table S10**). Replication analysis using metabolites at first repeat assessment showed similar findings to the primary analysis using metabolites at baseline (**Supplementary Table S11**).

Subgroup analyses and sensitivity analyses

After the analysis was stratified by baseline age (<60 years and ≥ 60 years) and MHT status (ever used and never used), we observed consistent findings as the primary analysis (**Supplementary Table S12**). In sensitivity analyses, the findings remained robust after excluding participants who developed incident CVD within the first three years of follow-up and with additional adjustment for medication use, including menopausal hormone therapy (MHT), lipid-lowering agents, and aspirin (**Supplementary Table S13-S14**).

Proportion mediated by metabolomic signature between TSM and CVD

Mediation analysis showed that the metabolomic signature partially mediated the association between TSM and CVD and its subtypes, accounting for 11.6% of CVD, 11.3% of ischemic heart disease, 10.6% of myocardial infarction, 9.2% of atrial fibrillation, 13.8% of heart failure, and 12.4% of stroke risk (**Supplementary Figure S6**). Further analyses of metabolite classes showed that relative lipoprotein lipid concentrations, lipoprotein subclasses, fatty acids, amino acids, and inflammation-related metabolites had higher mediation proportions

(**Supplementary Figure S7**). These proportions should be interpreted as indicative of potential mechanistic pathways rather than definitive causal effects, given the assumptions required in observational mediation analyses.

Mendelian randomization analyses

For the 95 metabolites identified as components of the metabolomic signature, we conducted two-sample Mendelian randomization to assess the causal associations between each metabolite and CVD as well as its subtypes. No evidence of horizontal pleiotropy and heterogeneity were observed. We identified causal associations for 29 metabolites with CVD, 23 metabolites with MI, 4 metabolites with AF, 6 metabolites with HF, 2 metabolites with IHD, and 7 metabolites with stroke (**Figure 4**). For example, genetically elevated levels of clinical LDL cholesterol, total lipids in medium LDL, saturated fatty acid were associated with a higher CVD risk and genetically elevated levels of phospholipids to total lipids in medium LDL percentage and cholesteryl esters to total lipids in small LDL percentage were found to be linked a lower risk of CVD.

10-year cardiovascular risk prediction with TSM and metabolomic signature

The addition of the metabolomic signature to SCORE2 significantly increased the AUC from 0.584 to 0.660 (DeLong test $P < 0.001$), and the addition of TSM increased the AUC from 0.584 to 0.657 (DeLong test $P < 0.001$). Net reclassification improvement (NRI) analysis further indicated that SCORE2 plus metabolomic signature increased the total NRI by 1.043, and SCORE2 plus TSM increased the total NRI by 0.027, relative to SCORE2 alone (**Figure 5**).

Discussion

This study identified 95 metabolites significantly associated with TSM using elastic net regression, which were primarily involved in lipid metabolism, fatty acids metabolism, and amino acid metabolism. A metabolomic signature was constructed based on these metabolites. We found that each SD increase in the TSM-related metabolomic signature was significantly associated with a higher risk of CVD. Additionally, the metabolomic signature mediated the

relationship between TSM and CVD risk, suggesting that it may be a critical mechanism underlying the increased CVD risk observed in postmenopausal women. Mendelian Randomization (MR) analysis further revealed that 37 of the 95 metabolites were potentially causally associated with CVD, supporting the potential utility of the metabolomic signature in assessing CVD risk among postmenopausal women.

Premature menopause (<40 years) has been recognized as a risk enhancer for atherosclerotic cardiovascular disease due to the cardiometabolic changes that occur earlier³. While 1% of women undergo premature menopause, up to 10% of women experience early menopause, defined as <45 years⁹. A growing body of evidence has demonstrated that early menopause (<45 years) significantly increases the risk of CVD. For instance, Honigberg et al. (2019) reported in a study involving 144,260 women that early menopause was associated with an increased risk of coronary heart disease (CHD) (HR: 1.32, 95% CI: 1.10-1.57) and stroke (HR: 1.23, 95% CI: 1.03-1.46)¹⁰. Similarly, Zhu et al. (2019), in a pooled analysis of individual patient data, found that each additional year of delay in menopause reduced CHD risk by 3% (HR: 0.97, 95% CI: 0.96-0.98) and stroke risk by 2% (HR: 0.98, 95% CI: 0.97-0.99)¹¹. Moreover, time since menopause has been identified as a critical predictor of CVD risk. Muka et al. (2016) reported that CVD risk significantly increased within the first 10 years post-menopause (HR: 1.12, 95% CI: 1.05-1.20) and plateaued thereafter⁸. The associations between age at menopause and CVD subtypes is not consistent. A nationwide cohort study of Korea found that postmenopausal women with a history of premature menopause had 33% higher risk of heart failure (HR 1.33, 95% CI 1.26–1.40) and 9% higher risk of Atrial fibrillation (HR 1.09, 95% CI 1.02–1.16)¹².

While the menopause transition contributes to activation of the renin-angiotensin-aldosterone system as a result of estrogen loss, the current analysis on MHT and CVD suggesting that the mechanism underlying the increased risk cannot be attributed solely to estrogen. Instead, estrogen deficiency may be just one component of the causal link between menopause and CVD. Additional intermediate or mediating factors likely contribute to the relationship, warranting

further investigation⁹. Despite growing epidemiological evidence, the mechanistic link between menopause and CVD remains unclear, with limited insights from omics, especially metabolomics. Existing studies indicated that postmenopausal women experience significant metabolic changes, including disruptions in lipid, amino acid, and glucose metabolism^{7,13}, which may underlie increased cardiovascular risk. These findings suggest that analyzing postmenopausal metabolomic profiles could unveil the mechanisms bridging menopause and CVD risk, offering insights for precision intervention strategies.

Menopause marks a key life transition for women, defined by significant endocrine changes, particularly a decline in estrogen. These hormonal shifts lead to extensive remodeling of the metabolomic profile. These metabolic disruptions not only elucidate the mechanisms underlying the elevated CVD risk observed in postmenopausal women but also highlight potential biomarkers and therapeutic targets for intervention.

First, menopause induces marked changes in lipid metabolism, evidenced by increased levels of low-density lipoprotein cholesterol (LDL-C) and very-low-density lipoprotein (VLDL) particles, as well as diminished functionality of high-density lipoprotein cholesterol (HDL-C)^{7,14}. Studies have demonstrated that HDL particles in postmenopausal women tend to shift toward smaller, less functional subtypes, thereby impairing cholesterol efflux capacity and contributing to atherogenesis^{15,6}. Concurrently, elevated triglyceride (TG) levels exacerbate dyslipidemia, a key driver of heightened CVD risk, particularly atherosclerotic cardiovascular disease (ASCVD)^{13,16}. Our study also found that over 50% of postmenopausal altered metabolites are lipid-related metabolites. Consistently, we observed more stable associations among participants with a higher baseline lipid burden. In addition, the mediation analysis indicated that lipoprotein lipids accounted for the largest proportion of the association between TSM and CVD.

Second, menopause significantly impacts amino acid metabolism, particularly branched-chain amino acids (BCAAs) such as leucine, isoleucine, and valine, as well as aromatic amino acids (AAAs) like phenylalanine and tyrosine^{17,18}. Elevated BCAA levels have been strongly linked

to insulin resistance, chronic inflammation, and adipose tissue dysfunction^{19,20}. Moreover, BCAAs may exacerbate oxidative stress and endothelial dysfunction, further amplifying CVD risk^{21,22}. Similarly, elevated AAA levels have been implicated in vascular damage and atherogenesis through the activation of inflammatory signaling pathways, such as nuclear factor kappa B (NF- κ B)^{16,23}. In the pathway-based analysis, our constructed amino acid signature showed stronger associations with CVD and its subtypes than other pathway-based signatures, indicating that postmenopausal amino acid metabolic changes are more closely associated with CVD.

Third, glucose metabolism dysregulation represents another hallmark of postmenopausal metabolic changes. Postmenopausal women frequently exhibit reduced insulin sensitivity and elevated fasting glucose levels^{15,24}. This insulin-resistant state not only directly promotes atherogenesis but also interacts with lipid metabolism to accelerate cardiovascular complications^{6,23}. Furthermore, estrogen deficiency has been shown to influence hepatic lipid metabolism by modulating gene expression, indirectly disrupting fatty acid synthesis and breakdown, thus compounding glucose-lipid metabolic dysfunction^{14,21}.

Finally, menopause-associated metabolic changes are intricately linked to inflammatory and oxidative stress pathways, which further exacerbate cardiovascular risk. Elevated levels of advanced glycation end products (AGEs) and glycoprotein acetyls (GlycA), both inflammatory markers strongly associated with CVD, have been observed in postmenopausal women^{18,20}. Additionally, the accumulation of lipid peroxidation products may activate the NF- κ B signaling pathway, triggering inflammatory responses and vascular damage^{17,22}.

A study revealed that the majority of metabolic alterations observed at the onset of menopause persisted or intensified throughout the postmenopausal period. This suggests the persistence of postmenopausal metabolomic changes and their potential cumulative impact on future cardiovascular disease (CVD) risk²⁴. We observed that the metabolomic signature accounted for 12% of the association between TSM and CVD risk, highlighting the unique biological pathways identified through metabolomics. This signature represents a comprehensive

depiction of overall metabolic homeostasis in response to TSM. Leveraging the metabolomic signature allows for a more objective and direct understanding of the cumulative metabolic effects associated with time since menopause.

The observed discrete effects (HRs 1.14–1.18) indicate small relative risk increases. Among high baseline risk populations or over extended follow-up periods, these small relative risks may accumulate into more appreciable absolute risk differences. Additionally, from a public health perspective, even small relative risks can have significant implications for overall disease burden when applied to large populations or prolonged exposures. Thus, these discrete effects are better interpreted as incremental signals for risk stratification and predictive modeling rather than standalone triggers for clinical intervention.

One strength of this study is its large-scale prospective cohort with long-term follow-up (~14 years) and standardized data collection, ensuring robust statistical power and reliable outcomes. The use of untargeted metabolomics reveals diverse metabolic changes beyond predefined pathways, aiding the discovery of novel biomarkers linked to postmenopausal aging. Its longitudinal design captures cumulative metabolic effects across different durations since menopause. Additionally, the well-characterized cohort and detailed clinical and biochemical assessments allow for comprehensive adjustments for potential confounders, including lifestyle factors, comorbidities, and MHT status. However, several limitations should be acknowledged.

First, our study was conducted in UK Biobank, a cohort that is predominantly White and generally healthier than the broader population. These characteristics may limit generalizability, especially for absolute risk predictions. Future work will focus on external validation and model adjustment in more diverse cohorts and healthcare settings to assess transportability and utility. Second, we observed sociodemographic differences between participants with and without CVD. Although these variables were adjusted for in our analysis, residual confounding may still exist due to unmeasured factors or their dynamic changes over time. Accordingly, the reported hazard ratios are best interpreted as relative associations, not causal effects, and should not be interpreted as direct evidence of biological mechanisms or intervention-responsive

causal impacts. Third, in this study, we used Cox proportional hazards models and treated non-CVD death as independent censoring. This does not account for non-CVD death as a competing event and may bias interpretation of cumulative incidence, particularly in populations with a high incidence of non-CVD death. Fourth, one limitation of our mediation analysis is the reliance on strong identification assumptions, including no unmeasured confounding among TSM, the metabolomic signature, and CVD, and the correct model specification. These assumptions cannot be fully verified in observational studies, and the reported mediated proportions should be interpreted as indicative of potential mechanistic pathways rather than definitive causal effects.

In this large-scale prospective cohort study of postmenopausal women, we demonstrated that time since menopause is significantly associated with adverse alterations in the metabolomic profile and increased cardiovascular disease (CVD) risk, independent of chronological age and traditional risk factors. We identified a metabolomic signature linked to time since menopause that not only predicted incident CVD but also partially mediated the association between menopause duration and CVD risk. Mendelian randomization further supported the causal roles of several metabolites in CVD development. These findings highlight the importance of metabolic changes following menopause in shaping long-term cardiovascular health and suggest that metabolomic profiling may provide novel biomarkers and mechanistic insights for CVD risk stratification and prevention in postmenopausal women.

Methods

Study design and participants

The UK Biobank is a prospective population-based cohort, recruited over 500,000 volunteers aged 40-69 years between 2006 and 2010. Individuals were invited to attend one of the 22 centers across England, Scotland, and Wales for baseline assessment. Written informed consent was obtained for collection of questionnaire and biological data. UK Biobank has approval from the North West Multicenter Research Ethics Committee (<https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us/ethics>). This research was conducted under UK Biobank application

number 227947. Women who were postmenopausal at baseline and had no missing key covariates were included. A prospective design was adopted based on participants with no CVD at baseline (**Figure 6**). This study is reported as per the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (**supplementary Table S1**).

Age at natural menopause and time since menopause

Natural menopause was defined as the absence of menstruation for a consecutive 12-month period, without a history of hysterectomy and/or oophorectomy prior to this time. The duration of the postmenopausal period, measured in time since menopause (TSM), was calculated by subtracting the age at natural menopause from the baseline age at enrollment. For analytical purposes, the calculated time since menopause was standardized by dividing the value by 5, such that a single unit corresponds to a 5-year interval. To limit misclassification from self-reported menopause, we used a standardized definition, excluded uncertain cases (e.g., hysterectomy without bilateral oophorectomy if <55 years), validated self-reports against age, reproductive/surgical history, and medication records, and adjusted for menopausal hormone therapy (MHT).

Metabolomics profiling

This study employed a high-throughput nuclear magnetic resonance (NMR) metabolomics platform (Nightingale Health Ltd, Finland) to analyze EDTA plasma samples from approximately 280,000 participants in the UK Biobank. A total of 251 metabolic biomarkers were measured, including 170 metabolites quantified in absolute levels and 81 composite ratio indices (**supplementary Table S2**). These biomarkers represent a diverse array of metabolic pathways, encompassing 14 subclasses of lipoprotein lipids, fatty acids and their compositions, as well as low-molecular weight metabolites such as amino acids, ketone bodies, and glycolysis-related metabolites²⁵. To ensure data integrity, rigorous quality control measures were implemented. Natural logarithm transformations and z-score standardization (mean 0, SD 1) were applied to the metabolite concentration data to mitigate systemic and technical variability.

The data analyzed were derived from phase 1 and phase 2 of the UK Biobank study, with around 16,000 participants completing a repeat assessment visit. All metabolic biomarkers were quantified in molar concentration units, providing a comprehensive characterization of the metabolic profiles of the participants. Detailed methodologies regarding the measurement of metabolic biomarkers and quality control procedures have been published in the relevant literature²⁶.

Ascertainment of cardiovascular diseases

The primary endpoint of the study was the occurrence of incident CVD, defined as a composite outcome encompassing myocardial infarction (MI), chronic ischemic heart disease (IHD), heart failure (HF), atrial fibrillation (AF), and stroke. Among these, MI and stroke were identified within category 47 of the UK Biobank, based on algorithmically defined outcomes derived from linkage of baseline self-reported data, hospital admission records, and death registry data²⁷. IHD, HF and AF were defined based on the first occurrence data assembled in category 1712 from the UK Biobank. The data includes the first reported occurrence of a disease in either the linked primary care data (cat. 3000), inpatient hospital admissions (cat. 2000), death registry (fields: 40001, 40002) or self-reported data through questionnaires (field: 20002). Detailed information on field IDs and outcome cases was provided in the **Supplementary Table S3**.

Covariates

We included the following factors in the analyses as covariates according to evidence from previous studies: age at baseline, race/ethnicity, years of education, income levels, Townsend deprivation index, smoking status, alcohol intakes, BMI, sleep status, diet, family history of CVD, and medications, including menopause hormone therapy status, aspirin use, and statins use. Ethnicity was categorized as White, and non-white. Years of education was categorized as ≤10, 11-12, >12. Annual household income level was divided into four categories as level 1 (Less than £18,000), level 2 (£18,000 to £30,999), level 3 (£31,000 to £51,999), and level 4 (greater than 52,000). Townsend deprivation index, which reflects the area-level socio-economic status, was based on participants' residential postcode at recruitment and categorized

based on quartile. Higher value indicates greater levels of deprivation. Smoking status was categorized as current, former, or never smokers. Alcohol intake was categorized as current, former, or never drinks. BMI was categorized according to the World Health Organization criteria as $< 18.5 \text{ kg/m}^2$, $18.5\text{-}24.9 \text{ kg/m}^2$, $25.0\text{-}29.9 \text{ kg/m}^2$ and $\geq 30 \text{ kg/m}^2$. Diet consumption was classified into ideal or poor according to whether adequate intake of at least half of 10 diet components²⁸. Sleep quality was defined by the criteria recommended by the National Sleep Foundation, which integrate five sleep behaviors (sleep duration, chronotype preference, insomnia, snoring, and daytime sleepiness)²⁹. Participants were divided into two groups: “good sleep quality” and “poor sleep quality”. The family history of CVD was categorized as yes or no. The status of medication usage was categorized into use and non-use. For detailed information on covariate collection and definitions see **Supplementary Table S4-S6**.

Statistical analysis

Identification of the metabolomic signature reflecting menopause duration

We identified metabolites associated with TSM in postmenopausal women using baseline metabolomics data from the UK Biobank. To exclude outliers, metabolite levels falling outside of 4 interquartile ranges from the median were removed. Prior to conducting the planned analyses, all 251 metabolites were log-transformed and standardized to the same scale using z-scores, and correlations among them were examined using the Pearson correlation coefficient. The associations of each metabolite with TSM were explored using multivariable linear regression, with a false discovery rate (FDR) adjusted p-value of 0.05 considered statistically significant.

To identify a metabolomic signature associated with TSM, we applied an elastic net regression model. This regularized regression approach combines the strengths of both Lasso and Ridge methods, effectively addressing multicollinearity, reducing overfitting, and selecting relevant features. Specifically, we regressed TSM on 251 named and standardized plasma metabolites. The optimal penalty parameter (lambda) was determined using a 10-fold cross-validation procedure, selecting the largest lambda value that produced a mean squared error within one

standard error of the minimum. The resulting metabolomic signature was calculated as the weighted sum of metabolites with nonzero coefficients, where the weights corresponded to the coefficients estimated by the elastic net model. Finally, the metabolomic signature was standardized using z-scores (mean 0, SD 1), with each unit change in the signature reflecting the combined effect of the selected metabolites' weighted contributions³⁰. We evaluated the Spearman correlation between TSM and the derived metabolomic signature at baseline (2006–2010) and the first repeat assessment visit (2012–2013).

To further investigate the patterns of the metabolomic signature, we stratified it into three categories based on quintiles: low (Q1), intermediate (Q2–Q4), and high (Q5). This categorization allowed us to examine whether the metabolic alterations associated with post-menopausal duration showed distinct patterns across different groups. Additionally, we built pathway-based signatures. The 251 metabolites were grouped into five categories—amino acids, fatty acids, lipoprotein subclasses, relative lipoprotein lipid concentrations, and other metabolites. For each group, we used the same workflow to derive TSM-associated signatures and assessed their associations with CVD.

Associations between time since menopause, the metabolomic signature and CVD risk

Cox proportional hazard models were also conducted to obtain HRs and 95% CIs for the associations of time since menopause and metabolomic signature with CVD risk. We first assessed the proportional hazards assumption by analyzing the relationship between standardized Schoenfeld residuals and time, confirming that the assumption was not violated. Four nested models were developed, sequentially including four sets of covariates to account for potential confounders. In model 1, age was adjusted; in model 2, ethnicity, socioeconomic status variables of education, income and Townsend index of deprivation were further adjusted based on model 1; in model 3, smoking status, alcohol intake, sleep status, diet status, BMI, and family history of CVD were further adjusted based on model 2; in model 4 (full adjusted model), we further included mutual adjustments for both TSM and metabolomic signature to assess their independent associations. We also examined potential non-linear associations

between metabolomic signature, TSM, and CVD risk using restricted cubic spline analysis. Meanwhile, Kaplan-Meier survival curves were plotted to estimate the CVD risk between low, intermediate, and high levels of metabolomic signature groups.

Subgroup analysis and sensitivity analysis

To distinguish the effects of chronological age from postmenopausal duration, we conducted a stratified analysis based on baseline age (<60 years and ≥ 60 years), as these two factors are closely intertwined. We also examined whether the associations of TSM and metabolomic signature with CVD varied by baseline MHT status. To assess the robustness of these findings, several sensitivity analyses were performed, including: (1) exclusion of participants who developed incident CVD within the first 3 years of follow-up; and (2) additional adjustment for medication use, including menopausal hormone therapy (MHT), lipid-lowering drugs, and aspirin.

Mediation analysis

We conducted mediation analyses to investigate whether the association between TSM and CVD could be mediated by metabolomic signature. Using the ‘mediation’ packages of R software, the proportion of mediation was determined by dividing the indirect effect by the total effect, with 95% CIs estimated via bootstrapping. Covariates from the multivariate-adjusted Cox model were included in the mediation analysis to account for potential confounders. Furthermore, we evaluated the contribution of specific metabolite classes in mediating the association between TSM and the risk of CVD, including its subtypes.

Mendelian randomization analyses

Two-sample MR analyses were performed to explore the potential causal associations between the identified metabolites and risk of CVD. We obtained the genetic variants that were considered as instrumental variables for MR from published genome-wide association study summary statistics for each identified metabolite. Summary statistics for CVD and its subtypes were derived from open genome wide association study projects (**Supplementary Table S7**).

To ensure independence, exposure and outcome GWAS were drawn from independent consortia with non-overlapping cohorts.

We conducted stringent linkage disequilibrium (LD) clumping thresholds (10 000 kb clumping window and r^2 threshold = 0.1) to select independent single nucleotide polymorphisms. The random-effects inverse-variance weighted (IVW) method was employed as the primary analysis^{31,32}, as it yields a more precise Mendelian randomization estimate when all instruments are valid³³.

10-year cardiovascular risk prediction with TSM and metabolomic signature

To assess the potential improvement in 10-year cardiovascular risk prediction by incorporating time since menopause (TSM) and a metabolomic signature, we added these variables to the recently introduced SCORE2 model, which was developed in 2021 for European individuals without prior cardiovascular disease³⁴. Individuals with a 10-year cardiovascular risk of $\geq 7.5\%$ were considered to be at high risk. Model performance improvement was evaluated using the area under the curve (AUC) and net reclassification improvement (NRI).

Two-sided $P < 0.05$ was considered statistically significant. P-values were corrected for multiple comparison using the Benjamini-Hochberg FDR adjustment. The primary R packages used in the analysis included ‘glmnet’ for elastic net regression and ‘survival’ for Cox regression.

Acknowledgements

This research has been conducted using the UK Biobank resource (application number 227947). The authors are grateful to UK Biobank participants. We acknowledge the financial support from the National Natural Science Foundation of China (82273702). The funding sources had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Author contributions

All the authors contributed substantially to the completion of this study. DZ conceived of the presented idea. QW and BX performed the data analyses, established the machine learning models, and drafted the manuscript. QW, BX, CF, ML, and XW participated in data collection. NS, SSV, SRM provided critical feedback. QW, BX, and DZ had primary responsibility for study design and data interpretation. All authors have read and approved the final version of the manuscript.

Competing interests

The authors have no conflict of interest to declare.

REFERENCE

- 1 Vervoort, D. *et al.* Addressing the Global Burden of Cardiovascular Disease in Women: JACC State-of-the-Art Review. *J Am Coll Cardiol* **83**, 2690-2707, doi:10.1016/j.jacc.2024.04.028 (2024).
- 2 Davis, S. R., Pinkerton, J., Santoro, N. & Simoncini, T. Menopause-Biology, consequences, supportive care, and therapeutic options. *Cell* **186**, 4038-4058, doi:10.1016/j.cell.2023.08.016 (2023).
- 3 El Khoudary, S. R. *et al.* Menopause Transition and Cardiovascular Disease Risk: Implications for Timing of Early Prevention: A Scientific Statement From the American Heart Association. *Circulation* **142**, e506-e532, doi:10.1161/CIR.0000000000000912 (2020).
- 4 Fuller, H. *et al.* Metabolomic epidemiology offers insights into disease aetiology. *Nat Metab* **5**, 1656-1672, doi:10.1038/s42255-023-00903-x (2023).
- 5 Ajoolabady, A., Pratico, D., Dunn, W. B., Lip, G. Y. H. & Ren, J. Metabolomics: Implication in cardiovascular research and diseases. *Obes Rev* **25**, e13825, doi:10.1111/obr.13825 (2024).
- 6 Iliou, A. *et al.* Metabolic phenotyping and cardiovascular disease: an overview of evidence from epidemiological settings. *Heart* **107**, 1123-1129, doi:10.1136/heartjnl-2019-315615 (2021).
- 7 Clayton, G. L., Borges, M. C. & Lawlor, D. A. The impact of reproductive factors on the metabolic profile of females from menarche to menopause. *Nat Commun* **15**, 1103, doi:10.1038/s41467-023-44459-6 (2024).
- 8 Muka, T. *et al.* Association of Age at Onset of Menopause and Time Since Onset of Menopause With Cardiovascular Outcomes, Intermediate Vascular Traits, and All-Cause Mortality: A Systematic Review and Meta-analysis. *JAMA Cardiol* **1**, 767-776, doi:10.1001/jamacardio.2016.2415 (2016).
- 9 Torbati, T., Shufelt, C., Wei, J. & Noel Bairey Merz, C. Premature menopause and cardiovascular disease: can we blame estrogen? *Eur Heart J* **43**, 4158-4160,

- doi:10.1093/eurheartj/ehac321 (2022).
- 10 Honigberg, M. C. *et al.* Association of Premature Natural and Surgical Menopause With Incident Cardiovascular Disease. *JAMA* **322**, 2411-2421, doi:10.1001/jama.2019.19191 (2019).
 - 11 Zhu, D. *et al.* Age at natural menopause and risk of incident cardiovascular disease: a pooled analysis of individual patient data. *Lancet Public Health* **4**, e553-e564, doi:10.1016/S2468-2667(19)30155-0 (2019).
 - 12 Shin, J. *et al.* Age at menopause and risk of heart failure and atrial fibrillation: a nationwide cohort study. *Eur Heart J* **43**, 4148-4157, doi:10.1093/eurheartj/ehac364 (2022).
 - 13 Auro, K. *et al.* A metabolic view on menopause and ageing. *Nat Commun* **5**, 4708, doi:10.1038/ncomms5708 (2014).
 - 14 Tian, X. *et al.* Association of lipid, inflammatory, and metabolic biomarkers with age at onset for incident cardiovascular disease. *BMC Med* **20**, 383, doi:10.1186/s12916-022-02592-x (2022).
 - 15 Streese, L. *et al.* Metabolic profiling links cardiovascular risk and vascular end organ damage. *Atherosclerosis* **331**, 45-53, doi:10.1016/j.atherosclerosis.2021.07.005 (2021).
 - 16 Ussher, J. R., Elmariah, S., Gerszten, R. E. & Dyck, J. R. The Emerging Role of Metabolomics in the Diagnosis and Prognosis of Cardiovascular Disease. *J Am Coll Cardiol* **68**, 2850-2870, doi:10.1016/j.jacc.2016.09.972 (2016).
 - 17 Lv, J. *et al.* Plasma metabolomics reveals the shared and distinct metabolic disturbances associated with cardiovascular events in coronary artery disease. *Nat Commun* **15**, 5729, doi:10.1038/s41467-024-50125-2 (2024).
 - 18 McGarrah, R. W., Crown, S. B., Zhang, G. F., Shah, S. H. & Newgard, C. B. Cardiovascular Metabolomics. *Circ Res* **122**, 1238-1258, doi:10.1161/CIRCRESAHA.117.311002 (2018).
 - 19 Ruiz-Canela, M. *et al.* Comprehensive Metabolomic Profiling and Incident Cardiovascular Disease: A Systematic Review. *J Am Heart Assoc* **6**, doi:10.1161/JAHA.117.005705 (2017).
 - 20 Tahir, U. A. *et al.* Metabolomic Profiles and Heart Failure Risk in Black Adults: Insights From the Jackson Heart Study. *Circ Heart Fail* **14**, e007275,

- doi:10.1161/CIRCHEARTFAILURE.120.007275 (2021).
- 21 Zhang, B. & Schmidlin, T. Recent advances in cardiovascular disease research driven by metabolomics technologies in the context of systems biology. *npj Metabolic Health and Disease* **2**, doi:10.1038/s44324-024-00028-z (2024).
 - 22 Muller, J. *et al.* Narrative review of metabolomics in cardiovascular disease. *J Thorac Dis* **13**, 2532-2550, doi:10.21037/jtd-21-22 (2021).
 - 23 Li, Y. *et al.* Metabolomic Profiles, Ideal Cardiovascular Health, and Risk of Heart Failure and Atrial Fibrillation: Insights From the Framingham Heart Study. *J Am Heart Assoc* **12**, e028022, doi:10.1161/JAHA.122.028022 (2023).
 - 24 Wang, Q. *et al.* Metabolic characterization of menopause: cross-sectional and longitudinal evidence. *BMC Med* **16**, 17, doi:10.1186/s12916-018-1008-8 (2018).
 - 25 Ritchie, S. C. *et al.* Quality control and removal of technical variation of NMR metabolic biomarker data in ~120,000 UK Biobank participants. *Sci Data* **10**, 64, doi:10.1038/s41597-023-01949-y (2023).
 - 26 Wurtz, P. *et al.* Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies. *Am J Epidemiol* **186**, 1084-1096, doi:10.1093/aje/kwx016 (2017).
 - 27 Rani, V., Deep, G., Singh, R. K., Palle, K. & Yadav, U. C. Oxidative stress and metabolic disorders: Pathogenesis and therapeutic strategies. *Life Sci* **148**, 183-193, doi:10.1016/j.lfs.2016.02.002 (2016).
 - 28 Said, M. A., Verweij, N. & van der Harst, P. Associations of Combined Genetic and Lifestyle Risks With Incident Cardiovascular Disease and Diabetes in the UK Biobank Study. *JAMA Cardiol* **3**, 693-702, doi:10.1001/jamacardio.2018.1717 (2018).
 - 29 Hirshkowitz, M. *et al.* National Sleep Foundation's sleep time duration recommendations: methodology and results summary. *Sleep Health* **1**, 40-43, doi:10.1016/j.sleh.2014.12.010 (2015).
 - 30 Hui, Z. & Trevor, H. Regularization and Variable Selection Via the Elastic Net. *J R Statist Soc B* **67**, 301–320, doi:<https://doi.org/10.1111/j.1467-9868.2005.00503.x> (2005).

- 31 Burgess, S. *et al.* Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol* **30**, 543-552, doi:10.1007/s10654-015-0011-z (2015).
- 32 Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* **37**, 658-665, doi:10.1002/gepi.21758 (2013).
- 33 Burgess, S., Bowden, J., Fall, T., Ingelsson, E. & Thompson, S. G. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology* **28**, 30-42, doi:10.1097/EDE.0000000000000559 (2017).
- 34 group, S. w. & collaboration, E. S. C. C. r. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. *Eur Heart J* **42**, 2439-2454, doi:10.1093/eurheartj/ehab309 (2021).

Figure legends

Figure 1. Overview of time since menopause (TSM), metabolomic signature and cardiovascular diseases (CVD) prediction in women. MI, myocardial infarction; IHD, chronic ischemic heart disease; HF, heart failure; AF, atrial fibrillation; MHT, menopausal hormone therapy, AUC: area under the curve; NRI: net reclassification improvement.

Figure 2. Correlations between time since menopause and the metabolomic signature using baseline data (a) and first repeat data (b)

Figure 3. Cumulative incidence of CVD according to metabolomic signature categorizes of low, intermediate and high

Figure 4. Mendelian randomization analyses of associations between genetically determined metabolites with CVD

Figure 5. Area under the curve (AUC) and net reclassification improvement (NRI) index relative to SCORE2 based on the 10-year CVD risk predicted. (A) AUC after adding time since menopause and metabolomic signature to SCORE2; (B) NRI compared (SCORE2+Time since menopause) to SCORE2 alone; (C) NRI compared (SCORE2+Metabolomic signature) to SCORE2 alone.

Figure 6. Flow chart of study design and analytical approach

Table 1. Characteristics of participants at baseline by CVD incidence

	Overall participants	No incident CVD		Incident CVD	
		N=59269	%	N=8313	%
Age at baseline	59.8±5.4	59.5±5.4		62.0±5.0	
Race/ethnicity					
White	65141	57107	96.4	8034	96.6
Non-white	2441	2162	3.6	279	3.4
Education level (years)					
≤10	34735	29876	50.4	4859	58.5
10-11	7766	6921	11.7	845	10.2
>12	25081	22472	37.9	2609	31.3
Income level (£)					
Less than 18,000	18598	15607	26.3	2991	36.0
18,000 to 30,999	19032	16680	28.1	2352	28.3
31,000 to 51999	16486	14773	24.9	1713	20.6
Greater than 52,000	13466	12209	20.7	1257	15.1
Socioeconomic status					
Q1	16781	14943	25.2	1838	22.1
Q2	16998	15025	25.3	1973	23.7
Q3	16897	14857	25.1	2040	24.5
Q4	16906	14444	24.4	2462	29.6
BMI level (kg/m²)					
<18.5	513	449	0.7	64	0.8
18.5-24.9	26632	24045	40.6	2587	31.1
25.0-29.9	25834	22713	38.3	3121	37.5
≥30	14603	12062	20.4	2541	30.6
Smoking status					
Never	39679	35230	59.4	4449	53.5
Former	22543	19603	33.1	2940	35.4
Current	5360	4436	7.5	924	11.1
Diet status					
Poor	42435	37094	62.6	5341	64.3
Good	25147	22175	37.4	2972	35.7
Alcohol status					
Never	3598	3021	5.1	577	7.0
Former	2216	1832	3.1	384	4.6
Current	61768	54416	91.8	7352	88.4
Sleep quality					
Poor	45445	39517	66.7	5928	71.3
Good	22137	19752	33.3	2385	28.7
Family history of CVD					
No	25819	22957	38.7	2862	34.4
Yes	41763	36312	61.3	5451	65.6
Statins use					
No	59796	53068	89.5	6728	80.9
Yes	7786	6201	10.5	1585	19.1
Aspirin use					
No	61872	54859	92.6	7013	84.4

Yes	5710	4410	7.4	1300	15.6
HRT use					
No	39158	34888	58.9	4270	51.4
Yes	28424	24381	41.1	4043	48.6
Time since menopause					
1	17272	16048	27.1	1224	14.7
2	19262	17258	29.1	2004	24.1
3	16826	14507	24.5	2319	27.9
4	9714	7910	13.4	1804	21.7
5	4508	3546	6.0	962	11.6

Table 2. Associations between time since menopause (TSM), metabolomic signature (at baseline) and CVD risk

	TSM per 1 increment*		Metabolomic signature per 1-SD increment	
	HR (95%CI)	P	HR (95%CI)	P
CVD				
Model 1	1.17 (1.15, 1.20)	<0.001	1.25 (1.22, 1.28)	<0.001
Model 2	1.15 (1.13, 1.18)	<0.001	1.24 (1.21, 1.27)	<0.001
Model 3	1.14 (1.12, 1.17)	<0.001	1.19 (1.17, 1.22)	<0.001
Model 4	1.14 (1.11, 1.16)	<0.001	1.18 (1.15, 1.21)	<0.001
MI				
Model 1	1.24 (1.17, 1.31)	<0.001	1.30 (1.23, 1.37)	<0.001
Model 2	1.21 (1.15, 1.28)	<0.001	1.28 (1.21, 1.35)	<0.001
Model 3	1.19 (1.13, 1.26)	<0.001	1.23 (1.17, 1.30)	<0.001
Model 4	1.18 (1.12, 1.25)	<0.001	1.21 (1.15, 1.28)	<0.001
AF				
Model 1	1.17 (1.13, 1.21)	<0.001	1.22 (1.18, 1.26)	<0.001
Model 2	1.16 (1.12, 1.20)	<0.001	1.21 (1.17, 1.25)	<0.001
Model 3	1.15 (1.12, 1.20)	<0.001	1.17 (1.13, 1.21)	<0.001
Model 4	1.15 (1.11, 1.19)	<0.001	1.15 (1.11, 1.20)	<0.001
HF				
Model 1	1.24 (1.18, 1.31)	<0.001	1.43 (1.36, 1.50)	<0.001
Model 2	1.21 (1.15, 1.27)	<0.001	1.40 (1.33, 1.47)	<0.001
Model 3	1.19 (1.13, 1.25)	<0.001	1.31 (1.25, 1.38)	<0.001
Model 4	1.18 (1.12, 1.24)	<0.001	1.30 (1.23, 1.36)	<0.001
IHD				
Model 1	1.21 (1.17, 1.26)	<0.001	1.30 (1.25, 1.34)	<0.001
Model 2	1.19 (1.14, 1.23)	<0.001	1.28 (1.23, 1.33)	<0.001
Model 3	1.17 (1.13, 1.21)	<0.001	1.23 (1.18, 1.27)	<0.001
Model 4	1.16 (1.12, 1.21)	<0.001	1.21 (1.17, 1.26)	<0.001
Stroke				
Model 1	1.16 (1.10, 1.23)	<0.001	1.23 (1.16, 1.30)	<0.001
Model 2	1.14 (1.08, 1.21)	<0.001	1.21 (1.14, 1.28)	<0.001
Model 3	1.13 (1.07, 1.19)	<0.001	1.19 (1.12, 1.26)	<0.001
Model 4	1.12 (1.06, 1.19)	<0.001	1.18 (1.11, 1.25)	<0.001

CI: confidence interval, HR: hazard ratio; CVD: cardiovascular diseases; MI: myocardial infarction; AF: atrial fibrillation; HF: heart failure; IHD: ischemic heart disease. Model 1 adjusted for age. Model 2 further adjusted for ethnicity, education, income and Townsend index of deprivation based on model 1. Model 3 further adjusted for smoking status, alcohol intake, sleep status, diet status, BMI, and family history of CVD based on model 2. Model 4 further adjusted for both time since menopause and metabolomic signature to assess their independent associations. * One unit representing a 5-year interval.

Study population



67,582 postmenopausal women with metabolites at baseline

4262 postmenopausal women with metabolites at repeat assessment

Exposure

Time since menopause (TSM)

Median follow-up 13.7 years

Outcome

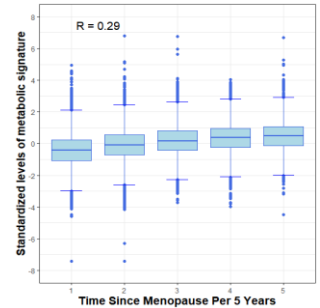
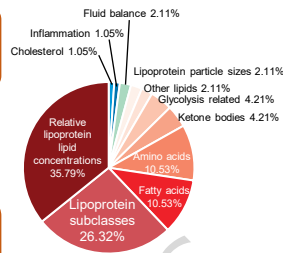
Incident CVD (MI, HF, AF, IHD, stroke)

To construct and validate a metabolic signature related to TSM

NMR-based quantification of 251 metabolites

Elastic Net regression

TSM-associated metabolic signature (95 metabolites)



To analyze the association of TSM and metabolomic signature with CVD

Cox models—TSM and CVD

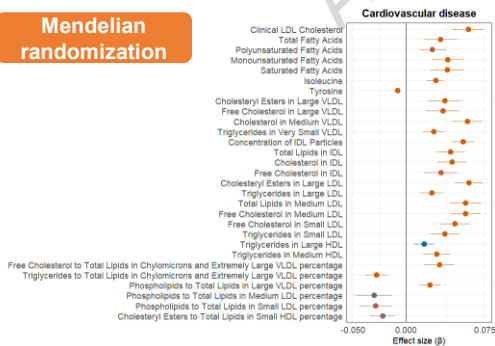
TSM per 1 increment	HR (95%CI)
CVD	1.14 (1.11, 1.16)
MI	1.18 (1.12, 1.25)
AF	1.15 (1.11, 1.19)
HF	1.18 (1.12, 1.24)
IHD	1.16 (1.12, 1.21)
Stroke	1.12 (1.06, 1.19)

Cox models—metabolomic signature and CVD

Metabolomic signature per 1-SD increment	HR (95%CI)
CVD	1.18 (1.15, 1.21)
MI	1.21 (1.15, 1.28)
AF	1.15 (1.11, 1.20)
HF	1.30 (1.23, 1.36)
IHD	1.21 (1.17, 1.26)
Stroke	1.18 (1.11, 1.25)

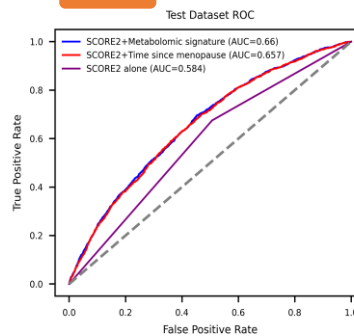
To identify the causal associations between metabolites and CVD

Mendelian randomization

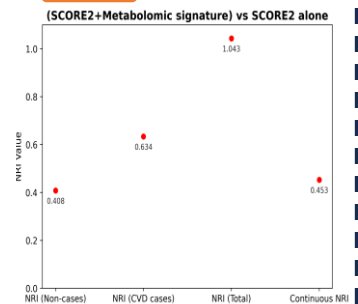


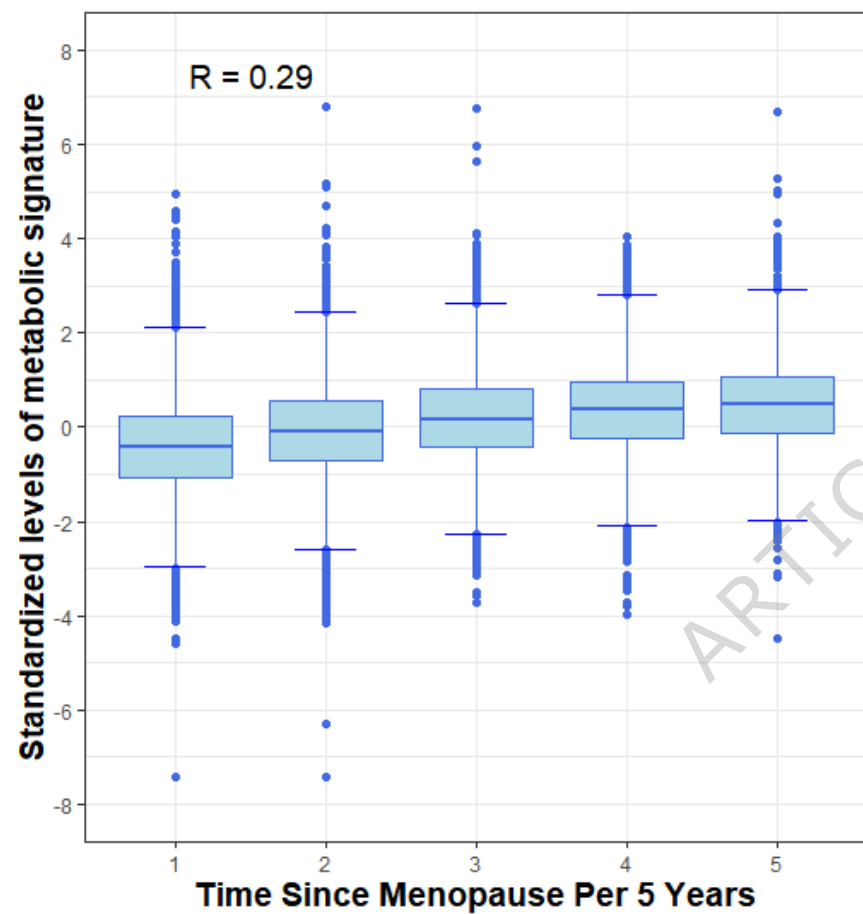
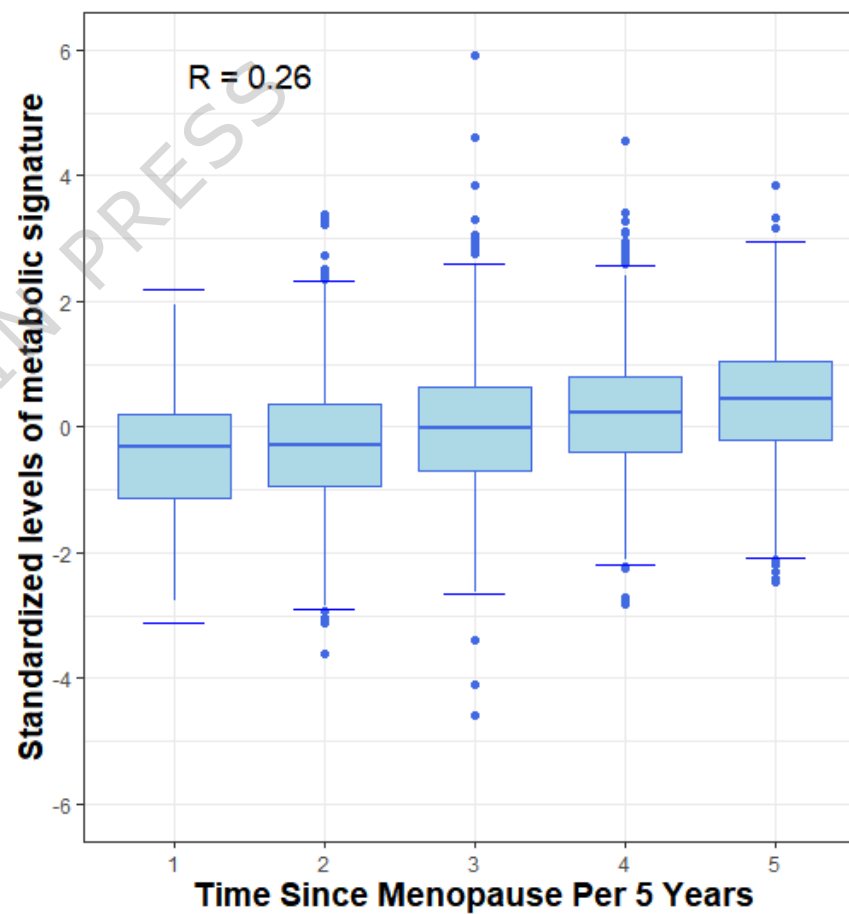
To assess improvement in CVD risk prediction after adding TSM and metabolomic signature to SCORE2 model

AUC



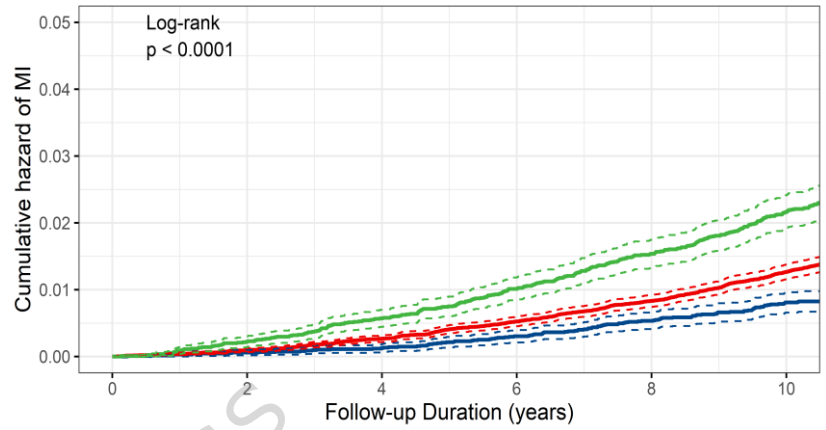
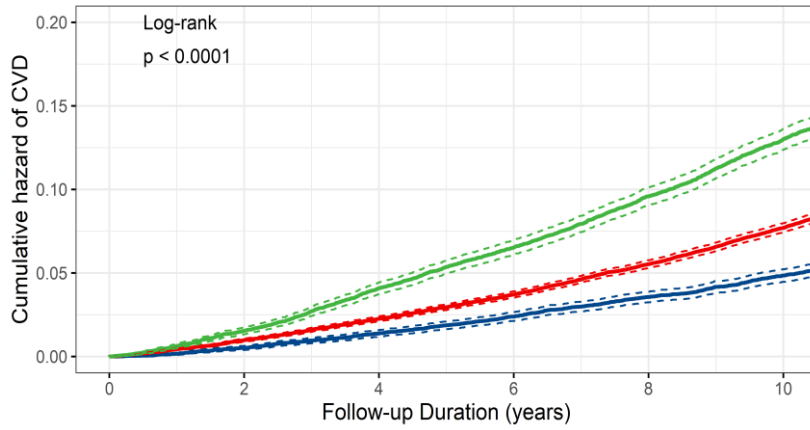
NRI



(a) baseline data**(b) repeat assessment**

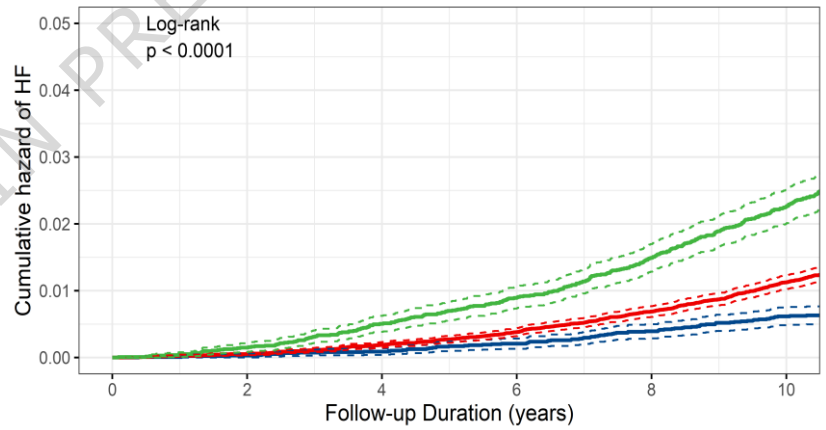
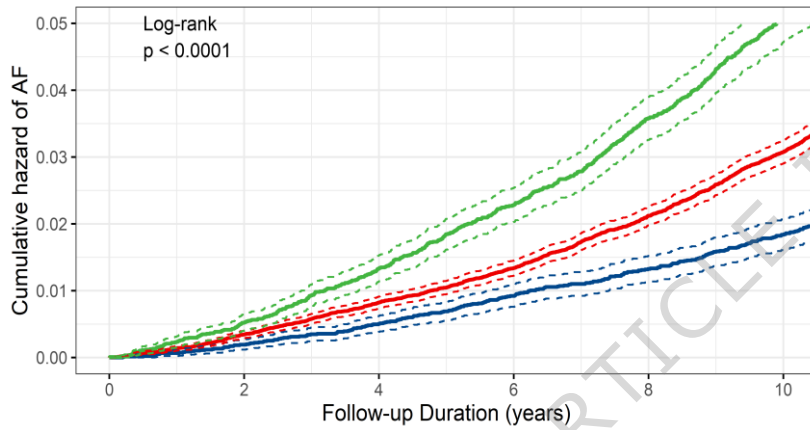
HR(95%CI) of CVD according metabolic signature categories — Low — Intermediate — High

HR(95%CI) of MI according metabolic signature categories — Low — Intermediate — High



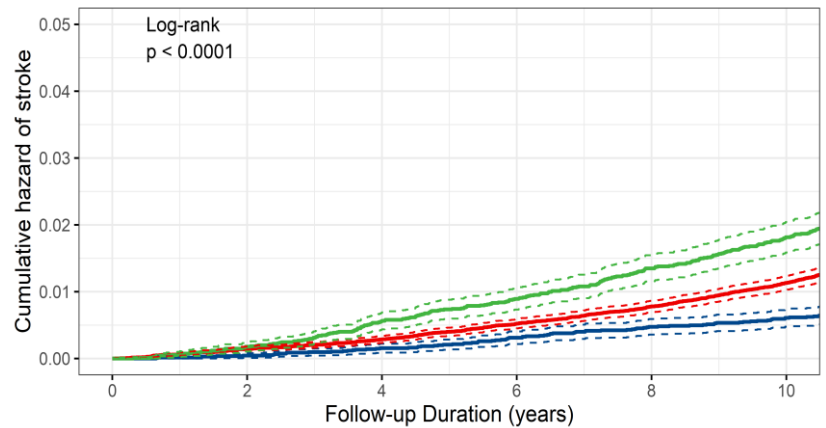
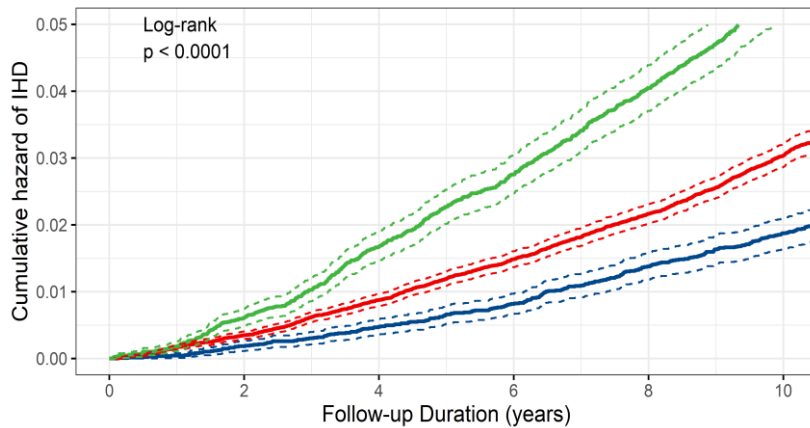
HR(95%CI) of AF according metabolic signature categories — Low — Intermediate — High

HR(95%CI) of HF according metabolic signature categories — Low — Intermediate — High

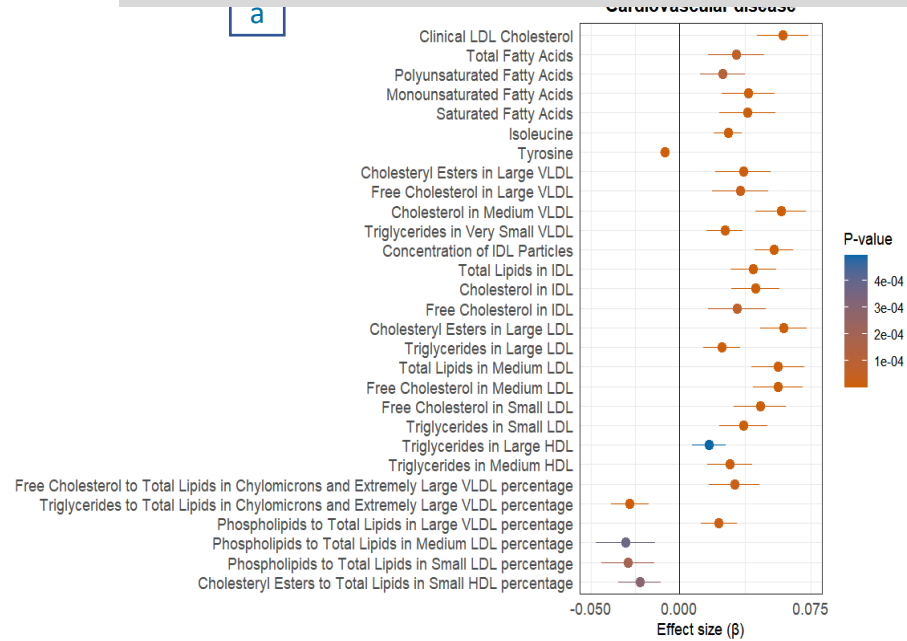


HR(95%CI) of IHD according metabolic signature categories — Low — Intermediate — High

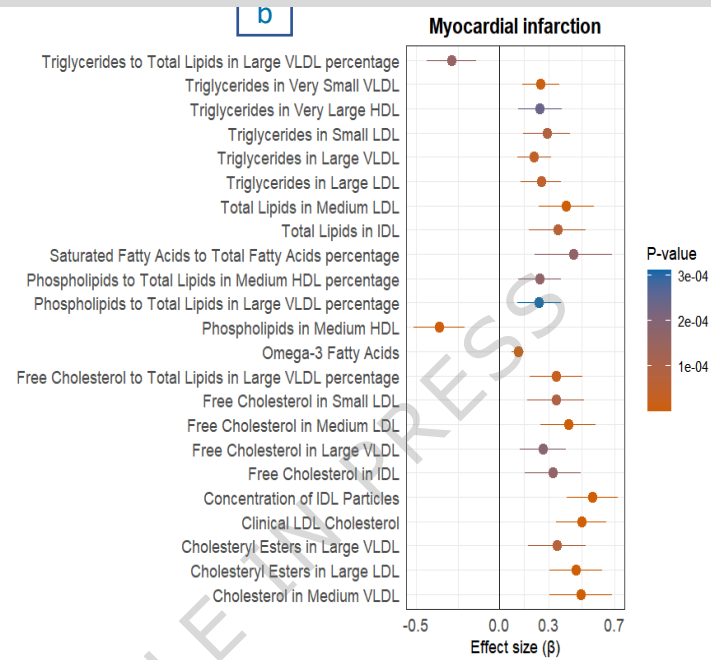
HR(95%CI) of stroke according metabolic signature categories — Low — Intermediate — High



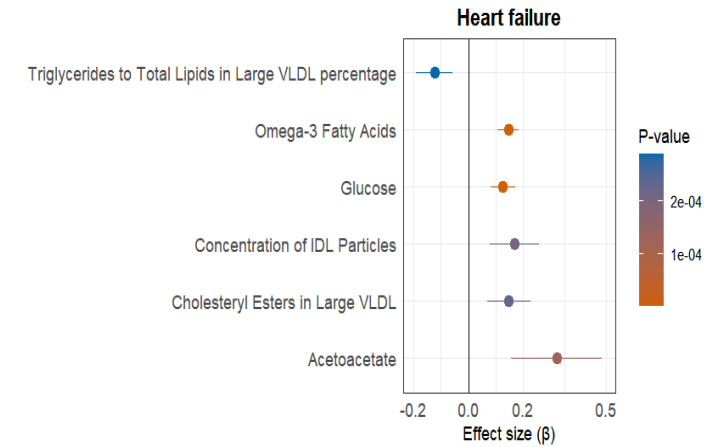
a



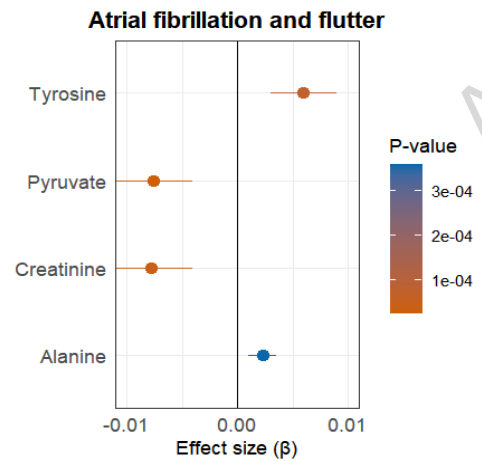
b



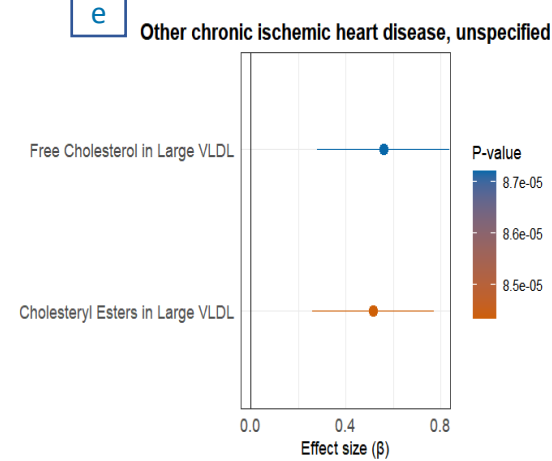
c



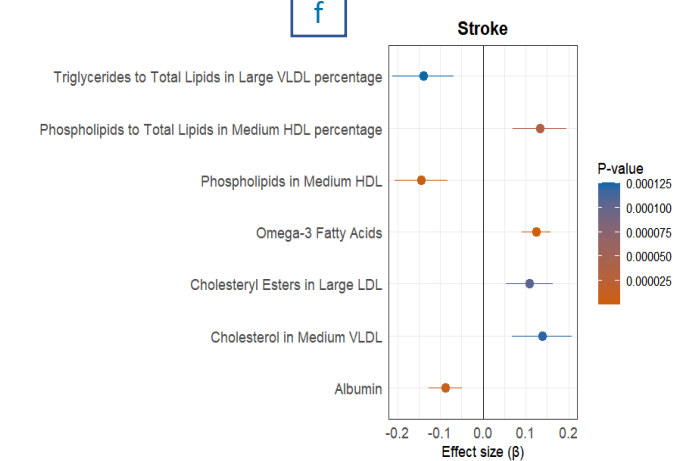
d



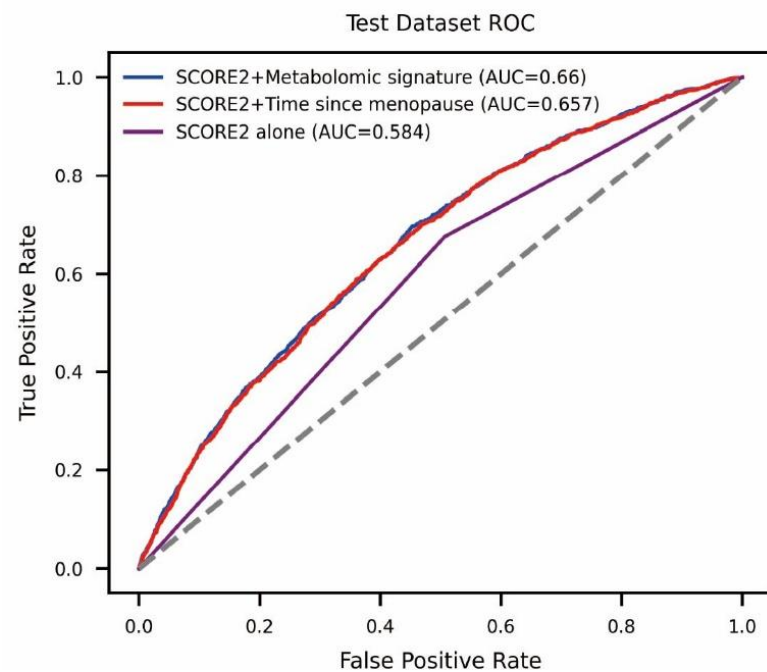
e



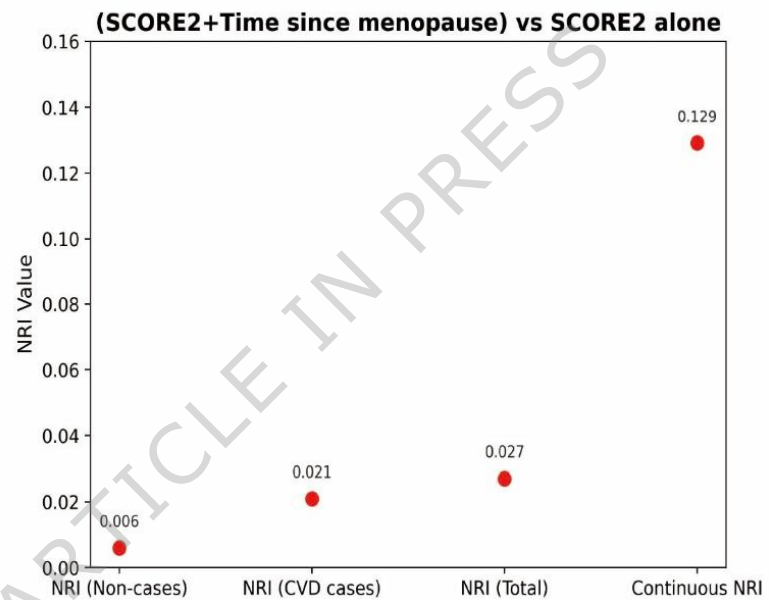
f



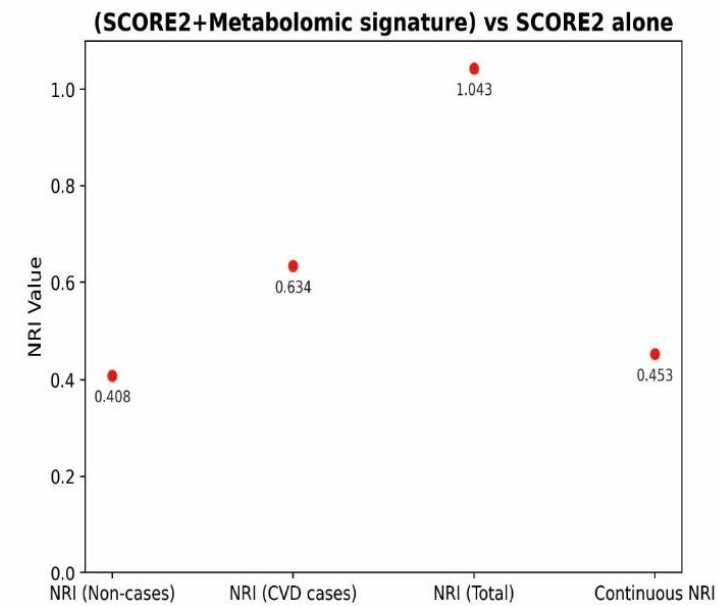
(A)



(B)



(C)



(n=273,255)

Missing metabolomics data $\geq 25\%$
(n=116,751)

Primary analysis (Metabolites at baseline)
(n=148,041)

Excluded

Women who were premenopausal,
experienced non-natural menopause,
or had abnormal menopause age
(<28 or >65 years) (n=70,984);
With CVD at baseline (n=6333);
Missing covariates (n=3142).

Participants with available metabolomics data and complete
information
(n=67,582)

Training set
(70%, n=47,307)

Elastic net regressions
10-fold cross-validation

Model applied interesting set to
calculate metabolite signature

Testing set
(30%, n=20,275)

Elastic net regressions Leave-one-out approach to calculate scores

67,582 participants included in the association of baseline metabolite
signature with CVD

Replication analysis (Metabolites at first repeat assessment)
(n=8463)

Excluded

Women who were premenopausal,
experienced non-natural menopause,
or had abnormal menopause age
(<28 or >65 years) (n=3791);
With CVD at baseline (n=314);
Missing covariates (n=96).

Participants with available metabolomics data and complete
information
(n=4262)

Elastic net coefficients
applied in repeat dataset
to calculate metabolite
signature score

4262 participants included in the association of metabolite
signature at first repeat with CVD