medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Age at Menarche and Coronary Artery Disease Risk: Divergent Associations with Different Sources of Variation

Ambreen Sonawalla, MD,[a,b] Daniel I. Chasman, PhD,[b,c] Yee-Ming Chan, MD, PhD[a,b].
[a]Division of Endocrinology, Department of Pediatrics, Boston Children's Hospital, [b]Departments of Pediatrics (A.S., Y.M.C.) and Medicine (D.I.C.), Harvard Medical School, [c]Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital

Short title: Age at Menarche and Coronary Artery Disease

Address for correspondence: Ambreen Sonawalla, MD; ambreen.sonawalla@childrens.harvard.edu; 300 Longwood Avenue, Boston, MA 02115. Phone: 617-355-0796, fax: 617-730-0194.

Total word count: 7097

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Abstract

Background: Both earlier and later age at menarche (AAM) are associated with increased risk of coronary artery disease (CAD) in women. This study sought to determine if the relationship of AAM with CAD and CAD risk factors differs for underlying sources of variation in AAM – specifically, variation attributable to common genetic variants as represented by a polygenic score (PGS) vs. variation independent of the PGS.

Methods: Primary analyses were conducted on data from 201,037 women in the UK Biobank and validation studies on data from 23,268 women in the Women's Genome Health Study (WGHS). For each individual, a PGS for AAM was calculated, then two variables were estimated from linear regression models: the PGS-associated change in AAM and the PGS-independent change in AAM. Logistic regression and linear splines were then used to study the relationships of these variables with CAD and CAD risk factors: hemoglobin A1c, triglycerides, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), systolic and diastolic blood pressure, body mass index, and waist-hip ratio.

Results: PGS-associated change in AAM demonstrated a linear relationship with CAD and linear or roughly linear relationships with CAD risk factors. In contrast, PGS-independent change in AAM demonstrated a U-shaped relationship with CAD and with hemoglobin A1c, triglycerides, HDL-C, and waist-hip ratio. Validation studies using WGHS data produced similar results.

Conclusions: These results suggest that later AAM itself does not cause increased risk of CAD; rather, upstream sources of variation other than common genetic variants can cause both later AAM and increased risk of CAD. Dysglycemia, dyslipidemia, and central adiposity are candidate mediators of the association of later AAM with increased risk of CAD.

Key words: pubertal timing, polygenic score, women, UK Biobank.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Non-standard Abbreviations and Acronyms:

AAM: age at menarche

BMI: body-mass index

CAD: coronary artery disease

DBP: diastolic blood pressure

HbA1c: hemoglobin A1c

HDL-C: high-density lipoprotein cholesterol

LDL-C: low-density lipoprotein cholesterol

PGS: polygenic score

SBP: systolic blood pressure

WGHS: Women's Genome Health Study

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

INTRODUCTION

Many pathologies manifesting in adulthood have antecedents in childhood. There is growing evidence that coronary artery disease (CAD) in women, a leading cause of morbidity and mortality in the world, is associated with both earlier and later age at menarche (AAM), a hallmark of pubertal timing (1–4). *Earlier* puberty is associated with increased risk of CAD in both men and women; however, the association between *later* puberty and increased risk of CAD appears to be unique to women; in men, later puberty is associated with a decreased risk of CAD (3). A deeper exploration of these childhood antecedents would allow a better understanding of the pathogenesis of CAD in adulthood, specifically identify factors that uniquely affect women and facilitate the development of targeted interventions, potentially as early as childhood.

Multiple studies have associated earlier AAM with a higher risk of developing components of the metabolic syndrome, namely obesity, type 2 diabetes mellitus, hypertension, and dyslipidemia (3,5–9). Studies have further suggested that the association between earlier AAM and risk of CAD is mediated by adiposity (10,11). In contrast, the associations of later AAM reported to date do not fit neatly into the paradigm of metabolic syndrome. Later AAM is associated with lower rather than higher body-mass index (BMI) (3,12,13), and studies on other components of the metabolic syndrome have produced conflicting results, with some studies showing association of later AAM with higher risk of hypertension (1,14), others showing a lower risk of hypertension (3) or type 2 diabetes (15), and yet others showing no association with hypertension, type 2 diabetes, or hypercholesterolemia (3,16). Hence, while earlier AAM has been associated with several CAD risk factors, there may be distinct mechanisms underlying the association of later AAM with increased risk of CAD.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Variation in AAM can stem from several upstream sources, including genetics (both common genetic variants and rare genetic variants), acquired factors such as chronic illness, chronic stress, underweight, and undernutrition, and environmental factors such as family composition (17). It is possible that some of these sources of variation may influence risk of CAD only through their influence on AAM, while others may directly influence CAD risk and features of the metabolic syndrome (Figure 1). Thus, dissecting variation in AAM based on underlying sources of variation could provide a clearer understanding of the relationship of AAM with CAD.

Genetics is a major source of variation in AAM, with half to three-quarters of variation attributable to genetics (18,19). The largest published genome-wide association study (GWAS) on AAM identified 389 independent single-nucleotide polymorphisms (SNPs) associated with AAM at genome-wide significance. The results from this GWAS allow the calculation of a polygenic score (PGS) for a given individual to reflect the cumulative contribution of common genetic variants to AAM.

Previous studies have used PGSs to dissect the influence of genetics vs. environmental factors (or other factors not captured by the PGS) that contribute to traits such as BMI and LDL-C (20,21). These studies have found that associations of these traits with health outcomes differ between genetically and environmentally influenced traits. For example, obesity driven by environmental factors was associated with more harmful cardiovascular outcomes than

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

genetically predicted obesity (20), suggesting that dissecting effects based on underlying source of variation can allow deeper insights into pathogenic mechanisms.

The aims of this study were two-fold: first, to determine if the association between later AAM and increased risk of CAD depends on the underlying source of variation in AAM – specifically, common genetic variation vs. other sources of variation; second, to study the relationships of these different sources of variation in AAM with CAD risk factors.

METHODS

To study how different sources of variation in the timing of menarche affect risk of CAD, variation in AAM was subdivided into variation attributable to the effects of common genetic influences, as estimated by a PGS for AAM, and variation due to other sources, i.e., independent of the PGS.

Study cohorts

This study used data from two cohorts: the UK Biobank for primary analyses and the Women's Genome Health Study (WGHS) for validation analyses. The UK Biobank is a population-based cohort of over 500,000 men and women in the UK 40 years and older at the time of recruitment, with extensive health-related phenotypic and laboratory data as well as individual-level genetic data (22). This study analyzed data from 201,037 unrelated women in the UK Biobank of non-Finnish European ancestry (as determined through principal component analysis) (23) who had

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

genetic and self-reported AAM data. During data collection in the UK Biobank, any value of AAM <5 years or >25 years was rejected, and any AAM entered as <6 years or >20 years required confirmation from participants (24). Women with missing AAM data were excluded from analysis. The UK Biobank obtained the multiple ethical and regulatory approvals required for recruitment and research procedures, and participants provided written consent (22). The WGHS is a cohort of initially healthy American women aged 45 years and older at enrollment with genetic and phenotypic data, followed over 26-28 years for cardiovascular and other outcomes (25). This analysis studied data from 23,268 women in the WGHS who had genetic and self-reported AAM data. Self-reported AAM ≤9 years or ≥ 17 years were entered as 9 years or 17 years, respectively. The WGHS was approved by the institutional review board of Brigham and Women's Hospital, and participants consented to ongoing analyses (25).

Polygenic score calculation

A PGS was calculated for each woman in the above cohorts in two steps. The first step used the PRS-CS algorithm, which allows the inclusion of all available SNPs from the GWAS on AAM (26), not just those that meet a given *p*-value threshold, and weights the SNP effect sizes based on their significance and adjusts for linkage disequilibrium (27). The use of this algorithm has the potential to explain more variability in AAM than algorithms that use only the SNPs that meet a given significance threshold. The second step calculated the PGS using PRSice-2 (without clumping or thresholding) to sum the weighted effect sizes for all SNPs in each individual with the ability to incorporate the probabilistic genotype dosages generated by imputation (28).

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Subdividing variation in age at menarche

Using each full cohort, regression of self-reported AAM was performed against the PGS for AAM, with the first 10 genetic principal components, assessment center and technical variables such as array number as covariates. The following two variables were then calculated for each individual:

1. PGS-associated change in AAM (PGS-associated ΔAAM): This represents the change in AAM attributable to the contributions of common genetic variants as estimated by the PGS. This was calculated as the difference between the individual's AAM predicted by the regression and the AAM corresponding to the mean PGS for the cohort (Supplemental Figure 1).
2. PGS-independent change in AAM (PGS-independent ΔAAM): This represents the change in AAM attributable to sources of variation not represented by the PGS. This was the residual for the individual of the regression of AAM against the PGS (Supplemental Figure 1).

Outcomes and analytical methods

The study's primary outcome variable was CAD risk. In the UK Biobank, prevalent CAD at baseline was determined as previously described using a combination of self-report, ICD-9/10 codes, and procedure codes (23). The WGHS recruited women with no history of CAD at baseline and identified validated incident CAD during 26-28 years of follow-up as described previously (25). Secondary outcome variables were CAD risk factors at baseline for both

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

cohorts: hemoglobin A1c (HbA1c), triglycerides, high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), systolic blood pressure (SBP), diastolic blood pressure (DBP), and BMI; data on waist-hip ratio was available at baseline in the UK Biobank and 6 years after recruitment in the WGHS.

The relationships of PGS-associated ΔAAM and PGS-independent ΔAAM with CAD prevalence were studied using logistic regression, with incident CAD using Cox proportional hazards models, and with each continuous variable (HbA1c, triglycerides, HDL-C, LDL-C, SBP, DBP, BMI, waist-hip ratio) using linear splines, with a knot at ΔAAM of 0 years. For analyses with LDL-C, results were corrected for self-reported use of cholesterol-lowering medications – these were specific LDL-lowering medications in the UK Biobank, (statins, ezetimibe, and bile-acid sequestrants) and collective cholesterol-lowering medications in the WGHS; additional analyses included only women not taking these medications.

Covariates of age and age^2 were used for all analyses. Analyses were conducted using R v.4.3.1. A significance threshold of 0.05 was used.

To determine if results in the UK Biobank were biased by overfitting as a result of the UK Biobank having contributed to the GWAS for AAM, we also repeated analyses of CAD using a PGS calculated using an earlier GWAS that did not include the UK Biobank (29). Furthermore, to determine if outlier values of AAM were disproportionately affecting results, analyses were repeated after excluding women with extreme values of AAM such that up to 0.1% of women were excluded at each extreme.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

RESULTS

To study the relationship between AAM and CAD risk, our study divided variation in AAM into 1) variation attributable to the effects of common genetic variants as estimated by a PGS for AAM, which we term PGS-associated ΔAAM, and 2) variation due to factors not captured by the PGS, which we term PGS-independent ΔAAM.

To examine the differential effects of these two sources of variation on CAD risk, we analyzed data from 201,037 unrelated, non-Finnish European women in the UK Biobank. We first calculated a PGS for AAM for each woman, then regressed self-reported AAM against the AAM PGS to calculate PGS-associated ΔAAM and PGS-independent ΔAAM (Supplemental Figure 1). For women in the UK Biobank, the regression of self-reported AAM against the AAM PGS demonstrated that the PGS accounted for 15.8% of the variation in AAM. PGS-associated ΔAAM had a mean of 0.016 years, with a standard deviation of 0.639 years, and the residual, PGS-independent ΔAAM, had a mean of 0 years with a standard deviation of 1.47 years (Supplemental Figure 2). We then studied the associations of PGS-associated ΔAAM and PGS-independent ΔAAM with risk of CAD and with CAD risk factors.

Risk of coronary artery disease

Risk of CAD demonstrated a linear relationship with PGS-associated ΔAAM but a non-linear, U-shaped relationship with PGS-independent ΔAAM (Figure 2). In the linear relationship of

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

PGS-associated ΔAAM with risk of CAD; each 1-year increase in PGS-associated ΔAAM was associated with an odds ratio (OR) for CAD of 0.91. In other words, for every 1 year that AAM was later due to the effects of common genetic variants, as estimated by the PGS, the odds of CAD were lower by 9% (Table 1). Linear spline analyses showed no difference in slopes when PGS-associated ΔAAM was negative vs. positive, i.e., driving earlier vs. later AAM (Table 1).

In contrast, the association between PGS-independent ΔAAM and risk of CAD was U-shaped, with both increasingly negative and increasingly positive values (driving earlier and later AAM, respectively) associated with increased CAD risk (Figure 2). For negative values of PGS-independent ΔAAM, each 1-year increase (causing AAM to be less early) was associated with an OR for CAD of 0.91 (Table 1). In other words, for every 1 year that AAM was less early due to factors independent of the PGS, the odds of CAD were lower by 9%. In contrast, for positive values of PGS-independent ΔAAM, each 1-year increase (causing AAM to be even later) was associated with an OR for CAD of 1.11, i.e., the odds of CAD were higher by 11% (Table 1).

Because the 2018 AAM GWAS meta-analysis included data from the UK Biobank, which could introduce bias in the above analyses, we conducted sensitivity analyses using the results from an AAM GWAS meta-analysis published in 2014 that did not include the UK Biobank. These analyses showed similar results for CAD (Supplemental Figure 3, Supplemental Table 1). We also conducted sensitivity analyses excluding women with extreme values of AAM and obtained similar results.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Hemoglobin A1c

We then examined the associations of PGS-associated ΔAAM and PGS-independent ΔAAM with CAD risk factors. Just as the relationship between PGS-associated ΔAAM and CAD was linear, we observed linear (or roughly linear) relationships between PGS-independent ΔAAM and most CAD risk factors. In contrast, for PGS-independent ΔAAM we observed mostly non-linear relationships with CAD risk factors.

For hemoglobin A1c, which correlates with average blood glucose over the preceding 3 months, PGS-associated ΔAAM demonstrated a roughly negative linear relationship, with higher PGS-associated ΔAAM associated with lower HbA1c (Figure 3). The slope of the association with HbA1c for negative values of PGS-associated ΔAAM was slightly steeper than for positive values ((-0.034 and -0.022, respectively; *p* for difference between slopes: 7×10^{-5}; Table 1). In contrast, PGS-independent ΔAAM demonstrated a U-shaped relationship with HbA1c: both increasingly negative and increasingly positive values of PGS-independent ΔAAM were associated with an increase in HbA1c (for negative values: slope = -0.022, for positive values: slope = 0.007; *p* for difference between slopes $< 2 \times 10^{-16}$; Figure 3, Table 1). This was similar to the U-shaped relationship of PGS-independent ΔAAM with CAD risk.

Comparing PGS-associated ΔAAM and PGS-independent ΔAAM, for negative values (driving earlier AAM) the slopes for the associations with HbA1c were comparable. However, for positive values (driving later AAM), different slopes were seen for PGS-associated vs. PGS-

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

independent ΔAAM; as noted above, for every 1 year that AAM was delayed, HbA1c *decreased* by 0.01% for PGS-associated ΔAAM but *increased* by 0.007% for PGS-independent ΔAAM.

Lipids

PGS-associated ΔAAM demonstrated linear relationships with triglycerides and HDL-C (Figure 3), with higher values of PGS-associated ΔAAM associated with a decrease in triglycerides and an increase in HDL-C (Figure 3) and no significant difference in slopes between negative and positive values of PGS-associated ΔAAM (*p* for difference in slopes for triglycerides = 0.83, for HDL-C = 0.38; Table 1). In contrast, PGS-independent ΔAAM demonstrated non-linear relationships with HDL-C and triglycerides – U-shaped for triglycerides and inverted-U-shaped for HDL-C (Figure 3), with both increasingly negative and increasingly positive values of PGS-independent ΔAAM associated with increases in triglycerides and decreases in HDL-C (Figure 3, Table 1).

For LDL-C, PGS-associated ΔAAM showed no association for negative values of PGS-associated ΔAAM and a negative linear relationship for positive values of PGS-associated ΔAAM (Table 1). For PGS-independent ΔAAM, there was an association for negative values – increasing PGS-independent ΔAAM was associated with lower LDL-C – and no significant association with LDL-C for positive values of PGS-independent ΔAAM (Figure 3, Table 1). Similar results were seen when analyses excluded those taking LDL-lowering medications (Table 1).

Blood pressure

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

For PGS-associated ΔAAM, increasing PGS-associated ΔAAM was associated with decreases in both SBP and DBP (Figure 3), with no significant difference between the slopes for negative vs. positive values of PGS-associated ΔAAM (Table 1). For PGS-independent ΔAAM, increasing PGS-independent ΔAAM was also associated with decreased SBP and DBP (Figure 3), but the slopes of the associations were steeper for negative values of PGS-independent ΔAAM compared to the slopes for positive values of PGS-independent ΔAAM (Table 1).

Adiposity

We studied two estimates of adiposity: BMI and waist-hip ratio. For BMI, we found a roughly linear negative relationship between PGS-associated ΔAAM and BMI, with higher PGS-associated ΔAAM associated with lower BMI (Figure 3, Table 1). The slope of the association with BMI was slightly steeper for negative values of PGS-associated ΔAAM than for positive values (Table 1). This pattern was similar to the relationship seen between PGS-associated ΔAAM and HbA1c.

For PGS-independent ΔAAM, a reverse-J-shaped relationship was seen with BMI (Figure 3). While higher values of PGS-independent ΔAAM were consistently associated with lower BMI, the slope of the association was ten-fold steeper for negative than for positive values of PGS-independent ΔAAM (-0.80 vs. -0.079 respectively; *p* for difference between slopes 2×10^{-16}; Table 1). This resembled the patterns seen with SBP and DBP.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

For waist-hip ratio, PGS-associated ΔAAM showed a negative linear association, with no significant difference in slopes for the association when PGS-associated ΔAAM was negative vs. positive (Figure 3, Table 1). In contrast, the relationship of PGS-independent ΔAAM with waist-hip ratio was U-shaped, with both increasingly negative and increasingly positive values of PGS-associated ΔAAM associated with an increase in waist-hip ratio (Figure 3, Table 1), similar to the associations seen with CAD, HbA1c, triglycerides, and HDL-C.

Validation in the Women's Genome Health Study

Validation studies using data from the WGHS produced results similar to the above results for the UK Biobank. In the WGHS, PGS-associated ΔAAM showed linear relationships with HbA1c, triglycerides, HDL-C, SBP, DBP, BMI and waist-hip ratio and no clear relationship with LDL-C (Figure 4). Just as in the UK Biobank, PGS-independent ΔAAM showed U-shaped relationships with triglycerides and HbA1c, an inverted-U-shaped relationship with HDL-C, and a reverse-J-shaped relationship with BMI in the WGHS (Figure 4).

There were two differences between results from the WGHS and the UK Biobank: 1) for SBP and DBP, PGS-independent ΔAAM showed roughly U-shaped relationships in the WGHS (Figure 4) rather than the reverse-J-shaped relationships seen in the UK Biobank (Figure 3), and 2) for waist-hip ratio, PGS-independent ΔAAM showed a negative linear relationship for negative values of PGS-independent ΔAAM but no significant relationship for positive values in the WGHS (Figure 4) compared to the clear U-shaped relationship seen in the UK Biobank (Figure 2).

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

For risk of CAD itself, the relationship of PGS-associated ΔAAM with CAD in the WGHS was similar to that in the UK Biobank (Figure 5, Supplemental Table 1), with a negative linear relationship in both cohorts. For the relationship of PGS-independent ΔAAM with CAD, analyses in the WGHS demonstrated a reverse-J shaped relationship (Figure 5, Supplemental Table 1), slightly different from the U-shaped relationship seen in the UK Biobank.

DISCUSSION

In this study, we dissected variation in AAM into variation attributable to common genetic variants, as estimated by a PGS, and variation independent of the PGS, and we found different relationships with these two sources of variation in AAM with risk of CAD and with CAD risk factors, particularly when causing AAM to occur later.

In general, later AAM showed favorable associations when attributable to PGS-associated variation and harmful or neutral associations when attributable to PGS-independent variation. If all sources of variation were affecting CAD and CAD risk factors wholly through AAM itself, the associations would be similar regardless of the source of variation studied. Our finding that these associations varied based on the underlying source of variation driving later AAM therefore indicates that it is not later AAM itself that causes increased risk of CAD. Rather, there appear to be PGS-independent factors that cause both later AAM and increased risk of CAD and unfavorable cardiometabolic risk profiles. Such factors could include environmental or acquired influences (such as chronic illnesses, chronic stress, undernutrition) as well as genetic influences

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

not captured by the PGS, and future studies will identify these factors and determine how they contribute to CAD risk.

Earlier AAM, whether attributable to PGS-associated or PGS-independent sources of variation, was consistently associated with greater cardiometabolic risk. Thus, it seems that earlier AAM itself is intrinsically associated with risk for CAD and worsening CAD risk factors; indeed, some Mendelian randomization studies suggest that earlier AAM is causative of these negative outcomes (10,11,30). However, other Mendelian randomization studies have suggested that the association of earlier AAM and increased risk of CAD is due to pleiotropic effects of genetic factors influencing both AAM and BMI, such that AAM itself may have only a small influence (31).

The U-shaped relationship of PGS-independent ΔAAM with CAD was mirrored by the relationships with HbA1c, triglycerides, HDL-C, and waist-hip ratio. This raises the possibility that dysglycemia, dyslipidemia, and central adiposity contribute to the relationship between later AAM and increased risk of CAD; further studies will be required to formally evaluate these CAD risk factors as potential mediators of this relationship. As noted above, previous studies evaluating associations of later AAM with these CAD risk factors have had conflicting results. While differences between study cohorts may have accounted for some of the differing results, our findings raise the additional possibility that the relationships between later AAM and these outcomes may have been obscured by opposing effects of PGS-associated and PGS-independent variation in AAM. Of note, the associations with waist-hip ratio differed from those with BMI,

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

suggesting that central obesity, reflected by waist-hip ratio, is more relevant for CAD, as has been suggested by prior studies (32,33).

Our analyses in the Women's Genome Health Study (WGHS) largely validated our results from the UK Biobank. PGS-associated change in AAM demonstrated linear relationships with CAD and CAD risk factors, and PGS-independent change in AAM demonstrated mostly non-linear relationships with these outcomes, supporting the conclusions described earlier.

However, the two cohorts also demonstrated some differences, most notably in the association of PGS-independent ΔAAM with the risk of CAD and waist-hip ratio In the UK Biobank, PGS-independent variation driving AAM *later* was associated with an increased risk of CAD, but there was no association in the WGHS. For waist-hip ratio, PGS-independent variation demonstrated a clear U-shaped association in the UK Biobank but a reverse J-shaped relationship in the WGHS. There are several potential reasons for this difference. First, it is possible that the smaller sample size and lower power in the WGHS affected the ability to find significant associations with PGS-independent variation in AAM and risk of CAD and waist-hip ratio. Second, environmental influences affecting AAM (which would contribute to PGS-independent variation in AAM) could differ between the two cohorts. The WGHS recruited women in the United States of America born in 1950 or earlier, while the UK Biobank recruited women in the UK who were born between 1932 and 1969. The different impact of global events such as World War II on the two countries could contribute to differences in the PGS-independent factors (which includes environmental factors) and, in turn, to different associations with risk of CAD.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Third, the WGHS excluded women with a history of CAD at the time of enrollment whereas the UK Biobank did not, and this may also have led to differences in PGS-independent factors between the cohorts. Fourth, the participants in the WGHS were health professionals while the UK Biobank drew from the general UK population, and this may have led to further differences in PGS-independent factors, such as higher socioeconomic status, greater knowledge of CAD and its risk factors, healthier diets and lifestyles, and use of preventative interventions, as well as potentially less variation in these factors.. While the analyses also differed in the CAD measure used from each cohort – prevalence of CAD was analyzed in the UK Biobank compared to incidence of CAD in the WGHS – it is unlikely to account for the difference in results, as an analysis of incident CAD in the Million Women Study in the UK also showed a U-shaped relationship between AAM and CAD (1). Despite these differences in the results between the two cohorts, results from both cohorts consistently demonstrated differences between the associations of PGS-independent variation driving AAM *earlier* (negative values) vs. driving AAM *later* (positive values) with these outcomes.

In 2021, Liang et al. used a PGS to represent genetically predicted AAM and examined associations with all-cause mortality, also using data from the UK Biobank. Interestingly, they found a U-shaped association with mortality, with both earlier and later genetically predicted AAM associated with higher risk; this finding contrasts with the linear relationship we found between PGS-associated ΔAAM and CAD. This difference suggests that later genetically predicted AAM increases the risk of causes of mortality other than CAD, and future studies of genetically predicted AAM are needed to identify these causes.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Prior studies that have also dissected influences on human traits into genetic vs. environmental influences have found differences in the magnitude of effect of genetics vs. environmental influences on health outcomes (20,21). Interestingly, our results show not only different magnitudes of influence of PGS-associated and PGS-independent variation in AAM on risk of CAD, but also opposite directionality. This further underscores the value of separating the effects of genetics vs. other influences while studying human traits as they can have starkly different effects.

Sensitivity analyses using a GWAS that did not include the primary cohort, the UK Biobank, produced similar results and suggested that these results were not biased by overfitting. Additionally, sensitivity analyses excluding women with extreme values of AAM also produced similar results, suggesting that results were not heavily influenced by outliers.

One limitation of this study is that while the PGS represents genetic influences on AAM, it does not represent all genetic factors that influence AAM. Our analyses demonstrated that the PGS explains 15.8% of the variation in the observed AAM, but prior studies suggest that 49-73% of variation in AAM is inherited (18,19); hence, a large amount of the variation in AAM due to genetic factors remains unexplained. Another limitation of this study is that the PGS represents just one method of capturing the effects of common genetic variants that affect AAM. There may be several pathways causing later AAM represented within these common genetic variants, and using a single PGS to represent all those effects may obscure relationships with each individual pathway. Future studies may identify these different pathways by methods such as clustering

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

analyses (34,35), which would then allow an estimation of multiple polygenic scores, each representing a different pathway, to study their associations with risk of CAD and CAD risk factors.

Distinguishing between sources of variation in AAM has provided a novel lens through which to study associations of AAM with CAD and has allowed us to uncover differences in the associations of PGS-associated vs. PGS-independent variation in AAM with CAD and CAD risk factors. Because later puberty in women, but not men, has been associated with an increased risk of CAD, studying these differences further may provide unique insights into mechanisms that affect CAD risk specifically in women.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

ACKNOWLEDGEMENTS

The authors thank Evan Schafer and Jia Zhu for their support with statistical analyses. This research has been conducted using data from UK Biobank, a major biomedical database: www.ukbiobank.ac.uk

SOURCES OF FUNDING

A.S. was supported by National Institutes of Health (NIH) grant T32DK007699. The WGHS and its parent cohort, the Women's Health Study (WHS), have been supported by the National Heart, Lung, and Blood Institute (HL043851 and HL080467) and the National Cancer Institute (CA047988 and UM1CA182913), with support and funding for genotyping provided by Amgen. Some cardiovascular endpoints in the WHS were funded by HL099355.

DISCLOSURES

A.S. has no disclosures or relationships with industry. D.I.C has no disclosures or relationships with industry. Y.-M.C. receives royalties from UpToDate on topics related to puberty.

SUPPLEMENTAL MATERIALS

Supplemental Figures 1-3

Supplemental Table 1

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

REFERENCES

1. Canoy D, Beral V, Balkwill A, Wright FL, Kroll ME, Reeves GK, et al. Age at menarche and risks of coronary heart and other vascular diseases in a large UK cohort. Circulation 2015;131(3):237–44.

2. Zhu J, Chan YM. Adult Consequences of Self-Limited Delayed Puberty. Pediatrics [Internet]. 2017 Jun 1;139(6).

3. Day FR, Elks CE, Murray A, Ong KK, Perry JRB. Puberty timing associated with diabetes, cardiovascular disease and also diverse health outcomes in men and women: the UK Biobank study. Sci Rep. 2015 Jun 18;5.

4. Lee JJ, Cook-Wiens G, Johnson BD, Braunstein GD, Berga SL, Stanczyk FZ, et al. Age at Menarche and Risk of Cardiovascular Disease Outcomes: Findings From the National Heart Lung and Blood Institute-Sponsored Women's Ischemia Syndrome Evaluation. J Am Heart Assoc. 2019 Jun 18;8(12).

5. Santos MP, Li Y, Bazzano LA, He J, Rexrode KM, Ley SH. Age at menarche, type 2 diabetes and cardiovascular disease complications in US women aged under 65 years: NHANES 1999-2018. BMJ Nutr Prev Health. 2023 Dec 1;6(2):293–300.

6. Janghorbani M, Mansourian M, Hosseini E. Systematic review and meta-analysis of age at menarche and risk of type 2 diabetes. Acta Diabetol. 2014;51(4):519–28.

7. Stöckl D, Meisinger C, Peters A, Thorand B, Huth C, Heier M, et al. Age at menarche and its association with the metabolic syndrome and its components: results from the KORA F4 study. PLoS One. 2011 Oct 18;6(10).

8. Liu D, Qin P, Liu Y, Sun X, Li H, Wu X, et al. Association of age at menarche with hypertension in rural Chinese women. J Hypertens. 2021 Mar 1;39(3):476–83.

9. Kheradmand M, Hamzehgardeshi Z, Shahhosseini Z, Mirjalili R, Moosazadeh M. The association between early menarche and higher-risk cardiometabolic profile: a dose-response analysis of the Tabari cohort at enrollment phase. Front Cardiovasc Med. 2023;10.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

10. Cao M, Cui B. Negative Effects of Age at Menarche on Risk of Cardiometabolic Diseases in Adulthood: A Mendelian Randomization Study. J Clin Endocrinol Metab. 2020 Jan 23;105(2).

11. Gill D, Brewer CF, Del Greco M F, Sivakumaran P, Bowden J, Sheehan NA, et al. Age at menarche and adult body mass index: a Mendelian randomization study. International Journal of Obesity 2018 42:9. 2018 Feb 26;42(9):1574–81.

12. Prentice P, Viner RM. Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. Int J Obes (Lond). 2013 Aug;37(8):1036–43.

13. Yang L, Li L, Millwood IY, Lewington S, Guo Y, Sherliker P, et al. Adiposity in relation to age at menarche and other reproductive factors among 300 000 Chinese women: findings from China Kadoorie Biobank study. Int J Epidemiol. 2017 Jan 1;46(2):502–12.

14. Chen L, Zhang L, Chen Z, Wang X, Zheng C, Kang Y, et al. Age at menarche and risk of hypertension in Chinese adult women: Results from a large representative nationwide population. J Clin Hypertens (Greenwich). 2021 Aug 1;23(8):1615–21.

15. Liu D, Zhao Y, Liu Y, Sun X, Li H, Yin Z, et al. Adiposity and insulin resistance as mediators between age at menarche and type 2 diabetes mellitus. Menopause. 2020 May 1;27(5):579–85.

16. Elks CE, Ong KK, Scott RA, Van Der Schouw YT, Brand JS, Wark PA, et al. Age at menarche and type 2 diabetes risk: the EPIC-InterAct study. Diabetes Care. 2013 Nov;36(11):3526–34.

17. Schlomer GL, Marceau K. Father absence, age at menarche, and genetic confounding: A replication and extension using a polygenic score. Dev Psychopathol. 2022 Feb 27;34(1):355–66.

18. Morris DH, Jones ME, Schoemaker MJ, Ashworth A, Swerdlow AJ. Familial concordance for age at menarche: Analyses from the breakthrough generations study. Paediatr Perinat Epidemiol. 2011 May;25(3):306–11.

19. Towne B, Czerwinski SA, Demerath EW, Blangero J, Roche AF, Siervogel RM. Heritability of age at menarche in girls from the Fels Longitudinal Study. Am J Phys Anthropol. 2005 Sep;128(1):210–9.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

20. Ojalehto E, Zhan Y, Jylhävä J, Reynolds CA, Dahl Aslan AK, Karlsson IK. Genetically and environmentally predicted obesity in relation to cardiovascular disease: a nationwide cohort study. EClinicalMedicine. 2023 Apr 1;58.

21. Hawkes G, Yengo L, Vedantam S, Marouli E, Beaumont RN, Tyrrell J, et al. Identification and analysis of individuals who deviate from their genetically-predicted phenotype. PLoS Genet. 2023 Sep 21;19(9).

22. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015 Mar 1;12(3).

23. Zhu J, Pujol-Gualdo N, Wittemans LBL, Lindgren CM, Laisk T, Hirschhorn JN, et al. Evidence From Men for Ovary-independent Effects of Genetic Risk Factors for Polycystic Ovary Syndrome. J Clin Endocrinol Metab. 2022 Mar 24;107(4):e1577–87.

24. UK Biobank age at menarche data collection: https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=2714.

25. Ridker PM, Chasman DI, Zee RYL, Parker A, Rose L, Cook NR, et al. Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. Clin Chem. 2008;54(2):249–55.

26. 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 Jun 1;49(6):834.

27. Ge T, Chen CY, Ni Y, Feng YCA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun. 2019 Dec 1;10(1).

28. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience. 2019 Jul 1;8(7).

29. Perry JRB, 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 Oct 10;514(7520):97.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

30. Zheng J, Chen K, Huang T, Shao C, Li P, Wang J, et al. Genetically Determined Lifestyle and Cardiometabolic Risk Factors Mediate the Association of Genetically Predicted Age at Menarche With Genetic Predisposition to Myocardial Infarction: A Two-Step, Two-Sample Mendelian Randomization Study. Front Cardiovasc Med. 2022 Apr 25;9.

31. Bell JA, Carslake D, Wade KH, Richmond RC, Langdon RJ, Vincent EE, et al. Influence of puberty timing on adiposity and cardiometabolic traits: A Mendelian randomisation study. PLoS Med. 2018 Aug 1;15(8).

32. Huxley R, Mendis S, Zheleznyakov E, Reddy S, Chan J. Body mass index, waist circumference and waist:hip ratio as predictors of cardiovascular risk--a review of the literature. Eur J Clin Nutr. 2010 Jan;64(1):16–22.

33. Sahakyan KR, Somers VK, Rodriguez-Escudero JP, Hodge DO, Carter RE, Sochor O, et al. Normal-Weight Central Obesity: Implications for Total and Cardiovascular Mortality. Ann Intern Med. 2015 Dec 1;163(11):827–35.

34. Chasman DI, Giulianini F, Demler O V., Udler MS. Pleiotropy-Based Decomposition of Genetic Risk Scores: Association and Interaction Analysis for Type 2 Diabetes and CAD. Am J Hum Genet. 2020 May 7;106(5):646–58.

35. Udler MS, Kim J, von Grotthuss M, Bonàs-Guarch S, Cole JB, Chiou J, et al. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis. PLoS Med. 2018 Sep 1;15(9).

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

FIGURE LEGENDS

Figure 1: Relationships of different sources of variation in age at menarche (AAM) with coronary artery disease (CAD) and CAD risk factors. Variation in AAM was dissected into variation associated with common genetic variants (as estimated by a polygenic score for AAM) and variation independent of the polygenic score, then associations of these sources of variation with CAD and CAD risk factors were studied. When associated with common genetic variants, later AAM was associated with a lower risk of CAD and favorable changes in CAD risk factors (blue arrows). However, when occurring independently of the polygenic score, later AAM was associated with higher risk of CAD and unfavorable changes in CAD risk factors (red arrows). The "other factors" causing later AAM independent of the PGS have yet to be identified and could potentially include chronic illness, underweight, and rare genetic variants.

Figure 2: Association of variation in AAM with odds of CAD in the UK Biobank. Variation in AAM was dissected into variation associated with common genetic variants as estimated by a PGS for AAM (PGS-associated ΔAAM) and variation independent of the PGS (PGS-independent ΔAAM), then associations of these sources of variation with risk of CAD were studied. PGS-associated ΔAAM showed linear or roughly linear associations with CAD whereas PGS-independent ΔAAM showed a U-shaped association with risk of CAD. The different associations of positive values of PGS-associated ΔAAM with risk of CAD demonstrates that it is not later AAM itself that causes the increased risk of CAD and that the increased risk of CAD observed in later AAM is being driven by factors other than common genetic variants and that. To achieve bin sizes of ≥1000 individuals, the first and last bins for each variable represent the

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

group of individuals with ΔAAM ≤ or ≥ the value listed on the x-axis, respectively. Dots represent estimates; bars represent standard errors.

AAM: age at menarche; ΔAAM: change in AAM; CAD: coronary artery disease; PGS: polygenic score.

Figure 3: Associations of variation in AAM with risk factors for CAD in the UK Biobank. We observed linear (or roughly linear) relationships between PGS-associated ΔAAM and most CAD risk factors. In contrast, for PGS-independent ΔAAM we observed mostly non-linear relationships with CAD risk factors. The U-shaped relationship of PGS-independent ΔAAM with CAD was mirrored by the relationships with HbA1c, triglycerides, HDL-C, and waist-hip ratio. This raises the possibility that dysglycemia, dyslipidemia and central adiposity may mediate the relationship between later AAM and increased risk of CAD. For LDL-C, n = 177,542 after excluding those taking LDL-C lowering medications. Dots represent estimates, bars represent standard errors.

AAM: age at menarche; ΔAAM: change in AAM; BMI: body mass index; CAD: coronary artery disease; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; PGS: polygenic score

Figure 4: Associations of variation in AAM with risk factors for CAD in the WGHS. PGS-associated ΔAAM showed linear or roughly linear associations with most CAD risk factors and PGS-independent ΔAAM showed mostly non-linear associations with most CAD risk factors. Similar to results in the UK Biobank, hemoglobin A1c, triglycerides and HDL-C showed U-shaped associations with PGS-independent ΔAAM which support conclusions from the primary

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

analyses that dysglycemia, dyslipidemia and central adiposity may mediate the relationship between later AAM and increased risk of CAD. For LDL-C, n = 22,495 after excluding those taking cholesterol-lowering medications. Dots represent estimates, bars represent standard errors.

AAM: age at menarche; ΔAAM: change in AAM; BMI: body mass index; CAD: coronary artery disease; HDL-C: high density lipoprotein cholesterol; LDL-C: low density lipoprotein cholesterol; PGS: polygenic score; WGHS: Women's Genome Health Study.

Figure 5: Associations of variation in AAM with hazard ratio for risk of CAD in the WGHS. PGS-associated ΔAAM demonstrated a negative linear association with risk of CAD, similar to the results in the UKBB. PGS-independent ΔAAM demonstrated a reverse-J shaped relationship, slightly different from the U-shaped relationship seen in the UK Biobank. Analyses in the WGHS largely validated the results from the UK Biobank. The different associations of PGS-independent ΔAAM driving AAM later with risk of CAD may have been due to differences in environmental factors between the two study cohorts.

AAM: age at menarche; ΔAAM: change in AAM; CAD: coronary artery disease; PGS: polygenic score. Dots represent estimates, bars represent standard errors.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

649 **TABLES**

Coronary artery disease					
	Change per year increase in AAM				
	Negative values of ΔAAM		**Positive values of ΔAAM**		***p* for difference between slopes**
	Odds ratio (95% CI)	*p*	Odds ratio (95% CI)	*p*	
PGS-associated ΔAAM	0.91 (0.83 to 0.99)	$3x10^{-2}$	0.92 (0.84 to 1)	0.063	0.87
PGS-independent ΔAAM	0.92 (0.88 to 0.95)	$3x10^{-6}$	1.11 (1.08 to 1.15)	$2x10^{-9}$	$7x10^{-10}$
Coronary artery disease risk factors					
	Change per year increase in AAM				
	Negative values of ΔAAM		**Positive values of ΔAAM**		***p* for difference between slopes**
	Slope (95% CI)	*p*	Slope (95% CI)	*p*	
Hemoglobin A1c (%)					
PGS-associated ΔAAM	-0.034 (-0.041 to -0.027)	$<2x10^{-16}$	-0.01 (-0.02 to -0.006)	$3x10^{-3}$	$6x10^{-5}$
PGS-independent ΔAAM	-0.022 (-0.025 to -0.019)	$<2x10^{-16}$	0.007 (0.004 to 0.01)	$1.3x10^{-6}$	$<2x10^{-16}$
Triglycerides (mg/dL)					
PGS-associated ΔAAM	-2.82 (-3.84 to -1.79)	$7x10^{-8}$	-2.63 (-3.62 to -1.64)	$2x10^{-7}$	0.83
PGS-independent ΔAAM	-4.25 (-4.7 to -3.8)	$<2x10^{-16}$	1.2 (0.78 to 1.62)	$2x10^{-8}$	$<2x10^{-16}$
High-density lipoprotein (HDL) cholesterol (mg/dL)					

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

PGS-associated ΔAAM	0.90 (0.69-1.11)	$<2x10^{-16}$	0.74 (0.54-0.95)	$8x10^{-13}$	0.38
PGS-independent ΔAAM	0.85 (0.76-0.95)	$<2x10^{-16}$	-0.28 (-0.37 to -0.20)	$2x10^{-10}$	$<2x10^{-16}$
Low-density lipoprotein (LDL) cholesterol (mg/dL)					
	Adjusted for LDL-lowering medications				
PGS-associated ΔAAM	0.27 (-0.15 to 0.69)	$2x10^{-1}$	-0.53 (-0.93 to -0.12)	$1x10^{-2}$	0.026
PGS-independent ΔAAM	-0.46 (-0.65 to -0.28)	$1x10^{-6}$	0.067 (-0.11 to 0.24)	0.45	$7x10^{-4}$
	Including only those not taking LDL-lowering medications (n = 177,542)				
PGS-associated ΔAAM	0.14 (-0.31 to 0.59)	0.054	-0.59 (-1.02 to -0.16)	$8x10^{-3}$	0.059
PGS-independent ΔAAM	-0.53 (-0.72 to -0.33)	$3x10^{-7}$	0.0065 (-0.18 to 0.19)	0.9	$2x10^{-3}$
Systolic blood pressure (mmHg)					
PGS-associated ΔAAM	-0.73 (-0.99 to -0.47)	$4x10^{-8}$	-0.71 (-0.96 to -0.45)	$5x10^{-8}$	0.91
PGS-independent ΔAAM	-0.69 (-0.81 to -0.58)	$<2x10^{-16}$	-0.21 (-0.31 to -0.10)	$2x10^{-4}$	$5x10^{-7}$
Diastolic blood pressure (mmHg)					
PGS-associated ΔAAM	-0.49 (-0.63 to -0.34)	$5x10^{-11}$	-0.49 (-0.63 to -0.35)	$9 x10^{-12}$	0.98
PGS-independent ΔAAM	-0.52 (-0.59 to -0.46)	$<2x10^{-16}$	-0.14 (-0.20 to -0.08)	$5 x10^{-6}$	$2x10^{-12}$
Body-mass index (kg/m^2)					
PGS-associated ΔAAM	-0.98 (-1.05 to -0.91)	$<2x10^{-16}$	-0.77 (-0.84 to -0.70)	$<2 x10^{-16}$	$4x10^{-4}$
PGS-independent	-0.80 (-0.83 to -0.77)	$<2x10^{-16}$	-0.079 (-0.11 to -0.051)	$4 x10^{-8}$	$<2x10^{-16}$

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

ΔAAM					
Waist-hip ratio					
PGS-associated ΔAAM	-0.0038 (-0.0047 to -0.0029)	1.1×10^{-15}	-0.0029 (-0.0038 to -0.0020)	2×10^{-10}	0.28
PGS-independent ΔAAM	-0.0046 (-0.0050 to -0.0042)	$<2 \times 10^{-16}$	0.0016 (0.0012 to 0.0020)	$<2 \times 10^{-16}$	$<2 \times 10^{-16}$

Table 1: Associations of variation in age at menarche with coronary artery disease and its risk factors.

AAM: age at menarche; ΔAAM: change in AAM; PGS: polygenic score

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

FIGURES

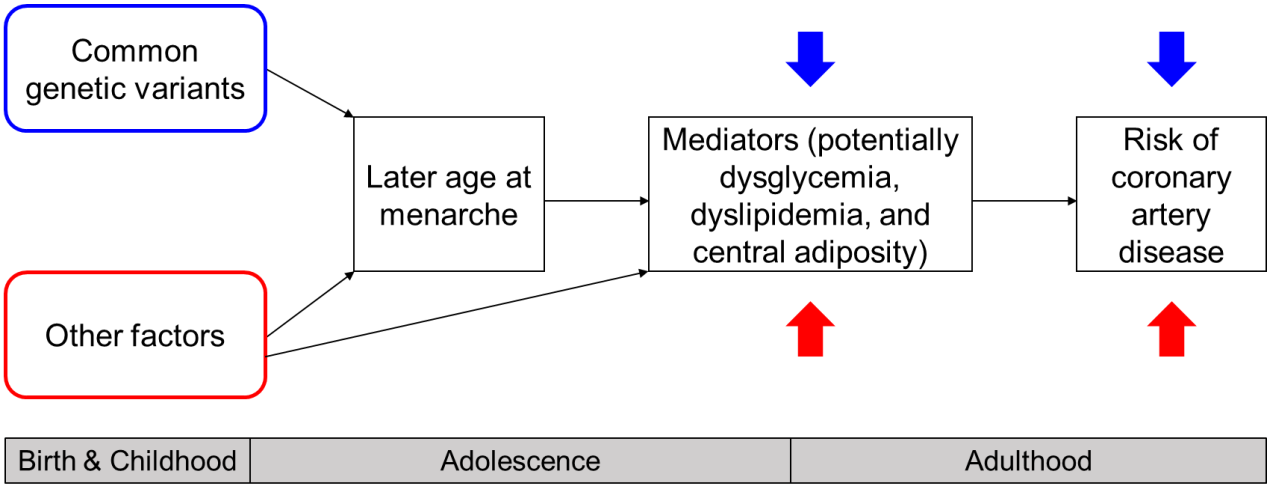
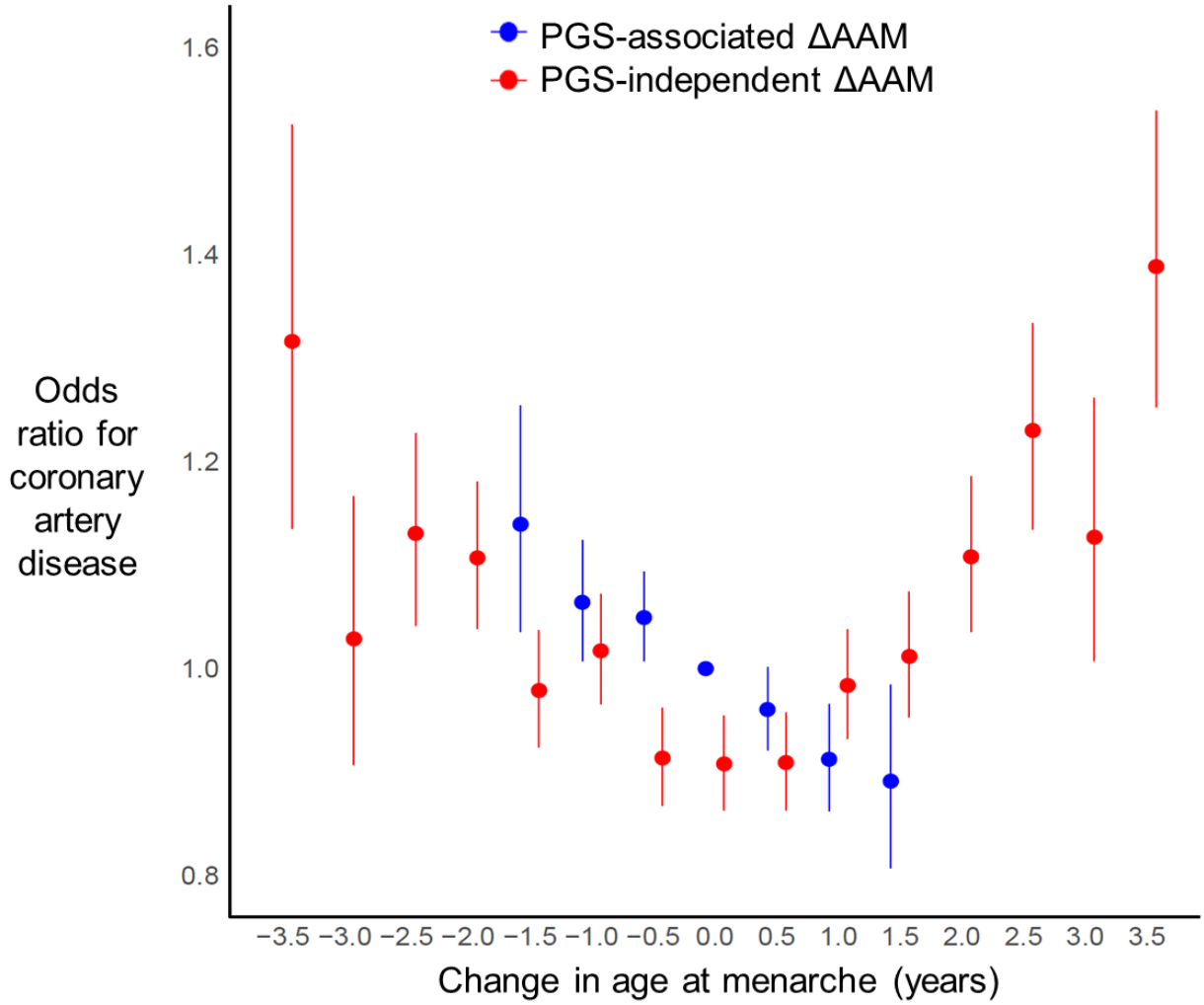


Figure 1: Relationships of different sources of variation in age at menarche (AAM) with coronary artery disease (CAD) and CAD risk factors. Variation in AAM was dissected into variation associated with common genetic variants (as estimated by a polygenic score for AAM) and variation independent of the polygenic score, then associations of these sources of variation with CAD and CAD risk factors were studied. When associated with common genetic variants, later AAM was associated with a lower risk of CAD and favorable changes in CAD risk factors (blue arrows). However, when occurring independently of the polygenic score, later AAM was associated with higher risk of CAD and unfavorable changes in CAD risk factors (red arrows). The "other factors" causing later AAM independent of the PGS have yet to be identified and could potentially include chronic illness, underweight, and rare genetic variants.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .



	-3.5	-3	-2.5	-2	-1.5	-1	-0.5	0	0.5	1	1.5	2	2.5	3	3.5
PGS-associated ΔAAM, *n*					4,703	18,503	44,781	61,427	46,406	19,729	5,488				
PGS-independent ΔAAM, *n*	1,642	2,896	6,681	12,093	17,793	22,004	25,491	27,186	25,083	21,310	15,410	10,458	6,298	3,351	3,341

Figure 2: Association of variation in AAM with odds of CAD in the UK Biobank. Variation in AAM was dissected into variation associated with common genetic variants as estimated by a PGS for AAM (PGS-associated ΔAAM) and variation independent of the PGS (PGS-independent ΔAAM), then associations of these sources of variation with risk of CAD were studied. PGS-associated ΔAAM showed linear or roughly linear associations with CAD whereas PGS-independent ΔAAM showed a U-shaped association with risk of CAD. The different

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

associations of positive values of PGS-associated ΔAAM with risk of CAD demonstrates that it is not later AAM itself that causes the increased risk of CAD and that the increased risk of CAD observed in later AAM is being driven by factors other than common genetic variants and that. To achieve bin sizes of ≥1000 individuals, the first and last bins for each variable represent the group of individuals with ΔAAM ≤ or ≥ the value listed on the x-axis, respectively. Dots represent estimates; bars represent standard errors.

AAM: age at menarche; ΔAAM: change in AAM; CAD: coronary artery disease; PGS: polygenic score.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

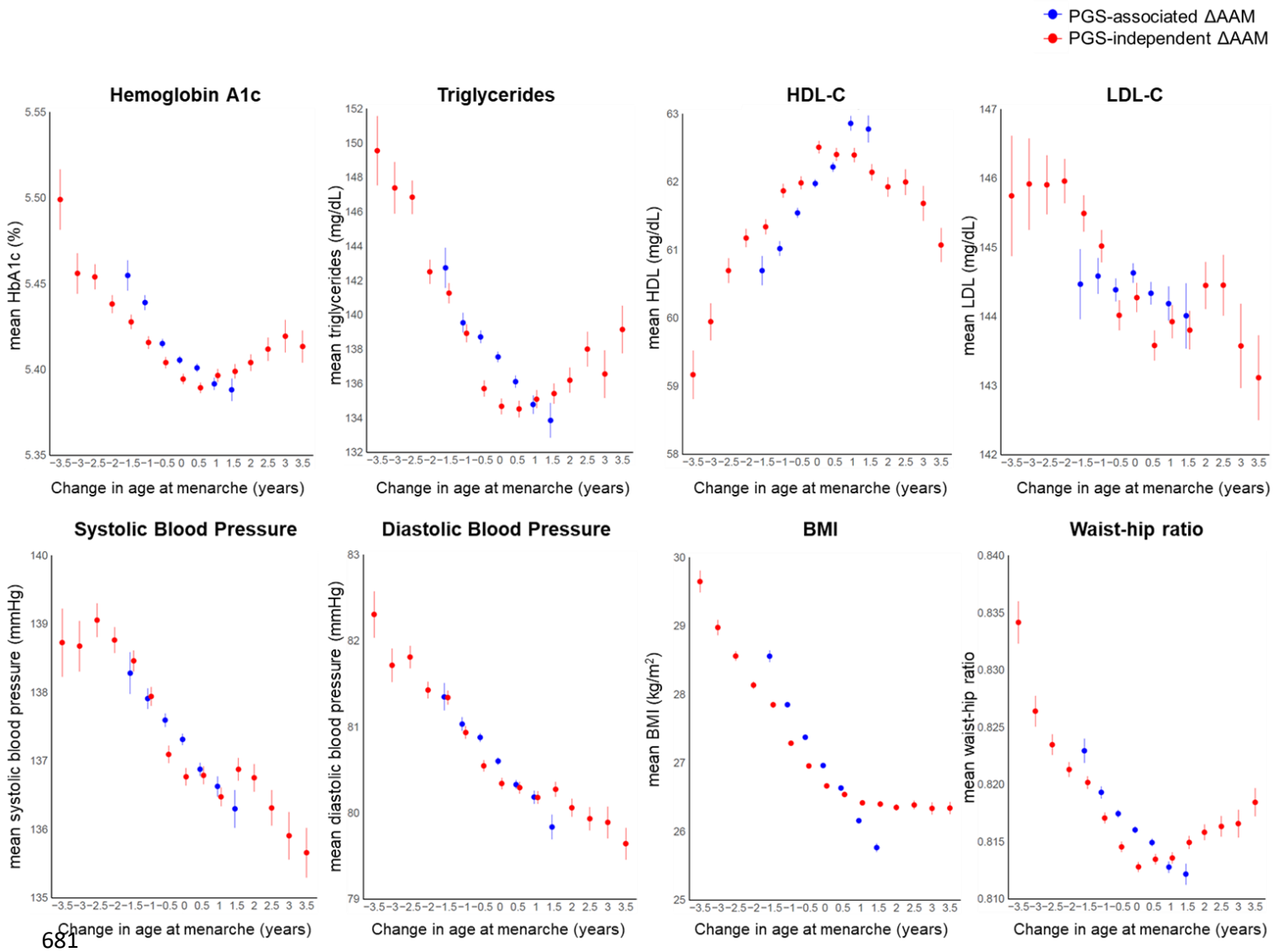


Figure 3: Associations of variation in AAM with risk factors for CAD in the UK Biobank. We observed linear (or roughly linear) relationships between PGS-associated ΔAAM and most CAD risk factors. In contrast, for PGS-independent ΔAAM we observed mostly non-linear relationships with CAD risk factors. The U-shaped relationship of PGS-independent ΔAAM with CAD was mirrored by the relationships with HbA1c, triglycerides, HDL-C, and waist-hip ratio. This raises the possibility that dysglycemia, dyslipidemia and central adiposity may mediate the relationship between later AAM and increased risk of CAD. For LDL-C, n = 177,542 after

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

excluding those taking LDL-C lowering medications. Dots represent estimates, bars represent standard errors.

AAM: age at menarche; ΔAAM: change in AAM; BMI: body mass index; CAD: coronary artery disease; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; PGS: polygenic score

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

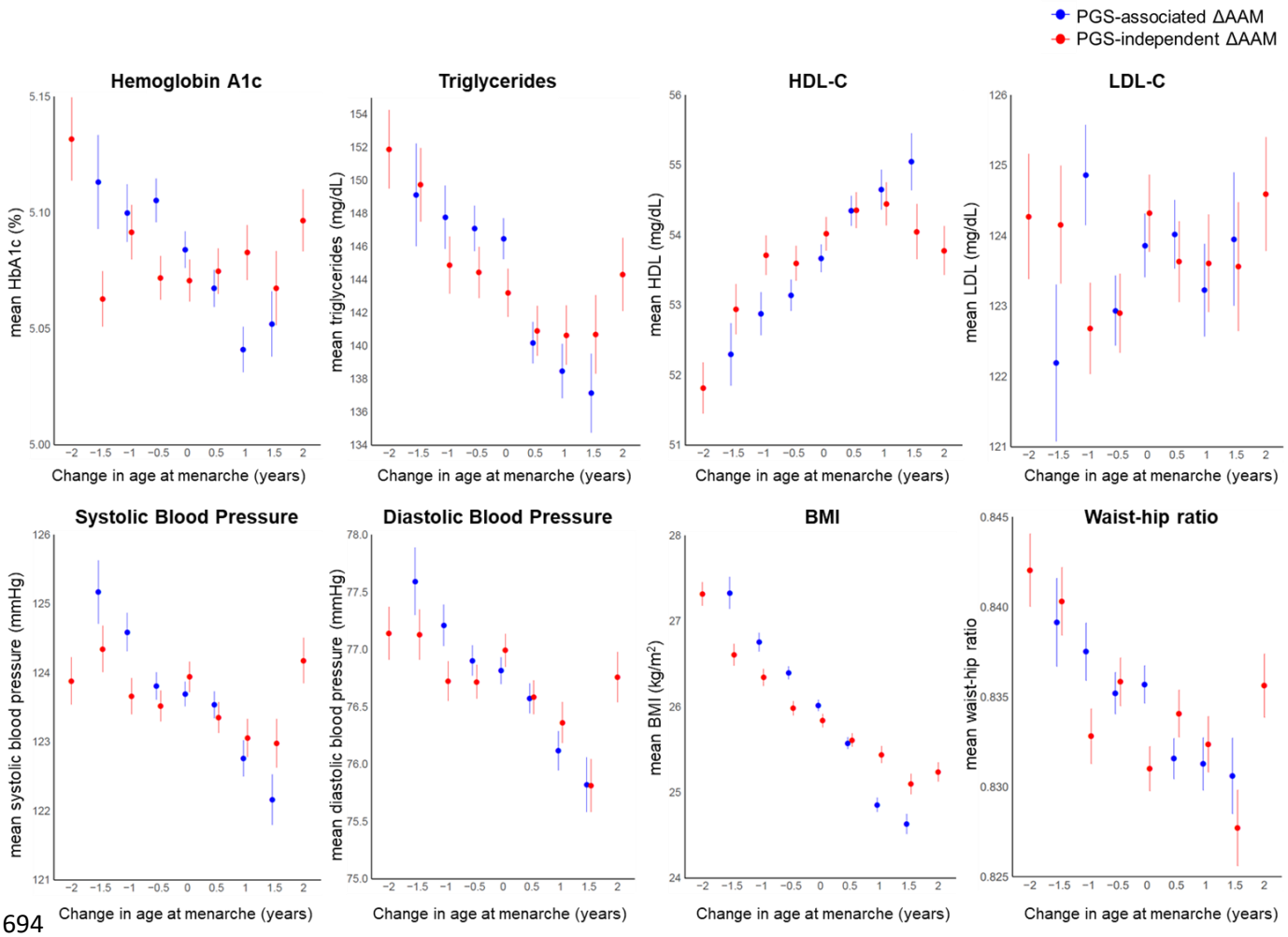
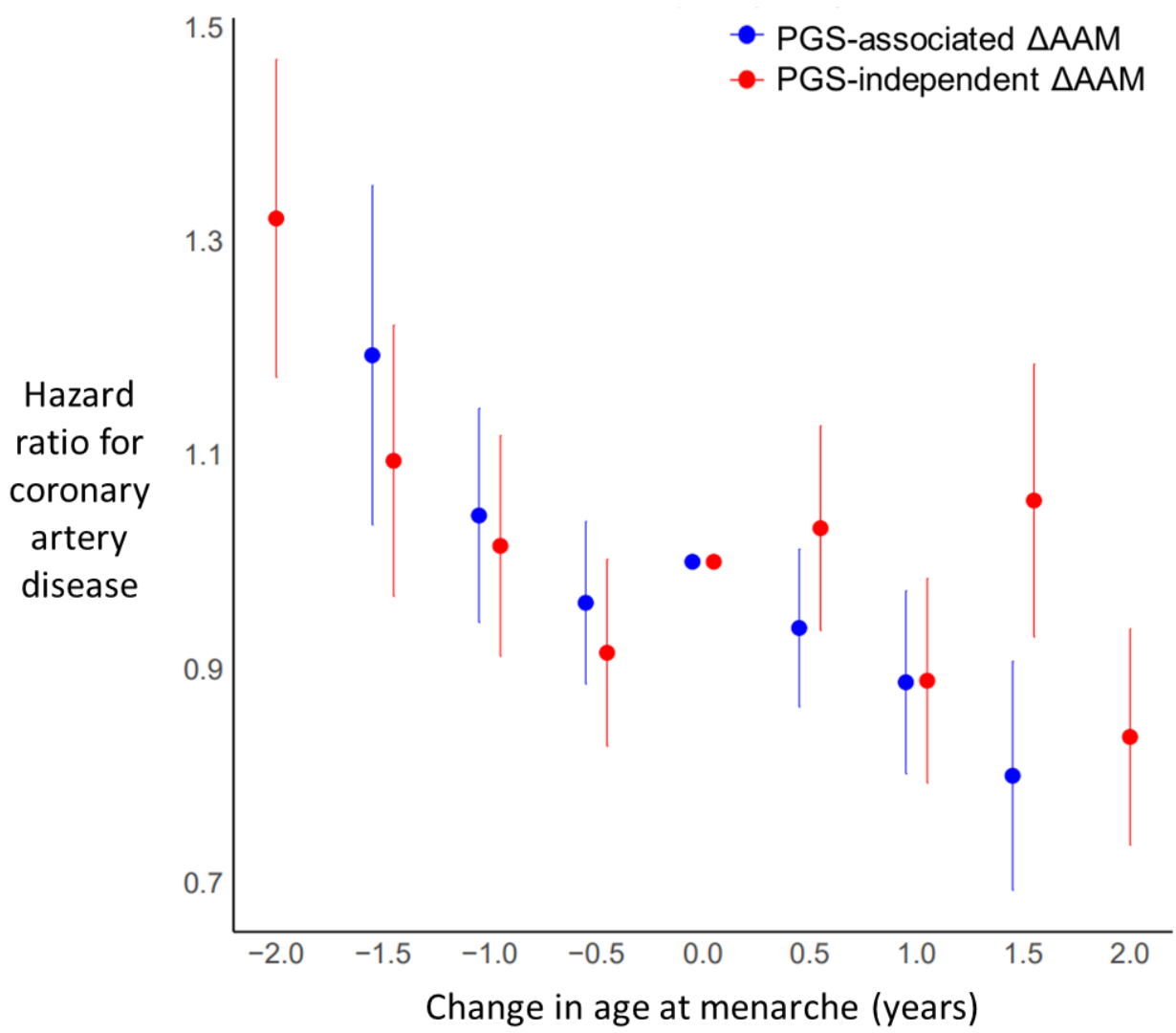


Figure 4: Associations of variation in AAM with risk factors for CAD in the WGHS. PGS-associated ΔAAM showed linear or roughly linear associations with most CAD risk factors and PGS-independent ΔAAM showed mostly non-linear associations with most CAD risk factors. Similar to results in the UK Biobank, hemoglobin A1c, triglycerides and HDL-C showed U-shaped associations with PGS-independent ΔAAM which support conclusions from the primary analyses that dysglycemia, dyslipidemia and central adiposity may mediate the relationship between later AAM and increased risk of CAD. For LDL-C, n = 22,495 after excluding those taking cholesterol-lowering medications. Dots represent estimates, bars represent standard errors.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

AAM: age at menarche; ΔAAM: change in AAM; BMI: body mass index; CAD: coronary artery disease; HDL-C: high density lipoprotein cholesterol; LDL-C: low density lipoprotein cholesterol; PGS: polygenic score; WGHS: Women's Genome Health Study.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .



	-2	-1.5	-1	-0.5	0	0.5	1	1.5	2
PGS-associated ΔAAM, *n*		954	2,454	4,784	5,945	5,016	2,784	1,331	
PGS-independent ΔAAM, *n*	1,620	1,752	2,736	3,755	4,055	3,614	2,475	1,470	1,791

Figure 5: Associations of variation in AAM with hazard ratio for risk of CAD in the WGHS. PGS-associated ΔAAM demonstrated a negative linear association with risk of CAD, similar to the results in the UKBB. PGS-independent ΔAAM demonstrated a reverse-J shaped relationship, slightly different from the U-shaped relationship seen in the UK Biobank. Analyses in the WGHS largely validated the results from the UK Biobank. The different associations of

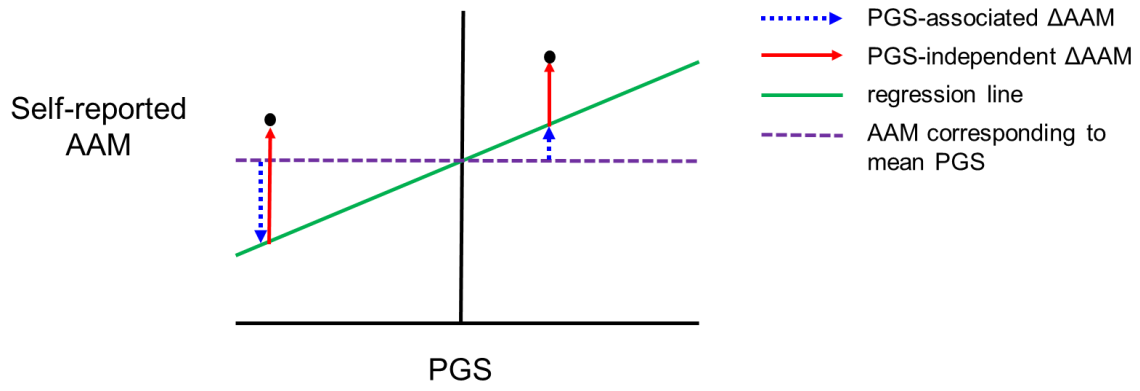
medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

PGS-independent ΔAAM driving AAM later with risk of CAD may have been due to differences in environmental factors between the two study cohorts.

AAM: age at menarche; ΔAAM: change in AAM; CAD: coronary artery disease; PGS: polygenic score. Dots represent estimates, bars represent standard errors.

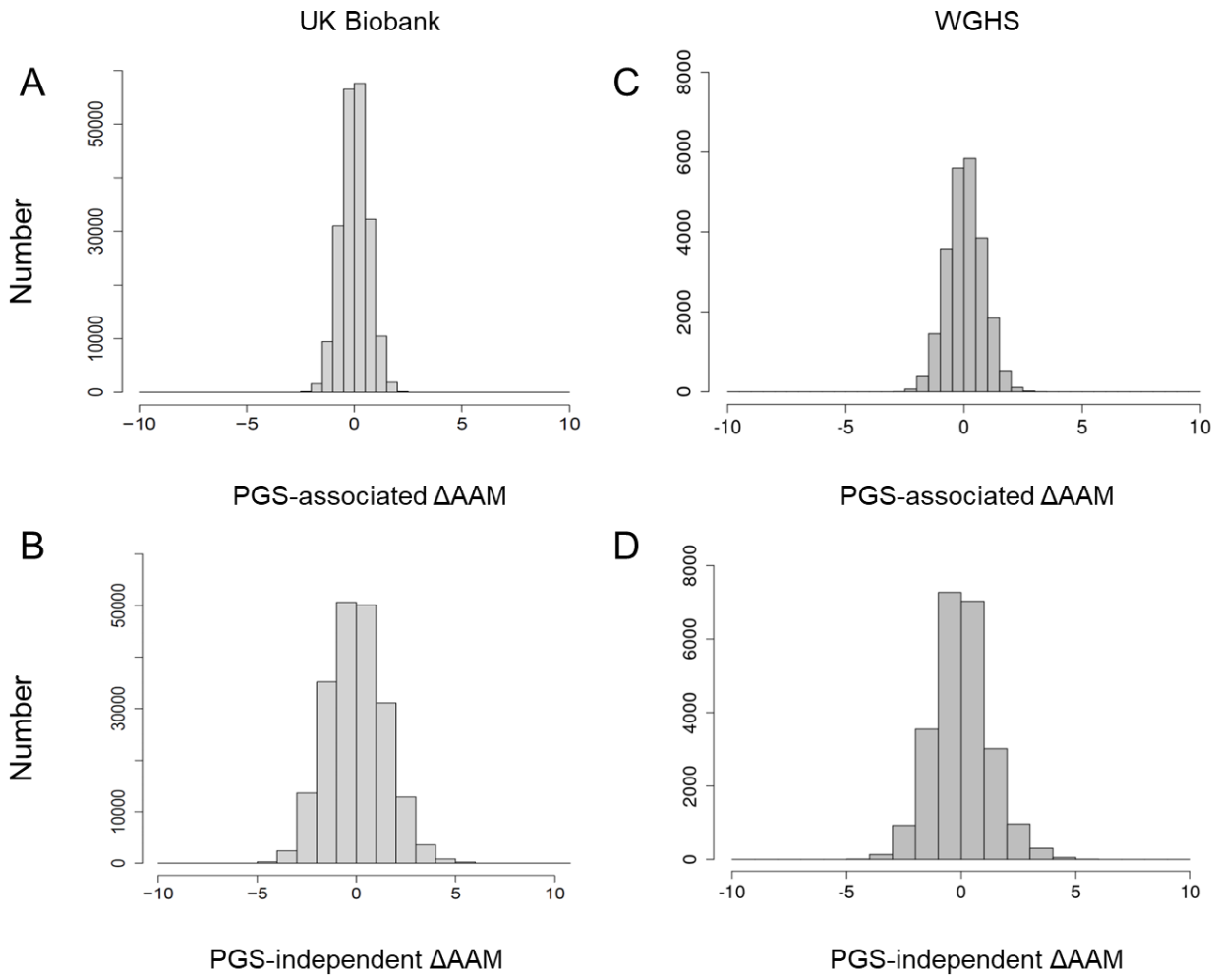
medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

SUPPLEMENTAL MATERIAL



Supplemental Figure 1: Derivation of PGS-associated and PGS-independent change in age at menarche. Self-reported age at menarche (AAM) was regressed against a polygenic score (PGS) for AAM. The green line indicates the regression line for self-reported AAM against PGS, the horizontal dashed line represents the observed AAM at the mean PGS for the cohort, the dots represent two examples of women, the blue dotted arrow represents PGS-associated change in AAM (ΔAAM), and the red dashed arrow represents PGS-independent ΔAAM. For the woman represented by the dot on the left, AAM predicted by the PGS (regression) is below the mean for the cohort, so PGS-associated ΔAAM is negative, and self-reported AAM is above the regression line, so PGS-independent ΔAAM is positive. For the woman represented by the dot on the right, AAM predicted by the PGS is above the mean for the cohort, so PGS-associated ΔAAM is positive, and self-reported AAM is above the regression line, so PGS-independent ΔAAM is also positive.

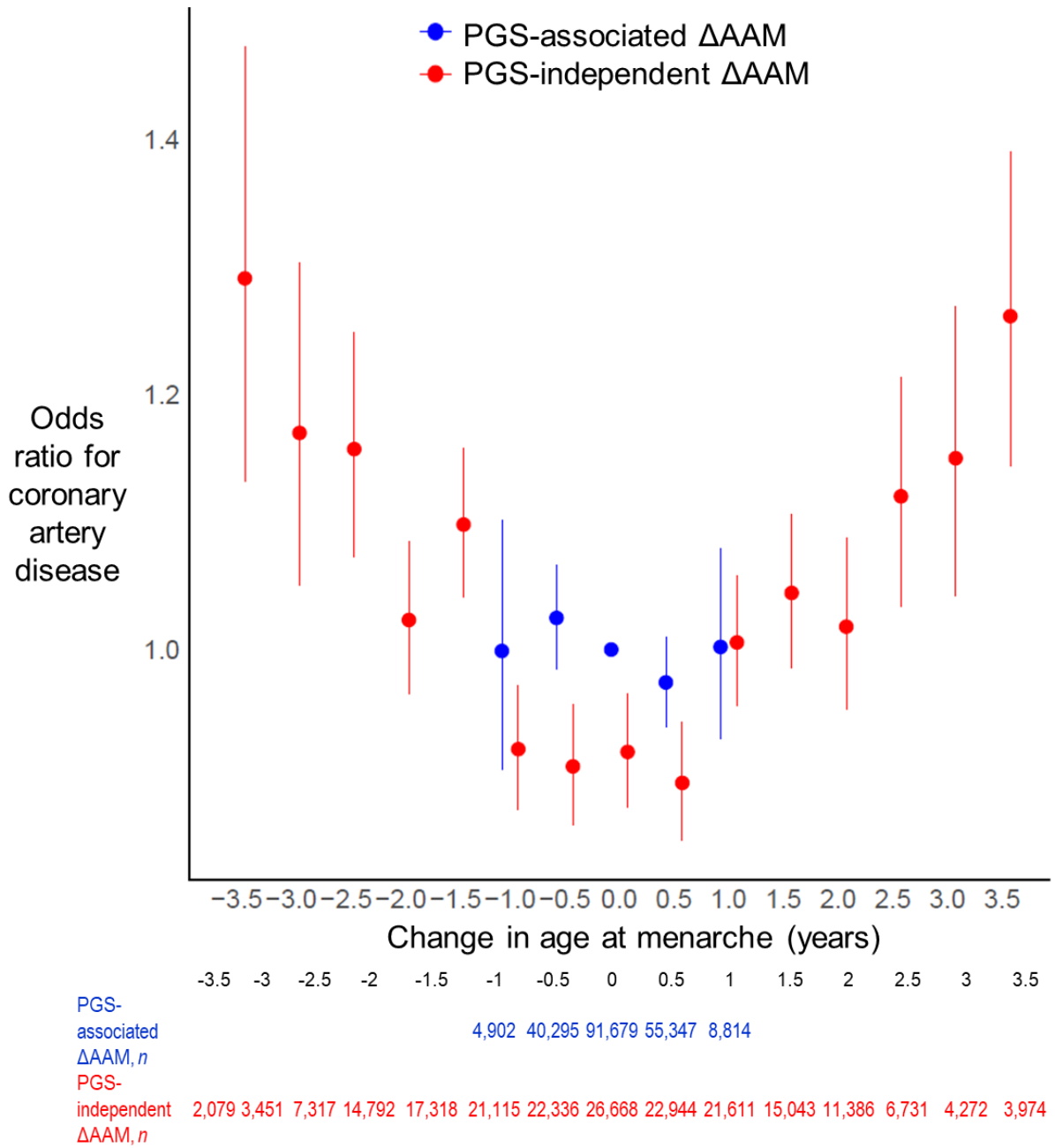
medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .



Supplemental Figure 2: Distribution of PGS-associated ΔAAM in the UK Biobank (A), PGS-independent ΔAAM in the UK Biobank (B), PGS-associated ΔAAM in the WGHS (C), and PGS-independent ΔAAM in the WGHS (D).

AAM: age at menarche; ΔAAM: change in AAM; PGS: polygenic score, WGHS: Women's Genome Health Study.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .



Supplemental Figure 3: Association of variation in age at menarche with coronary artery disease in the UK Biobank using a PGS based on a 2014 genome-wide association study meta-analysis that did not include data from the UK Biobank.

AAM: age at menarche; ΔAAM: change in AAM; PGS: polygenic score. Dots represent estimates, bars represent standard errors.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.14.24312022; this version posted August 20, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

	Change per year increase in AAM				
	Negative values of ΔAAM		**Positive values of ΔAAM**		***p* for difference between slopes**
Coronary artery disease					
	Odds ratio (95% CI)	*p*	Odds ratio (95% CI)	*p*	
PGS-associated ΔAAM	0.89 (0.77 to 1.04)	0.14	0.98 (0.86 to 1.1)	0.79	0.43
PGS-independent ΔAAM	0.91 (0.88 to 0.94)	$5x10^{-7}$	1.09 (1.06 to 1.13)	$4x10^{-7}$	$5x10^{-9}$

Supplemental Table 1: Associations between variation in age at menarche and coronary artery disease using a PGS based on a 2014 genome-wide association study meta-analysis that did not include data from the UK Biobank.

AAM: age at menarche; ΔAAM: change in AAM; PGS: polygenic score