Cardiovascular Diabetology

RESEARCH

Open Access

Additive association of blood group A allele with 15 cardiometabolic diseases: a UK Biobank life-course study

Ran Zhao[1†], Wenyan Xian[1†], Yihao Ma[2], Valerio Napolioni[3], Patrick W. C. Lau[4], Xiao-Li Tian[5], Yann Le Guen[6], Andre Franke[7] and Jie Huang[1*]

Abstract

Background Although existing studies have reported associations between blood group A and cardiometabolic diseases (CMD), most have focused on dominant inheritance models. However, genome-wide association studies have mostly been based on additive genotypes. This study aims to investigate the association between the blood group A allele and 15 CMD using recessive, dominant, and additive models and identify potential mediators.

Methods This study leveraged data from over 320,000 participants with O and A blood groups in the UK Biobank to investigate the association between blood group A allele and 15 major CMD under recessive, dominant, and gene dosage (additive) models. Protein data from nearly 30,000 participants were used to analyze the association between ABO protein levels and CMD. Mediation analysis further explored whether blood cell count traits and blood biochemistry mediate the association between the number of A allele and CMD.

Results The additive model demonstrates a dose–response association of the blood group A allele with venous thromboembolism (VTE), myocardial infarction (MI), ischemic stroke (IS), type 2 diabetes mellitus (T2DM), and heart failure (HF), among others. Each additional A allele increased disease risk, particularly for VTE (HR = 1.273, P[FDR] = 4.43×10^{-96}). ABO protein levels also correlated with five CMD outcomes, particularly VTE and coronary artery disease (CAD). Mediation analyses revealed that blood cell traits (e.g., hemoglobin, hematocrit) and biochemistries (e.g., aspartate aminotransferase to alanine aminotransferase ratio, apolipoprotein B) significantly mediated the associations for specific CMD, suggesting shared biological mechanisms.

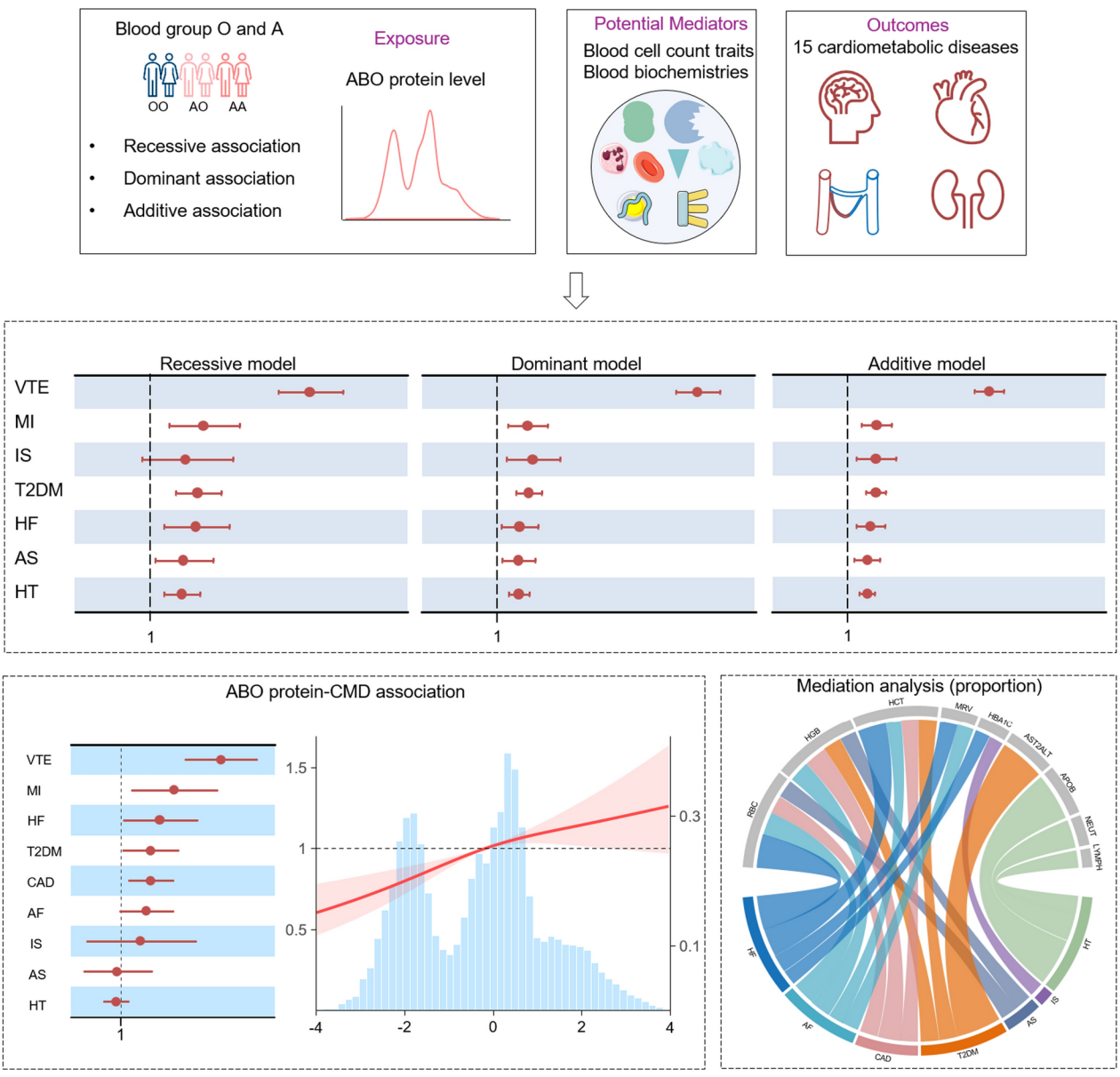
Conclusions Our findings reveal that blood group A allele is associated with an increased risk of multiple CMD, particularly under the additive model. Some blood cell count traits and blood biochemistries play significant mediating roles.

†Ran Zhao and Wenyan Xian contributed equally to this study.

*Correspondence:
Jie Huang
huangj@sustech.edu.cn

Full list of author information is available at the end of the article

Graphical abstract



Research insights

What is currently known about this topic? Blood group A is linked to higher risks of thrombosis and cardiovascular diseases, most have focused on dominant inheritance models.

What is the key research question? How does the blood group A allele influence the risk of 15 cardiometabolic diseases under different inheritance models?

What is new? The additive model demonstrates a dose-dependent association of the A allele with certain CMD risk.

ABO protein levels correlate with VTE, CAD, and T2DM.

Blood cell traits mediate A allele's effect on CMD.

How might this study influence clinical practice? Findings may enable personalized risk stratification and targeted screening for CMD based on blood group and biomarkers.

Keywords Blood group, ABO A allele, Additive association, Cardiometabolic disease, Mediation analysis

Introduction

Cardiometabolic diseases (CMD) have become leading causes of death and disability worldwide and are the primary contributors to the burden of noncommunicable diseases (NCD), including coronary heart disease, stroke, and diabetes [1]. These conditions often share common etiological factors, such as tobacco use, obesity, physical inactivity, and dietary habits [2], as well as overlapping genetic traits among certain diseases [3].

Several studies have established associations between blood group A and certain CMD. A comprehensive study revealed that individuals with blood group A have a higher risk of thrombosis, heart failure (HF), and myocardial infarction (MI), while those with blood group O exhibit a higher risk of hypertension (HT) [4]. A phenome-wide association study has also identified complex relationships of ABO blood groups with cardiovascular diseases (CVD) and their risk factors [5]. Additionally, two genome-wide association studies have pinpointed *ABO* gene loci linked to coronary artery disease (CAD) [6, 7]. These findings provide valuable insights into the association between blood group and CMD.

However, several knowledge gaps remain. First, genetic tests now allow precise differentiation of blood group haplotypes (e.g., AO and AA) [8], offering an opportunity to explore the association between ABO blood groups and diseases through various inheritance models (including recessive, dominant, and additive models). Despite this, most studies have focused on the associations under the dominant model (treating haplotype AA the same as AO). Conversely, in numerous genome-wide association studies, the additive model has been the default [9, 10]. Thus, we hypothesize that the additive model can best explain the association between blood group A and CMD. A cross-sectional study reported differences in the risk of pulmonary embolism (PE) and deep vein thrombosis (DVT) between heterozygous and homozygous allele carriers [11]. Secondly, no studies have yet examined the association between ABO protein levels and CMD. Moreover, the underlying mechanisms linking blood group A and CMD have not been investigated. Accumulating evidence suggests strong associations of blood groups with several blood cell traits and blood biochemistries [12, 13], such as leukocyte counts and lipid traits, which are critical biomarkers or risk factors of CMD [14, 15]. Yet evidence regarding whether these blood traits contribute to associations between blood groups and CMD remains scarce. Comprehensive longitudinal study is hence imperative to scrutinize associations between blood group A allele and CMD, as well as their potential mechanisms.

This study leverages data from over 320,000 participants with O and A blood groups in the UK Biobank, categorizing them into three groups (OO, AO, and AA) based on the number of A allele. The associations between ABO A allele and 15 major CMD are investigated under dominant, recessive, and gene dosage (additive) models. Individual-level protein data from nearly 30,000 participants are utilized to analyze the association between ABO protein levels and CMD. Mediation analysis further reveals whether 29 blood cell count traits and 29 biochemistries drive the mechanisms by which the number of A allele is associated with CMD.

Methods

Study population

The study sample was drawn from the UK Biobank, which recruited approximately 500,000 participants aged 37–73 between 2006 and 2010 [16]. We restricted the analysis to Caucasian participants with blood types A and O, excluding individuals with blood types B and AB due to their smaller sample sizes. Ancestry was defined based on principal component analysis-determined genetic clustering. In this study, we included only Caucasian participants to avoid confusion due to genetic ancestry differences, as the Caucasian population accounts for nearly 95%. Participants with missing disease diagnosis information and covariates (including those who replied "prefer not to answer" or "do not know" in questionnaires) were excluded, as were participants who were lost to follow-up or withdrew from the study during the follow-up period.

Ascertainment of blood group A

Genetically, blood group A was determined based on three key single nucleotide polymorphisms (SNP): rs505922, rs8176719, and rs8176746 on chromosome 9q34.2 [8]. A deletion at rs8176719 and a T allele at rs505922 indicated the O haplotype. The B haplotype was indicated by a T allele at rs8176746. The number of A allele was determined as follows: individuals with rs8176746 C/C and rs8176719 G/G were classified as having homozygous AA blood group genotype (A allele = 2); those with rs8176746 C/A or C/C and rs8176719 del/G or G/G were classified as heterozygous AO blood group genotype (A allele = 1); and those with rs8176746 C/C and rs8176719 del/del were classified as OO blood type (A allele = 0).

Ascertainment of clinical endpoints

The incidence of 15 major CMD was analyzed as outcomes, including all types of stroke (AS), ischemic stroke (IS), hemorrhagic stroke (HS), transient ischemic attack (TIA), CAD, MI, atrial fibrillation (AF), HF, cardiomyopathy (CM), aortic aneurysm (AA), HT, venous thromboembolism (VTE), peripheral vascular disease (PAD), type 2 diabetes mellitus (T2DM), and chronic kidney disease (CKD). Diagnoses for all conditions were determined based on ICD-10 codes and self-reported data. Detailed

ICD codes and self-report codes can be found in the Supplementary Table 1. The date of first occurrence for each disease was determined using the 'first occurrence date' fields from the UK Biobank, which integrates data from hospital records, death registries, and primary care to accurately identify the timing of the first occurrence of each disease.

Ascertainment of covariates

Age was determined based on the date of attendance and self-reported date of birth. Sex and education level were self-reported. The Townsend deprivation index was calculated before participants joined the UK Biobank. Body mass index (BMI) and waist-to-hip ratio (WHR) are both important factors that may influence outcomes [17, 18]. BMI values here were constructed from height and weight measured during the initial Assessment Centre visit. WHR was calculated from waist and hip circumference measurements taken during the same visit. Smoking and drinking status were self-reported and categorized into current, previous and never status. Weekly moderate physical activity refers to the number of days in a week during which moderate physical activity (excluding walking) lasting 10 min or longer was performed, as recorded in a questionnaire. Secretor status was determined by rs601338 within the *FUT2* gene, where genotypes G/G and G/A correspond to secretors, and genotype A/A corresponds to non-secretors. Lastly, the first five principal components were selected to control confounding effects due to population structure and differences in genetic background.

Follow-up

In this study, follow-up time was defined as the duration from birth to the end of follow-up. Since blood type affects one's health from birth, we did not exclude individuals who have already been diagnosed at baseline. This is different from many other lifestyle studies. We therefore categorized our study as "life-course". The follow-up period was terminated under the following conditions: (1) The occurrence of the outcome event, where follow-up ended at the date of the first recorded outcome; (2) Death, where follow-up ended at the date of death if it occurred before the outcome; (3) Loss to follow-up, where follow-up ended at the date of withdrawal or the last contact; and (4) Reaching the study end date, set as December 31, 2023, for participants without a recorded outcome, death, or loss to follow-up.

Association of blood group A allele with CMD

Cox proportional hazards model was used to estimate the association between blood group A allele and CMD. Three inheritance models were used: recessive model, dominant model, and additive model. The recessive model compares AA (A allele = 2) with AO/OO (A allele = 1 or 0), testing if only AA is associated with CMD. The dominant model compares AO/AA with OO, evaluating whether having at least one A allele (AO or AA) is associated with CMD. The additive model treats the number of A allele (0, 1, 2) as a continuous variable, assessing whether the association between A allele and CMD is an additive effect. In addition, CMD with significant associations in the additive model were further used to analyze the association with normalized ABO protein levels, and a restricted cubic spline (RCS) was introduced into the Cox models to examine the non-linear effects. For sensitivity analysis, we included an additional set of covariates for model adjustment, covering blood cell count traits that may influence CMD, including red blood cell count (RBC), hemoglobin (HGB), white blood cell count (WBC), neutrophil count (NEUT), and platelet count (PLT) [19, 20].

Age, sex, education, Townsend deprivation index, BMI, WHR, smoking status, drinking status, weekly moderate physical activity days, secretor status, and the first five principal components were included as covariates in all models. To compare the goodness of fit of the models, we calculated the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for each model, where lower values indicate a better fit. The Schoenfeld residuals test was used to check the proportional hazards assumption in the Cox models. Some covariates, including sex and smoking status, violated this assumption. For these covariates, stratified Cox models were applied to ensure that baseline hazards remained constant over time. Variance inflation factors across all models indicated acceptable levels of low multicollinearity (VIF < 1.8).

Association of potential mediators with blood group A allele and CMD

To further investigate the mechanisms through which the blood group A allele influences CMD incidence, we selected 29 blood cell count traits and 29 biochemistries as potential mediators. The blood cell count traits cover those related to red blood cells, white blood cells, and platelets, while the blood biochemistries include liver and kidney function markers, glycolipid metabolites, inflammatory factors, and hormones, among others. A detailed list of the selected blood biomarkers and blood count indicators is provided in Supplementary Table 2. All potential mediators were standardized, multivariable linear regression models were applied to analyze the relationship between the A allele and each potential mediator, adjusting for all covariates. Cox models were subsequently used to assess the association between potential mediators and CMD.

Mediation analysis

Given the exploratory nature of mediation analysis, we applied a data-driven approach to comprehensively analyze all available blood cell count traits and blood biochemistries as potential mediators. For the mediation analyses involving blood group A allele, potential mediators, and CMD, linear regression was used for association between the exposure and mediators, and survival regression model was employed for association between the CMD and potential mediators. Unlike Cox models, survival regression models were used to estimate time-to-event outcomes with a focus on linear relationships between exposure and outcome. A negative association indicates that the exposure accelerates the onset of the outcome. All covariates were adjusted to control confounding effects in the mediation pathways.

All analyses were performed on R 4.4.1. Mediation analysis was conducted using the "mediation" package in R. All p-values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) correction. The study overview is shown in Fig. 1.

Results

Study population

This study included 326,526 individuals with blood group O and A from the UK Biobank. Of these, 160,640 individuals had blood group OO, 136,929 had haplotype AO, and 28,957 had haplotype AA. The mean age of participants was 56.75 years, and 46.3% were male (N = 151,155). The number of cases of different CMD ranged from 2914 (CM) to 129,684 (HT), and the median age of onset ranged from 60.4 years (VTE) to 72.9 years (CM). The protein levels of ABO were available for further analysis in 29,060 individuals. Baseline characteristics are shown in Table 1. A detailed flowchart of the

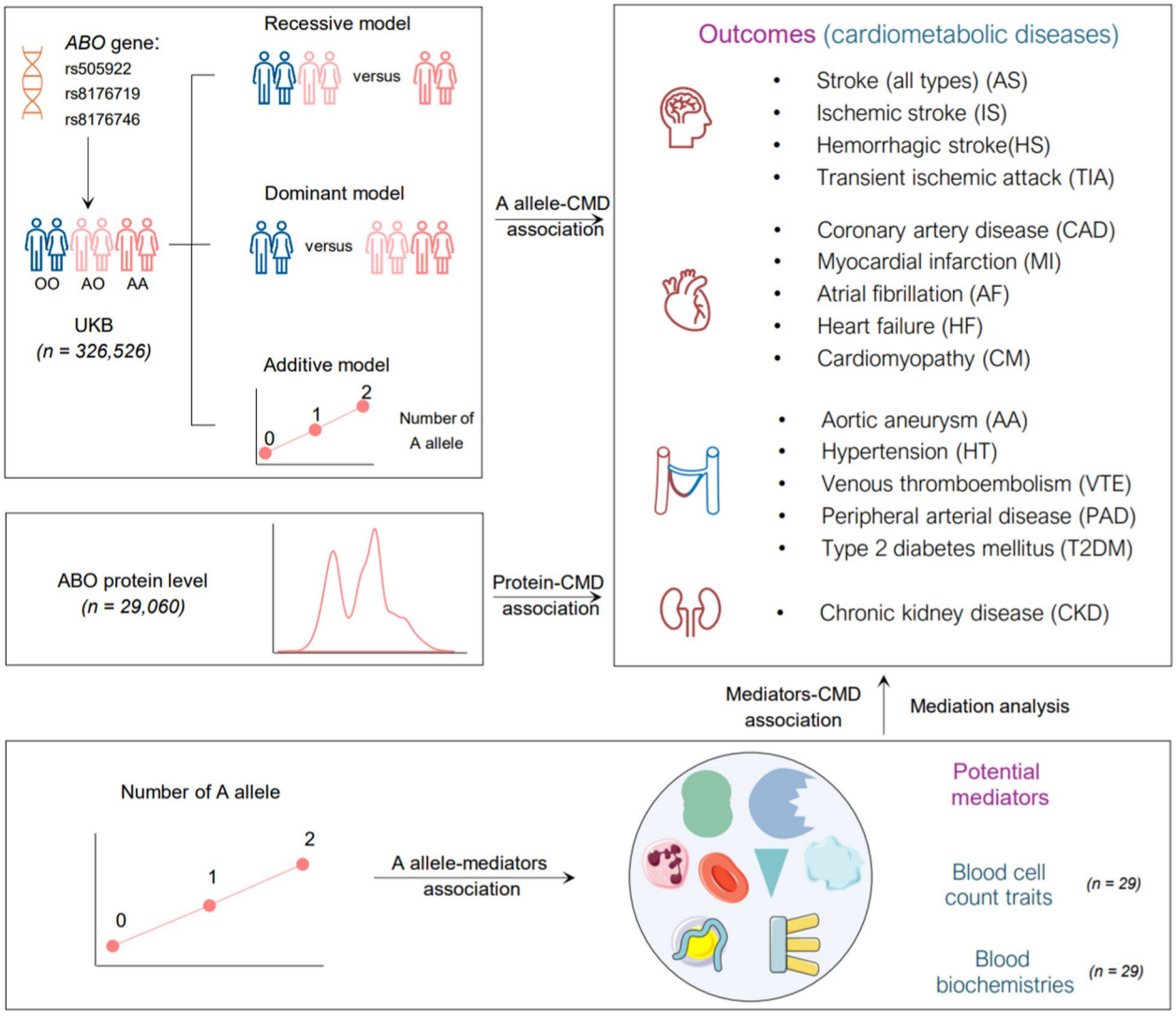


Fig. 1 Study overview

Table 1 Baseline characteristics of study cohort*

Characteristics	All (N = 326,526)	OO (N = 160,640)	AO (N = 136,929)	AA (N = 28,957)
Age	56.75 (8.02)	56.75 (8.03)	56.75 (8.02)	56.76 (8.02)
Male	151,155 (46.3%)	74,533 (46.4%)	63,194 (46.2%)	13,428 (46.4%)
BMI	27.31 (4.70)	27.30 (4.70)	27.31 (4.70)	27.31 (4.74)
WHR	0.87 (0.09)	0.87 (0.09)	0.87 (0.09)	0.87 (0.09)
Education	1266,692 (38.8%)	62,489 (38.9%)	53,402 (39%)	11,091 (38.3%)
College degree or above	199,834 (61.2%)	98,151 (61.1%)	83,527 (61%)	17,506 (61.7%)
Other degree				
Townsend deprivation index	− 1.55 (2.90)	− 1.53 (2.91)	− 1.56 (2.89)	− 1.57 (2.90)
Smoke status	176,370 (54.0%)	86,508 (53.9%)	74,098 (54.1%)	15,764 (54.4%)
Never	117,604 (36.0%) 32,552 (10.0%)	57,962 (36.1%)	49,269 (36.0%)	10,373 (35.8%)
Previous		16,170 (10.1%)	13,562 (9.9%)	2,820 (9.7%)
Current				
Drinking status	9,882 (3.0%)	4,874 (3.0%)	4,115 (3.0%)	893 (3.1%)
Never	10,750 (3.3%)	5,283 (3.3%)	4,515 (3.3%)	952 (3.3%)
Previous	305,894 (93.7%)	150,483 (93.7%)	128,299 (93.7%)	27,112 (93.6%)
Current				
Weekly physical activity days	3.65 (2.33)	3.66 (2.33)	3.64 (2.33)	3.63 (2.33)
Secretor status	243,537 (74.6%)	119,378 (74.3%)	102,361 (74.8%)	21,798 (75.3%)
Secretor	82,989 (25.4%)	41,262 (25.7%)	34,568 (25.2%)	7,159 (24.7%)
Non-secretor				

*Continuous variables are presented as mean (standard deviation), and categorical variables are presented as N (%). *BMI* Body mass index, *WHR* Waist to hip ratio

participant selection process can be found in Supplementary Fig. 1. Follow-up age is shown in Supplementary Fig. 2.

Association of blood group A allele with CMD

The association between blood group A allele and CMD varies across different inheritance models. In the traditional dominant model, compared to individuals with blood group O, those with blood group A have a 1.388-fold increased risk of VTE (HR = 1.388 [1.346–1.432], P[FDR] = 5.18×10^{-95}). For other CMD, including MI, IS, T2DM, HF, AS, CAD, and AF, the risks show a slight increase, ranging from 1.037 to 1.066-fold, with all FDR-corrected P-values < 0.05. Only the risk of HT is decreased (HR = 0.981 [0.971–0.992], P[FDR] = 2.7×10^{-3}) (Fig. 2b). However, in the recessive model, IS and HT are no longer significant (Fig. 2a). In the additive model, each additional A allele increases the risk of VTE by 1.273-fold (HR = 1.273 [1.245–1.302], P[FDR] = 4.43×10^{-96}), with lower P-value compared to the dominant model. This means individuals homozygous for the A allele (AA) would have a 1.546-fold higher cumulative VTE risk compared to OO carriers. A similar pattern is observed for MI, IS, T2DM, HF, AS, CAD, and AF where the additive effect is more pronounced than the dominant effect (Fig. 2c). Yet HT shows the lowest P-value in dominant model. The additive models provided the best fit and stability for VTE, MI, IS, T2DM, HF, AS, CAD, and AF (lowest AIC and BIC) (Supplementary Table 3), and the dominant model was optimal for HT. Similarly, AS and AF are no longer significant in the recessive model but retain significance in the additive model. Sensitivity analyses, which added blood cell counts as covariates, showed that HF was no longer significant in the dominant model but remained significant in the recessive and additive models (see Supplementary Fig. 1).

Further analyses involving 29,060 individuals explored the association between ABO protein levels and CMD that were significant in the additive model. Figure 3a shows the standardized ABO protein levels across different blood groups. Significant associations were identified between ABO protein levels and VTE, MI, HF, T2DM, and CAD. For VTE, each standard deviation increase in ABO protein level raises the risk by 1.097-fold (HR = 1.097 [1.062–1.132], P[FDR] = 1.35×10^{-8}) (Fig. 3b). Non-linearity was tested using RCS, but the risk for all diseases increased linearly (P for non-linear > 0.05) (Fig. 3c). In sensitivity analyses, the association between HF and ABO protein level was no longer significant (see Supplementary Fig. 4).

Association of potential mediators with blood group A allele and CMD

Nine CMD significantly associated with the number of A allele in the additive model were analyzed for their relationships with potential mediators. Multiple blood cell count traits and blood biochemistries showed significant associations with outcomes. For example, WBC, NEUT, and eosinophil (EO) counts exhibited strong correlations with diseases such as T2DM, HF, CAD, and HT. Certain lipid metabolism biochemistries, such as apolipoprotein B (APOB), total cholesterol (TC), and low-density lipoprotein cholesterol (LDL), were associated with almost all outcomes. The number of A allele was significantly

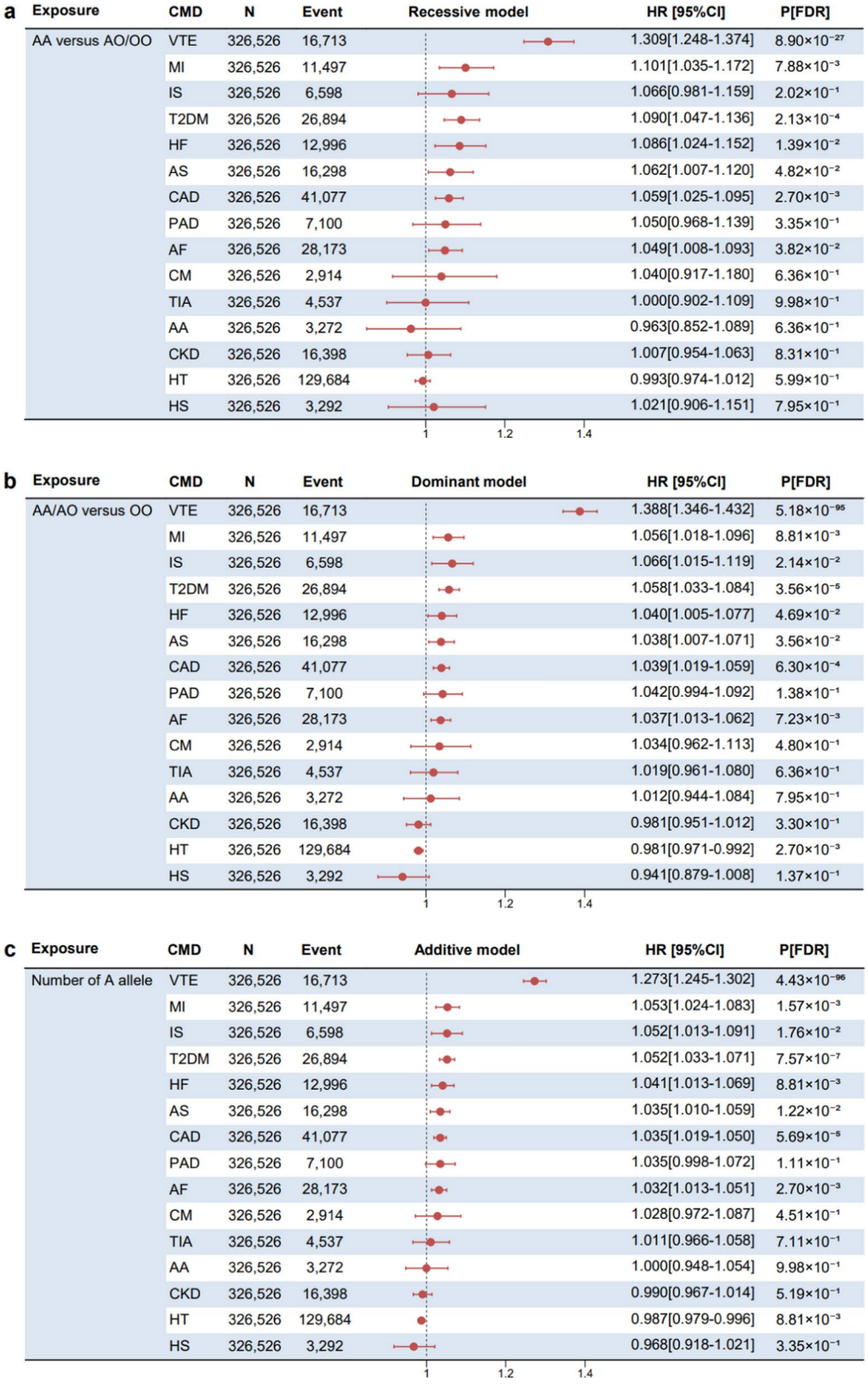


Fig. 2 Forest plots show association of blood group A allele and 15 major cardiometabolic diseases in different inheritance models, expressed as hazard ratios (HR) with 95% confidence intervals (CI). **a** Recessive mode; **b** Dominant model; **c** Additive model. *VTE* Venous thromboembolism, *MI* Myocardial infarction, *IS* Ischemic stroke, *T2DM* Type 2 diabetes mellitus, *HF* Heart failure, *AS* All types of stroke, *CAD* Coronary artery disease, *PAD* Peripheral vascular disease, *AF* Atrial fibrillation, *CM* Cardiomyopathy, *TIA* Transient ischemic attack, *AA* Aortic aneurysm, *CKD* Chronic kidney disease, *HT* hypertension, *HS* Hemorrhagic stroke

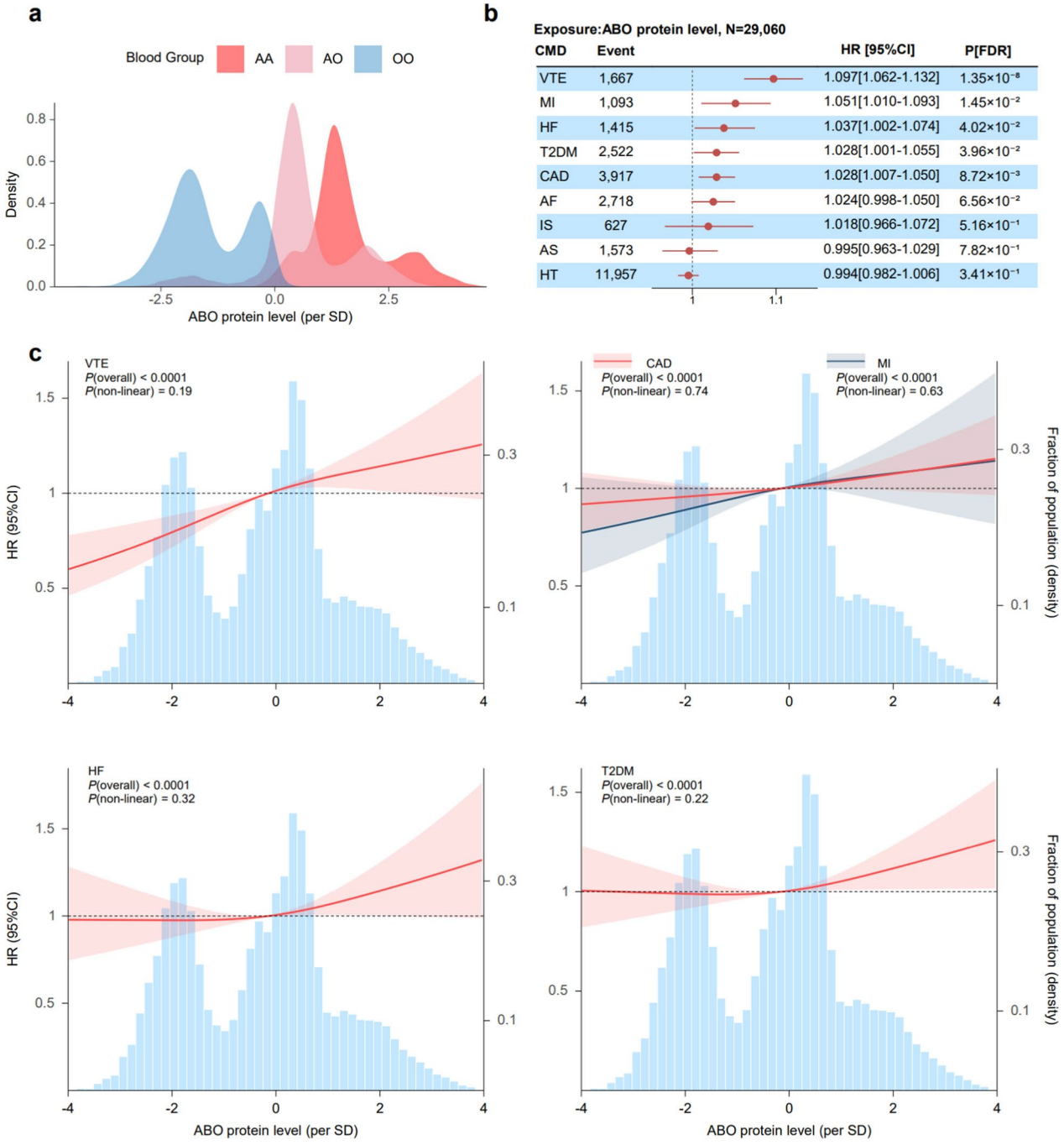


Fig. 3 **a** Density plot shows ABO protein levels in different blood group; **b** Forest plot shows association of ABO protein level with cardiometabolic diseases, expressed as hazard ratios (HR) with 95% confidence intervals (CI); **c** Restricted cubic spline curves show the non-linearity association of ABO protein level with VTE, CAD, MI, HF, and T2DM. **P*[FDR] < 0.05, ***P*[FDR] < 0.01, ****P*[FDR] < 0.001. SD = standard deviation, *VTE* Venous thromboembolism, *MI* Myocardial infarction, *HF* Heart failure, *T2DM* Type 2 diabetes mellitus, *CAD* Coronary artery disease, *AF* Atrial fibrillation, *IS* Ischemic stroke, *AS* All types of stroke, *HT* Hypertension

negatively associated with RBC, HGB, hematocrit (HCT), and the aspartate aminotransferase to alanine aminotransferase ratio (AST/ALT). The most significant negative correlation was observed with alkaline phosphatase (ALP). Detailed results are presented in Fig. 4a and Supplementary Table 4.

Mediation analysis

RBC, HGB, and HCT mediated the associations between the blood group A allele and certain CVD, including HF, AF, and CAD. The mediation proportions ranged from 11.23% (HCT for AF) to 24.6% (RBC for HF). AST/ALT mediated 33.01% of the association between the number

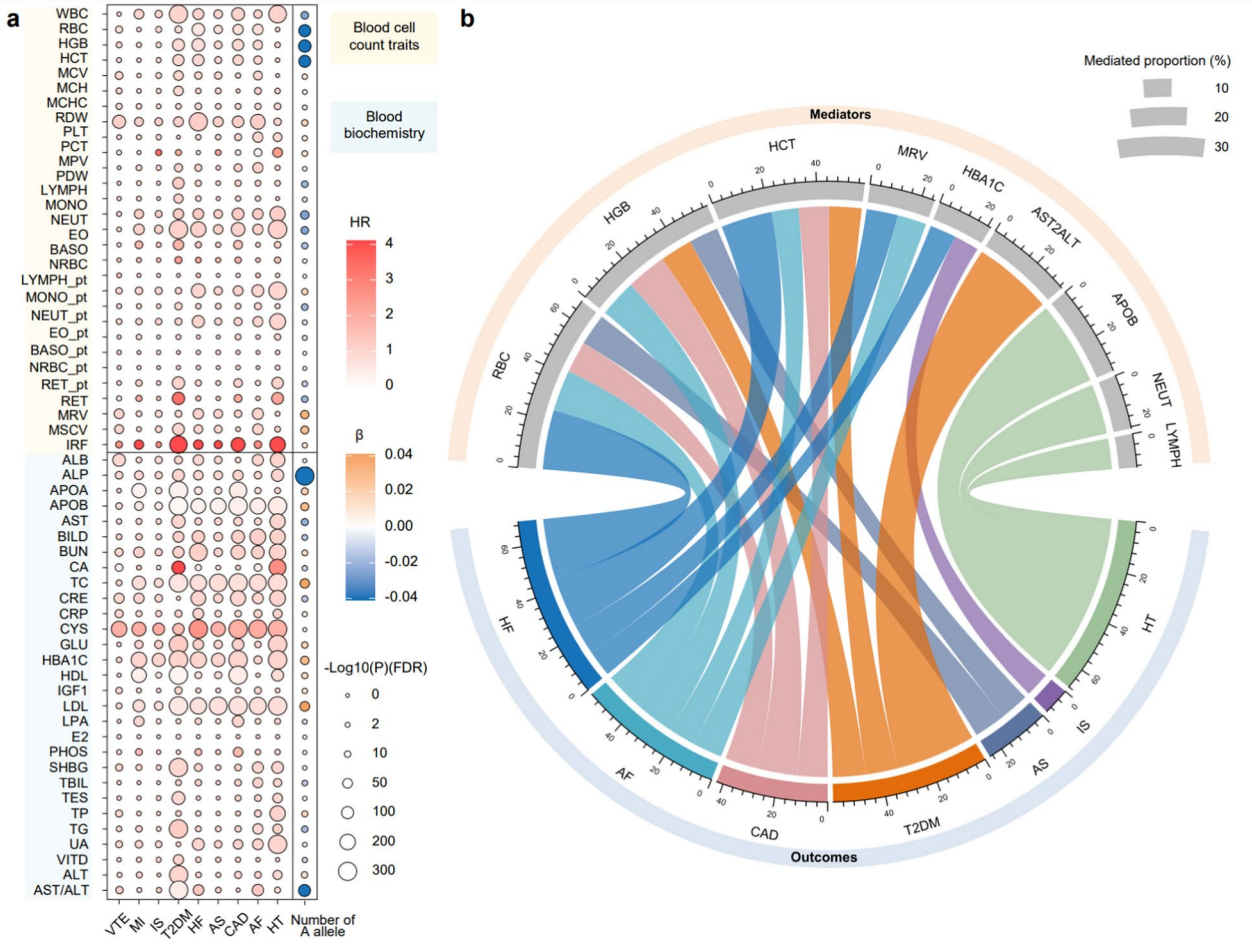


Fig. 4 **a** Bubble heat map shows association of blood cell counts traits and blood biochemistries with cardiometabolic diseases and number of A allele. **b** Chord plot shows mediation analyses results that mediated proportion greater than 10%. *WBC* White blood cell count, *RBC* Red blood cell count, *HGB* Hemoglobin, *HCT* Hematocrit, *MCV* Mean corpuscular volume, *MCH* Mean corpuscular hemoglobin, *MCHC* Mean corpuscular hemoglobin concentration, *RDW* Red cell distribution width, *PLT* Platelet count, *PCT* plateletcrit, *MPV* Mean platelet volume, *PDW* Platelet distribution width, *LYMPH* Lymphocyte count, *MONO* Monocyte count, *NEUT* Neutrophil count, *EO* Eosinophil count, *BASO* Basophil count, *NRBC* Nucleated red blood cells, *LYMPH_pt* Lymphocyte percentage, *MONO_pt* Monocyte percentage, *NEUT_pt* Neutrophil percentage, *EO_pt* Eosinophil percentage, *BASO_pt* Basophil percentage, *NRBC_pt* Nucleated red blood cells percentage, *RET_pt* Reticulocyte percentage, *RET* Reticulocyte count, *MRV* Mean reticulocyte volume, *MSCV* Mean spherical cell volume, *IRF* Immature reticulocyte fraction, *ALB* Albumin, *ALP* Alkaline phosphatase, *APOA* Apolipoprotein A, *APOB* Apolipoprotein B, *AST* Aspartate aminotransferase, *BILD* Direct bilirubin, *BUN* Blood urea nitrogen, *CA* Calcium, *TC* Total cholesterol, *CRE* Creatinine, *CRP* C-reactive protein, *CYS* Cystatin C, *GLU* Glucose, *HBA1C* Glycated hemoglobin A1c, *HDL* High-density lipoprotein cholesterol, *IGF1* Insulin-like growth factor 1, *LDL* Low-density lipoprotein cholesterol, *LPA* Lipoprotein(a), *E2* Estradiol, *PHOS* Phosphate, *SHBG* Sex hormone-binding globulin, *TBIL* Total bilirubin, *TES* Testosterone, *TP* Total protein, *TG* Triglycerides, *UA* Uric acid, *VITD* Vitamin D, *ALT* Alanine aminotransferase, *AST/ALT* AST to ALT ratio. *VTE* Venous thromboembolism, *MI* Myocardial infarction, *IS* Ischemic stroke, *T2DM* Type 2 diabetes mellitus, *HF* Heart failure, *AS* All types of stroke, *CAD* Coronary artery disease, *AF* Atrial fibrillation, *HT* Hypertension

of A allele and T2DM. The associations between the number of A allele and HT were primarily mediated by APOB, NEUT, and LYMPH. No mediation effects were observed for blood cell count traits or blood biochemistries in the associations between the number of A allele and VTE or MI. Figure 4b presents the results for all mediation proportions greater than 10%, with detailed mediation analysis results provided in Supplementary Fig. 5 and Supplementary Table 5.

Discussion

This study is the first to compare the associations between blood group A allele and 15 major CMD under three different inheritance models. We identified significant associations for 9 diseases in both dominant and additive models, whereas the recessive model failed to capture associations with IS and HT. Generally, the additive models have lower P-values, including those for VTE, MI, IS, T2DM, HF, AS, and CAD. Thus, the association between blood group A and these diseases reflects an additive effect, rather than a simple difference in disease

risk among blood groups. Each additional A allele further increases disease risk.

Consistent with previous studies, individuals with blood group A have a higher risk of VTE [21–23]. A study by Höglund et al. demonstrated a stepwise increase in DVT and PE risk from OO to AO and AA [11]. Goumidi et al. conducted comprehensive haplotype and diplotype analyses, highlighting the differences between A1 and A2 alleles within blood group A and the risk of venous thrombosis, with the A1 haplotype predominantly driving the increased VTE risk [24]. In our results, individuals with blood group A had a 38.8% increased risk of VTE compared to those with blood group O, with each A allele increasing the risk by 27.3%. This effect is lower than previously reported, likely due to the inclusion of additional covariates. The additive model's effect sizes (e.g., 27.3% increased VTE risk per A allele) highlight the clinical relevance of blood group A allele dosage. For instance, individuals homozygous for the A allele (AA) would have a cumulative 54.6% higher VTE risk compared to OO carriers, underscoring the need for genotype-specific risk stratification. Similarly, some studies suggest that individuals with blood group A have a higher risk of CAD and MI [25–27], while Reilly et al. [7] identified significant loci associated with MI in the *ABO* gene. A plasma proteome study also found a causal association of ABO antigens with CAD and HF [28]. A cohort study further reported a 10% increased risk of T2DM for individuals with blood group A compared to blood group O (HR = 1.10 [1.02–1.18]), consistent with our dominant model results (HR = 1.058 [1.033–1.084]). Proteome-wide Mendelian randomization also identified a causal association between ABO antigens and T2DM [29]. Regarding blood group A and stroke, prior evidence has been inconsistent. Some studies reported no association between blood group A and stroke [30], while others, consistent with our findings, indicated that the A allele is associated with increased IS risk [31]. Our results further suggest significant dominant and additive associations between blood group A and AS/IS, but no associations with hemorrhagic and transient stroke (including HS and TIA). This study emphasizes that blood group genotypes provide more information than phenotypes when examining associations between blood groups and diseases. Furthermore, the effects of blood group A on disease risk may vary under different inheritance models, with most CMD associations better explained by an additive model.

Based on individual-level protein data, we found that ABO protein levels were significantly and linearly associated only with the risk of VTE, CAD, MI, HF, and T2DM. There were differences in ABO protein levels across different haplotypes, but the differences in disease risk between blood groups may not solely be attributed to antigen levels. Previous Mendelian randomization studies have established an association between genetically determined ABO protein levels and T2DM [32], yet they could not distinguish between antigen A and antigen B. Our study, which focused solely on individuals with blood groups O and A, offers some insight into the effects of antigen A.

Numerous studies have demonstrated that the expression of ABO antigens affects the levels of certain blood cell count traits and blood biochemistries [33, 34]. However, to our knowledge, the mediating factors between blood group A and CMD have not been previously reported. Mediation analysis revealed that some red blood cell count traits simultaneously mediated the associations between the number of A allele and major CVD, including HF, AF, and CAD, suggesting shared mechanisms underlying the associations between blood group A and these diseases. Some studies have found that blood cell counts can predict the onset of CVD [35]. AST/ALT, a liver function marker, mediated 33.01% of the effect of exposure on T2DM, further reinforcing the close link between liver function and T2DM [36, 37]. These findings have not been previously reported. We did not observe any credible mediating effect of these traits on the relationship between A allele and VTE, though this association is likely mediated by coagulation factor VIII and von Willebrand factor [38, 39], which warrants further investigation.

This study's strength lies in its comprehensive evaluation of blood group A allele associations with 15 CMD across recessive, dominant, and additive models. By demonstrating the superiority of the additive model in capturing genetic risk and integrating ABO protein-level data, we provide novel mechanistic insights into how ABO antigens influence disease pathways. Clinically, our findings highlight that blood group A allele dosage is a critical, heritable risk factor for CMD, particularly VTE, CAD, and T2DM. A two-letter formatted blood group testing report would be preferred, because "AA" blood type carries significantly different risk for complex diseases compared with "AO" blood type. Integrating ABO genotyping into existing risk stratification tools could enhance early identification of high-risk individuals, enabling targeted interventions such as intensified biomarker monitoring (e.g., AST/ALT) or preventive therapies.

There are some limitations to our study. Due to current limitations in data availability, we were unable to further analyze the dose–response relationship of different combinations of A1 and A2 alleles using genotype data, which may be addressed in future research. The number of participants with blood types B and AB was significantly lower than those with blood types A and O, so they were not included in the study. Since only the total ABO protein level was measured, we could not determine the individual antigen A levels, and thus were unable to

further identify which specific protein drives the association between ABO protein levels and CMD. There are many potential mediators that were not explored further, such as coagulation factors, which could be addressed in subsequent research. While our analysis identified potential mediators such as AST/ALT and HGB, the lack of established biological pathways for other traits suggests caution in interpreting these associations. Future studies should prioritize validating mediators with plausible mechanisms (e.g., coagulation factors) and leverage multi-omics data (e.g., proteomics) to refine causal inference. Additionally, this study only included individuals from the Caucasian population in the UK Biobank, so the findings may not be generalizable to other populations. Lastly, there may still be residual confounding factors.

Future studies should dissect the interplay between specific *ABO* variants (e.g., A1 vs. A2 subtypes) and environmental/lifestyle factors in modulating CMD risk. Expanding analyses to diverse populations will clarify the generalizability of these associations across ethnicities, addressing current limitations in representation. Additionally, elucidating the roles of unexamined mediators, such as coagulation factors [40], and developing blood group-specific biomarkers (e.g., antigen A quantification assays) could refine risk prediction and therapeutic targeting. These efforts align with precision medicine goals, paving the way for tailored prevention strategies that account for genetic, molecular, and demographic heterogeneity.

Conclusion

In conclusion, our findings reveal that the blood group A allele is associated with an increased risk of multiple CMD, particularly under the additive model. Some blood cell count traits and blood biochemistries play significant mediating roles, providing novel insights into mechanisms between blood group A and these diseases.

Abbreviations

CMD	Cardiometabolic diseases
NCD	Noncommunicable diseases
HF	Heart failure
MI	Myocardial infarction
HT	Hypertension
CVD	Cardiovascular diseases
CAD	Coronary artery disease
PE	Pulmonary embolism
DVT	Deep vein thrombosis
AS	All types of stroke
IS	Ischemic stroke
HS	Hemorrhagic stroke
TIA	Transient ischemic attack
AF	Atrial fibrillation
CM	Cardiomyopathy
AA	Aortic aneurysm
VTE	Venous thromboembolism
PAD	Peripheral vascular disease
T2DM	Type 2 diabetes mellitus
CKD	Chronic kidney disease
BMI	Body mass index
WHR	Waist-to-hip ratio
RCS	Restricted cubic spline
RBC	Red blood cell count
HGB	Hemoglobin
WBC	White blood cell count
NEUT	Neutrophil count
PLT	Platelet count
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
FDR	False discovery rate
EO	Eosinophil
APOB	Apolipoprotein B
TC	Total cholesterol
LDL	Low-density lipoprotein cholesterol
HCT	Hematocrit
AST/ALT	Aspartate aminotransferase to alanine aminotransferase ratio
ALP	Alkaline phosphatase

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12933-025-02669-w .

Supplementary file 1.

Supplementary file 2.

Supplementary file 3.

Acknowledgements

We thank the participants of the UK Biobank cohort for sharing their health-related information. The analysis of this study was supported by Center for Computational Science and Engineering at Southern University of Science and Technology.

Author contributions

Ran Zhao: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing—original draft, Writing—review & editing. Wenyan Xian: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Writing—review & editing. Yihao Ma: Investigation, Validation, Writing—review & editing. Valerio Napolioni: Investigation, Methodology, Writing—review & editing. Patrick W. C. Lau: Methodology, Writing—review & editing. Xiao-Li Tian: Writing—review & editing. Yann Le Guen: Writing—review & editing. Andre Franke: Investigation, Methodology, Project administration, Writing—review & editing. Jie Huang: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing—review & editing.

Funding

This work was supported by National Key Research and Development Program of China [2022YFC3600800].

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

The UK Biobank has received approval from the North West Multicenter Research Ethics Committee (No. 11/NW/0382) and accorded to the Declaration of Helsinki. Informed consent was obtained from all participants. This research was conducted using the UK Biobank resources under application 66137.

Consent for publication

Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]School of Public Health and Emergency Management, Southern University of Science and Technology, Shenzhen, Guangdong Province, China
[2]College of Letters & Science, University of California, Berkeley, CA, USA
[3]School of Biosciences and Veterinary Medicine, University of Camerino, Camerino, Italy
[4]Department of Sport, Physical Education & Health, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China
[5]Aging and Vascular Diseases, Human Aging Research Institute (HARI) and School of Life Science, Nanchang University, Nanchang, Jiangxi, China
[6]Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA, USA
[7]Institute of Clinical Molecular Biology (IKMB), Christian-Albrechts-University of Kiel and University Hospital Schleswig-Holstein (UKSH), Kiel, Germany

Received: 12 January 2025 / Accepted: 26 February 2025
Published online: 10 March 2025

References

1. Ralston J, Nugent R. Toward a broader response to cardiometabolic disease. Nat Med. 2019;25:1644–6.
2. Sattar N, Gill JMR, Alazawi W. Improving prevention strategies for cardiometabolic disease. Nat Med. 2020;26:320–5.
3. Zheng S, Tsao PS, Pan C. Abdominal aortic aneurysm and cardiometabolic traits share strong genetic susceptibility to lipid metabolism and inflammation. Nat Commun. 2024;15:5652.
4. Groot HE, Villegas Sierra LE, Said MA, Lipsic E, Karper JC, van der Harst P. Genetically determined ABO blood group and its associations with health and disease. Arterioscler Thromb Vasc Biol. 2020;40:830–8.
5. Li S, Schooling CM. A phenome-wide association study of ABO blood groups. BMC Med. 2020;18:334.
6. Mega JL, Stitziel NO, Smith JG, Chasman DI, Caulfield M, Devlin JJ, et al. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy. Lancet Lond Engl. 2015;385:2264.
7. Reilly MP, Li M, He J, Ferguson JF, Stylianou IM, Mehta NN, et al. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. Lancet. 2011;377:383–92.
8. Wolpin BM, Kraft P, Gross M, Helzlsouer K, Bueno-de-Mesquita HB, Steplowski E, et al. Pancreatic cancer risk and ABO blood group alleles: results from the pancreatic cancer cohort consortium. Cancer Res. 2010;70:1015–23.
9. Hill WG, Goddard ME, Visscher PM. Data and theory point to mainly additive genetic variance for complex traits. PLOS Genet. 2008;4:e1000008.
10. Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, et al. Genome-wide association studies. Nat Rev Methods Primer. 2021;1:1–21.
11. Höglund J, Karlsson T, Johansson T, Ek WE, Johansson Å. Characterization of the human genotypes and their association to common inflammatory and cardiovascular diseases in the UK biobank. Am J Hematol. 2021;96:1350–62.
12. Chen Y, Chen C, Ke X, Xiong L, Shi Y, Li J, et al. Analysis of circulating cholesterol levels as a mediator of an association between ABO blood group and coronary heart disease. Circ Cardiovasc Genet. 2014;7:43–8.
13. Li S, Xu R-X, Guo Y-L, Zhang Y, Zhu C-G, Sun J, et al. ABO blood group in relation to plasma lipids and proprotein convertase subtilisin/kexin type 9. Nutr Metab Cardiovasc Dis NMCD. 2015;25:411–7.
14. Luo J, Thomassen JQ, Nordestgaard BG, Tybjærg-Hansen A, Frikke-Schmidt R. Neutrophil counts and cardiovascular disease. Eur Heart J. 2023;44:4953–64.
15. Halade GV, Lee DH. Inflammation and resolution signaling in cardiac repair and heart failure. EBioMedicine. 2022;79:103992.
16. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK biobank resource with deep phenotyping and genomic data. Nature. 2018;562:203–9.
17. Yoo JE, Han K, Jung J-H, Hur Y-I, Kim YH, Kim ES, et al. Body mass index, waist circumference and cardiovascular diseases in transitional ages (40 and 66 years). J Cachexia Sarcopenia Muscle. 2023;14:369–81.
18. Lv Y, Zhang Y, Li X, Gao X, Ren Y, Deng L, et al. Body mass index, waist circumference, and mortality in subjects older than 80 years: a Mendelian randomization study. Eur Heart J. 2024;45:2145–54.
19. Rattazzi M, Villalta S, Galliazzo S, Del Pup L, Sponchiado A, Faggin E, et al. High neutrophils and low CD34+ cell count are associated with increased risk of venous thromboembolism. Eur Heart J. 2013;34(suppl 1):P4891–P4891.
20. Lee G, Choi S, Kim K, Yun J, Son JS, Jeong S, et al. Association of hemoglobin concentration and its change with cardiovascular and all-cause mortality. J Am Heart Assoc. 2018;7:e007723.
21. Dentali F, Sironi AP, Ageno W, Turato S, Bonfanti C, Frattini F, et al. Non-O blood type is the commonest genetic risk factor for VTE: results from a meta-analysis of the literature. Semin Thromb Hemost. 2012;38:535–48.
22. Englisch C, Moik F, Nopp S, Raderer M, Pabinger I, Ay C. ABO blood group type and risk of venous thromboembolism in patients with cancer. Blood Adv. 2022;6:6274–81.
23. Vasan SK, Rostgaard K, Majeed A, Ullum H, Titlestad K-E, Pedersen OBV, et al. ABO blood group and risk of thromboembolic and arterial disease. Circulation. 2016;133:1449–57.
24. Goumidi L, Thibord F, Wiggins KL, Li-Gao R, Brown MR, van Hylckama VA, et al. Association between ABO haplotypes and the risk of venous thrombosis: impact on disease risk estimation. Blood. 2021;137:2394–402.
25. Hong X-L, Li Y, Fu G-S, Wu H, Wang Y, Gu C-X, et al. Association of ABO blood groups with the severity of coronary artery disease: a cross-sectional study. J Geriatr Cardiol JGC. 2019;16:701–5.
26. Carpeggiani C. The complexity of ABO in coronary heart disease. Lancet. 2011;377:1492–3.
27. Dai X. ABO blood group predisposes to COVID-19 severity and cardiovascular diseases. Eur J Prev Cardiol. 2020;27:1436–7.
28. Lind L, Mazidi M, Clarke R, Bennett DA, Zheng R. Measured and genetically predicted protein levels and cardiovascular diseases in UK biobank and China kadoorie biobank. Nat Cardiovasc Res. 2024;3:1189–98.
29. Yuan S, Xu F, Li X, Chen J, Zheng J, Mantzoros CS, et al. Plasma proteins and onset of type 2 diabetes and diabetic complications: proteome-wide Mendelian randomization and colocalization analyses. Cell Rep Med. 2023;4: 101174.
30. Zakai NA, Judd SE, Alexander K, McClure LA, Kissela BM, Howard G, et al. ABO blood type and stroke risk: the reasons for geographic and racial differences in stroke study. J Thromb Haemost JTH. 2014;12:564–70.
31. Williams FMK, Carter AM, Hysi PG, Surdulescu G, Hodgkiss D, Soranzo N, et al. Ischemic stroke is associated with the ABO locus: the EuroCLOT study. Ann Neurol. 2013;73:16–31.
32. Ghanbari F, Yazdanpanah N, Yazdanpanah M, Richards JB, Manousaki D. Connecting genomics and proteomics to identify protein biomarkers for adult and youth-onset type 2 diabetes: a two-sample Mendelian randomization study. Diabetes. 2022;71:1324–37.
33. Franchini M, Mengoli C, Capuzzo E, Terenziani I, Bonfanti C, Lippi G. Correlation between ABO blood group, and conventional hematological and metabolic parameters in blood donors. Semin Thromb Hemost. 2015;42:75–86.
34. Nagano N, Tagahara A, Shimada T, Miya M, Tamei N, Muto S, et al. Comparison of serum alkaline phosphatase levels between two measurement methods in chronic hemodialysis patients in Japan: involvement of ABO blood group system and relationship with mortality risk. Clin Exp Nephrol. 2024. https://doi.org/10.1007/s10157-024-02540-4.
35. Madjid M, Fatemi O. Components of the complete blood count as risk predictors for coronary heart disease. Tex Heart Inst J. 2013;40:17–29.
36. Ko S-H, Baeg MK, Han K-D, Ko S-H, Ahn Y-B. Increased liver markers are associated with higher risk of type 2 diabetes. World J Gastroenterol. 2015;21:7478–87.
37. Niu H, Zhou Y. Nonlinear relationship between AST-to-ALT ratio and the incidence of type 2 diabetes mellitus: a follow-up study. Int J Gen Med. 2021;14:8373–82.
38. Ward SE, O'Sullivan JM, O'Donnell JS. The relationship between ABO blood group, von willebrand factor, and primary hemostasis. Blood. 2020;136:2864–74.
39. Azad AB, Akhter QS, Islam MA, Mukta FY, Afroz L, Layla KN, et al. ABO blood groups and their relationship with coagulation factor VIII in healthy adults. J Biosci Med. 2021;9:49–58.

40. Valenti L, Tripodi A, La Mura V, Pelusi S, Bianco C, Scalambrino E, et al. Clinical and genetic determinants of the fatty liver–coagulation balance interplay in individuals with metabolic dysfunction. JHEP Rep. 2022;4:100598.

Publisher's Note