BMC Medicine

A phenome-wide association study of ABO blood groups

Shun Li[1] and C. M. Schooling[1,2]*

Abstract

Background: ABO blood group is associated with differences in lifespan, cardiovascular disease, and some cancers, for reasons which are incompletely understood. To gain sex-specific additional insight about potential mechanisms driving these common conditions for future interventions, we characterized associations of ABO blood group antigen across the phenotype sex-specifically.

Methods: We performed a phenome-wide association study (PheWAS) assessing the association of tag single nucleotide polymorphisms (SNPs) for ABO blood group antigens (O, B, A_1, and A_2) with 3873 phenotypes.

Results: The tag SNP for the O antigen was inversely associated with diseases of the circulatory system (particularly deep vein thrombosis (DVT)), total cholesterol, low-density lipoprotein cholesterol (LDL-C), and ovarian cancer, and positively associated with erythrocyte traits, leukocyte counts, diastolic blood pressure (DBP), and healthy body composition; the tag SNP for the A_1 antigen tended to have associations in reverse to O. Stronger associations were more apparent for men than women for DVT, DBP, leukocyte traits, and some body composition traits, whereas larger effect sizes were found for women than men for some erythrocyte and lipid traits.

Conclusion: Blood group has a complex association with cardiovascular diseases and its major risk factors, including blood pressure and lipids, as well as with blood cell traits and body composition, with some differences by sex. Lower LDL-C may underlie some of the benefits of blood group O, but the complexity of associations with blood group antigen suggests overlooked drivers of common chronic diseases.

Keywords: ABO blood group, Phenome-wide association study, Cardiovascular disease, Sex-specific analysis

Background

ABO blood group is associated with several diseases [1]. People with blood group O have a lower risk of cardiovascular disease (CVD) [2], including myocardial infarction (MI), peripheral vascular disease, cerebral ischemic events [3], and venous thromboembolism [4], as well as of digestive system neoplasms (gastric and pancreatic cancer) [5] and ovarian cancer [6, 7], than people with other blood groups. Reasons for these differences are

not entirely clear, although lower von Willebrand factor (vWF) is thought to be one contributing factor [8]. However, the mechanisms behind these differences have not been entirely elucidated, although they could shed light on the causes of cardiovascular disease, which is increasingly realized to be incompletely understood [9, 10].

Physiologically, blood groups are manifest as many differences [11], including in red blood cell (RBC) traits, vWF, low-density lipoprotein cholesterol (LDL-C), total cholesterol, apolipoprotein E [12], and some inflammatory markers [13], such as soluble E-selectin and intercellular adhesion molecule-1. However, exactly how these differences relate to common diseases and traits has not been systematically examined, although consistent characterization may help identify informative overall patterns or overlooked causes of

* Correspondence: cms1@hku.hk
[1]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 7 Sassoon Rd, Pokfulam, Hong Kong, Special Administrative Region, China
[2]School of Public Health and Health Policy, The City University of New York, 55 W 125 St, New York, NY 10027, USA

common conditions. ABO blood group is also associated with several diseases, such as CVD [2, 4, 14] and some cancers [7, 15, 16], whose incidence differs by sex [17, 18], highlighting the possibility of sex-specific causes and the importance of sex-specific analysis, which has rarely been conducted previously, although is increasingly realized to be relevant [19] given shorter life expectancy in men than women. To address this gap, we conducted a phenome-wide association study (PheWAS), a genotype-to-phenotype approach [12, 20] which can be performed using summary statistics [21], to examine systematically the associations of tag single nucleotide polymorphisms (SNPs) for ABO blood group antigen with a wide range of diseases and related traits, using the largest available genome-wide association studies (GWAS), with sex-specific analysis and validation where possible.

Methods

Blood group antigen
Four tag SNPs were used for the main ABO blood group antigens shown in Table 1 (rs8176746 for B, rs687289 for O, rs507666 for A_1, and rs8176704 for A_2), as previously [22]. rs507666 perfectly marks the A_1 allele [22, 23]. Both rs8176704 [22, 24] and rs8176750 [25] ($r^2 = 0.99$) have been previously used to mark the A_2 allele, but the latter is limited by data availability. rs8176746 distinguishes the B allele from A [26] and was used to mark the B antigen. rs687289 is highly correlated with rs8176719 ($r^2 = 0.97$), which determines the O allele [27, 28], but is limited by data availability. Additional file 1: Table S1 shows how the combination of alleles at these four tag SNPs for each blood group antigen corresponds to each blood group antigen by taking the unique allele [22].

Outcomes: data sources
Large condition-specific GWAS are available from consortia for CVD, including coronary artery disease and stroke, and their risk factors (lipids, blood pressure, diabetes, and glucose metabolism), major cancers, and major contributors to the global burden of disease (including mental health, Alzheimer's disease, and some auto-immune diseases). However, these consortia GWAS do not usually provide sex-specific summary statistics. In contrast, the UK Biobank encompasses sex-specific summary statistics for a wide range of conditions and attributes, but as a

cohort study of half a million people from Great Britain intended to be aged 40 to 69 years when recruited in 2006 and 2010 [29] has a limited number of cases for rarer conditions and for diseases of old age. To be comprehensive, we largely conducted the primary analysis using the UK Biobank and then used other publicly available consortia data for the replication.

UK Biobank summary statistics were available 19,586 different phenotypes encompassing diagnoses, family history, lifestyle, current health status, anthropometrics, physical characteristics, treatment records, biochemical assays, psychological health, and physical measurements provided by Neale Lab [30]. Information concerning the source, original questionnaire or measurement, of these phenotypes is available on the official website of the UK Biobank (https://biobank.ndph.ox.ac.uk/show-case/search.cgi) keyed on the phenotype and ID (Additional file 1: Table S2). To verify previous findings of blood group on ovarian cancer [6, 7], where only self-reported ovarian cancer is available in the UK Biobank, another European ancestry-based publicly available consortium, Ovarian Cancer Association Consortium (OCAC), was included in the primary analysis.

Outcomes: categorization
Subcategories recommended by the UK Biobank were used for inclusion and exclusion. However, the subcategories are largely reflective of the information collection approach. To make the categorization more etiologically coherent, binary outcomes were considered in groups, corresponding to selected International Classification of Diseases (ICD)-9/10 chapters, that might share causes and similar associations with blood group, i.e., circulatory, endocrine, respiratory, neoplasms, digestive, neurological, musculoskeletal, gynecologic and obstetric, hematopoietic, dermatologic, genitourinary, mental health, infectious diseases, sense organs, injuries and poisonings, symptoms, and others.

In addition to the categories above, continuous and categorical ordered phenotypes were classified in groups using the recommended categories for the UK Biobank [31], i.e., blood count, blood biochemistry, and physical measures.

Inclusion and exclusion
To ensure all phenotypes examined came from the same population, non-UK-Biobank and duplicated phenotypes were excluded, unless specified otherwise. To ensure adequate power, binary phenotypes with fewer than 100 cases and continuous and categorical ordered phenotypes with the sample sizes fewer than 10,000 were excluded [32]. The ICD-coded binary phenotypes without main ICD codes or with external causes (codes as V01–Y98, Z00–Z99, or U00–U99) were also excluded. The subcategories

Table 1 ABO blood group antigens and corresponding tag SNPs

Blood group antigen	Tag SNP	Effect allele/non-effect allele
A_1	rs507666	A/G
A_2	rs8176704	A/G
B	rs8176746	T/G
O	rs687289	G/A

of family history, household, summary administration, other sociodemographic factors, smoking, diet-related subcategories, pollution, employment, physical activity, and any phenotypes unrelated to health status were excluded as well. Some self-reported phenotypes were replaced by the corresponding diagnosis phenotypes in the Neale Lab consortium where available. A flowchart showing the selection of phenotypes is given in Fig. 1.
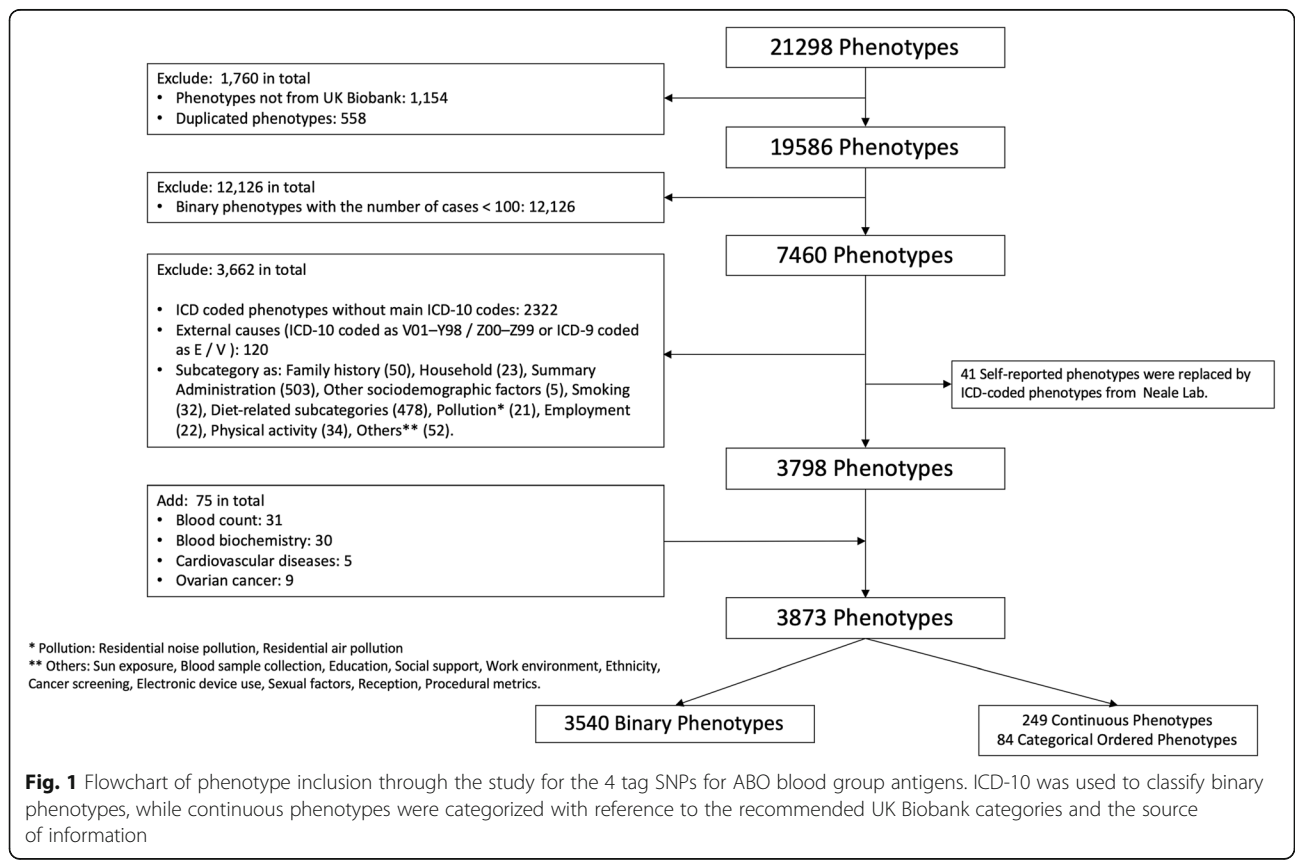
Statistical analysis

Estimates, standard errors, effect alleles, sample sizes, and P values were extracted from the GWAS summary statistics using the MR-base platform [21], followed by harmonization of alleles to ensure estimates corresponded to the same effect allele for each SNP. To account for multiple comparisons, the level of statistical significance was computed as a Bonferroni correction: $P = 0.05/Np/Ns$, where Np is the number of phenotypes tested for each SNP and Ns is the number of SNPs examined [33]. As such, the significance level was 3.2×10^{-6} (0.05/3873/4). Differences by sex were assessed using a two-sided z test [34]. Estimates for binary traits obtained using linear regression were converted to odds ratios (ORs), for presentation, as necessary, using an established approximation [35].

Similar to the data extraction and harmonization process above, replication was performed using the non-UK-Biobank phenotypes excluded from the main analyses to obtain summary statistics for the same phenotypes from different consortia, i.e., the HaemGen consortium, Coronary Artery Disease Genome wide Replication and Meta-analysis plus the Coronary Artery Disease Genetics (CARDIoGRAMplusC4D), the Global Lipids Genetics Consortium (GLGC), the Nuclear Magnetic Resonance (NMR)-GWAS summary statistics, the Genetic Investigation of ANthropometric Traits (GIANT) consortium, the Meta-Analysis of Glucose and Insulin related traits Consortium (MAGIC), the UK Biobank Lung Exome Variant Evaluation (UK BiLEVE) consortium, and the BioBank Japan project. A two-sided z test was used to compare the primary and replication results.

Results

In total, 3540 binary, 84 categorical ordered, and 249 continuous phenotypes were included (Additional file 1: Tables S2-S4). Of the 3540 binary phenotypes considered, 335 were circulatory, 171 endocrine, 179 respiratory, 108 neoplasms, 369 digestive, and the remaining 2378 were neurological, musculoskeletal, gynecologic and obstetric, hematopoietic, dermatologic, genitourinary, mental health,

Fig. 1 Flowchart of phenotype inclusion through the study for the 4 tag SNPs for ABO blood group antigens. ICD-10 was used to classify binary phenotypes, while continuous phenotypes were categorized with reference to the recommended UK Biobank categories and the source of information
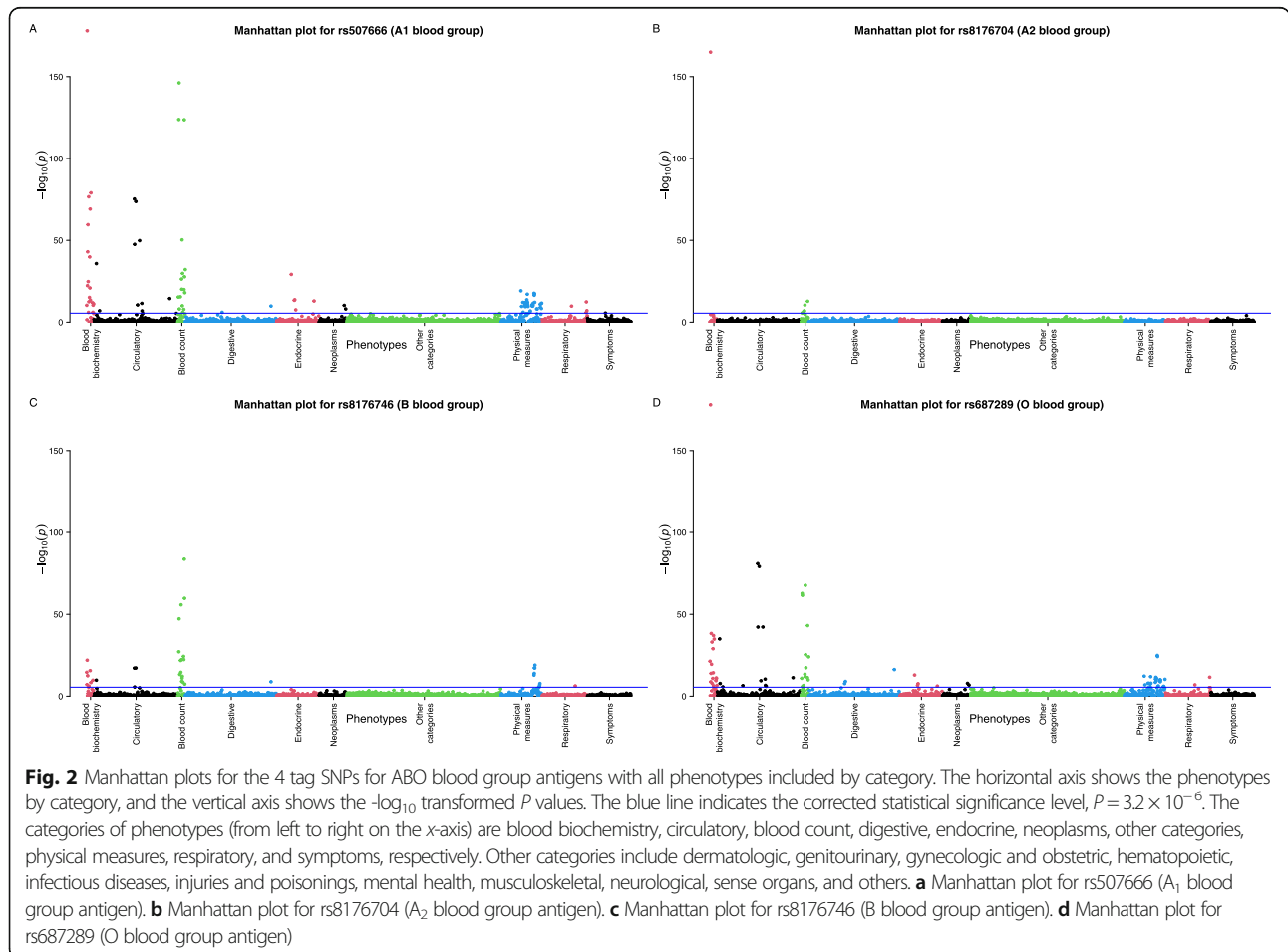
infectious diseases, sense organs, injuries and poisonings, description of symptoms, and others. Figure 2 gives Manhattan plots by tag SNP for each ABO antigen.

For the binary phenotypes considered, ABO antigens were mainly associated with circulatory diseases (Tables 2 and 3). The direction of association differed for A_1 and O antigens for most associations with binary phenotypes (Fig. 3), especially CVD phenotypes. The tag SNP for A_1 antigen was positively associated with diseases of the circulatory system (OR per allele = 1.04), and with higher risks of thrombotic disorders and related treatment (blood clot in the leg, blood clot in the lung, phlebitis and thrombophlebitis, self-reported deep venous thrombosis (DVT), warfarin treatment; ORs = 1.37, 1.47, 1.64, 1.37, 1.22, respectively), but lower risk of high blood pressure (OR = 0.96). Conversely, O antigen was negatively associated with diseases of the circulatory system (OR = 0.97), and with lower risk of these thrombotic disorders and related treatments (ORs = 0.73, 0.70, 0.59, 0.73, 0.85, respectively), but higher risk of high blood pressure (OR = 1.03). Sex-specific analyses were available for high blood pressure and DVT. Directions were the same by sex with no differences by sex for high blood pressure but the effects of A_1 and B on DVT (P = 0.004, 0.023, respectively) were larger in men than women.

O was associated with lower risk of ovarian cancer (OR = 0.93), especially for high-grade serous ovarian cancer (OR = 0.91). Conversely, A_1 was associated with higher risk of ovarian cancer and high-grade serous ovarian cancer (ORs = 1.10, 1.14, respectively). However, no other associations with cancer were found. O was negatively associated with allergic rhinitis/eczema (OR = 0.97), while A_1 was positively associated with this phenotype (OR = 1.04). A_1 was also associated with early diagnosis of allergic rhinitis (effect size per allele = − 0.06). B was associated with snoring (OR = 1.06).

For treatment or medication phenotypes, A_1 was positively associated with cholesterol lowering medication, particularly simvastatin (ORs = 1.07, 1.06, respectively), while O was negatively associated with those medications (ORs = 0.96, 0.97, respectively). Corresponding to higher risk of A_1 antigen on thrombotic disorders, A_1 was positively associated with aspirin (OR = 1.04). Moreover, A_1 was associated with tonsillectomy (OR = 1.05), whereas O was negatively associated with tonsillectomy (OR = 0.96).

Fig. 2 Manhattan plots for the 4 tag SNPs for ABO blood group antigens with all phenotypes included by category. The horizontal axis shows the phenotypes by category, and the vertical axis shows the -log$_{10}$ transformed P values. The blue line indicates the corrected statistical significance level, $P = 3.2 \times 10^{-6}$. The categories of phenotypes (from left to right on the x-axis) are blood biochemistry, circulatory, blood count, digestive, endocrine, neoplasms, other categories, physical measures, respiratory, and symptoms, respectively. Other categories include dermatologic, genitourinary, gynecologic and obstetric, hematopoietic, infectious diseases, injuries and poisonings, mental health, musculoskeletal, neurological, sense organs, and others. **a** Manhattan plot for rs507666 (A_1 blood group antigen). **b** Manhattan plot for rs8176704 (A_2 blood group antigen). **c** Manhattan plot for rs8176746 (B blood group antigen). **d** Manhattan plot for rs687289 (O blood group antigen)

Table 2 Number of significant associations of tag blood group antigens with binary phenotypes by phenotype category

Binary category	Total number	Number significant by blood group antigen			
		A_1	A_2	B	O
Circulatory	335	10	0	4	13
Endocrine	171	5	0	0	4
Respiratory	179	3	0	1	2
Neoplasms	108	2	0	0	2
Digestive	369	2	0	1	3
Symptoms	176	1	0	0	0
Other categories*	2202	0	0	0	0

*Other categories include neurological, musculoskeletal, gynecologic and obstetric, hematopoietic, dermatologic, genitourinary, neoplasms, mental health, infectious diseases, sense organs, injuries and poisonings, and others

Blood group antigen had many associations with blood cell attributes and anthropometrics (Fig. 4). The tag SNP for A_1 was associated with lower RBC traits, including hemoglobin (Hb) concentration, hematocrit (Hct) percentage, and erythrocyte count (effect sizes = − 0.063, − 0.060, − 0.062, respectively), and with lower white blood cell (WBC) traits (counts of total WBC, monocyte, neutrophil, and eosinophil, effect sizes = − 0.037, − 0.045, − 0.035, − 0.018, respectively). Differences for A_1 in total WBC count were greater for men than women (P = 0.037). Similar to A_1, A_2 was negatively associated with RBC traits but with smaller absolute values of effect sizes (− 0.017, − 0.018, − 0.029, respectively). Conversely, B and O were associated with higher RBC traits, with B having larger effect sizes (B: 0.058, 0.044, 0.083, respectively; O: 0.034, 0.035, 0.030, respectively). B was also associated with higher mean corpuscular hemoglobin concentration (MCHC) and reticulocyte count (effect sizes = 0.039, 0.016, respectively), but lower platelet count, plateletcrit, mean corpuscular volume, mean corpuscular hemoglobin, red cell distribution width, and monocyte count (effect sizes = − 0.050, − 0.047, − 0.078, − 0.048, − 0.081, − 0.033, respectively). In contrast to A_1, O was positively associated with those WBC traits (effect sizes = 0.026, 0.044, 0.022, 0.013, respectively). Moreover, A_2 was positively associated with mean corpuscular volume and mean corpuscular hemoglobin (both effect

sizes were 0.024). Directions of association were consistent by sex, but some differences in effect sizes were found (Fig. 5). For A_2 and B, the associations with erythrocyte count, Hb concentration, and Hct (P_{A2} = 0.039, 0.020, 0.017, P_B = 0.002, 0.028, 0.013, respectively) were stronger in women than men. For A_1 on WBC count and neutrophil count, the associations were stronger in men than women (P = 0.037, 0.002, respectively).

ABO blood group was also related to several aspects of blood biochemistry (Fig. 4). A_1 was associated with higher levels of many lipoprotein traits, particularly LDL-C, high-density lipoprotein cholesterol (HDL-C), total cholesterol, apolipoprotein A (ApoA), apolipoprotein B (ApoB) (effect sizes = 0.058, 0.022, 0.058, 0.030, 0.042, respectively), glucose phenotypes, alanine aminotransferase (ALT), and C-reactive protein (effect sizes = 0.031, 0.019, 0.032, respectively), while O was negatively associated with these phenotypes (effect sizes = − 0.032, − 0.012, − 0.030, − 0.015, − 0.020, − 0.021, − 0.023, − 0.024, respectively). A_1 was associated with lower levels of liver biochemistry, i.e., alkaline phosphatase (ALP), aspartate aminotransferase (AST), creatinine, triglycerides, direct bilirubin, and total bilirubin (effect sizes = − 0.322, − 0.048, − 0.018, − 0.020, − 0.026, − 0.021, respectively), whereas O had opposite directions of association with these phenotypes (effect sizes = 0.254, 0.032, 0.012, 0.017, 0.018, 0.016, respectively). Among the above, the

Table 3 The number of significant associations of tag blood group antigens with continuous and categorical ordered phenotypes by category

Continuous and categorical ordered category	Total number	Number significant by blood group antigen			
		A_1	A_2	B	O
Blood count	31	17	6	16	14
Blood biochemistry	30	19	1	8	18
Physical measures	161	35	0	9	22
Respiratory	4	1	0	0	0
Other categories*	107	1	0	0	0

*Other categories include circulatory, endocrine, gynecologic and obstetric, neoplasms, mental health, and sense organs

absolute value of effect sizes for A_1 (effect sizes = 0.057, 0.022, respectively) was almost double than that for O (effect sizes = − 0.032, − 0.012, respectively) with opposite directions of association with LDL-C and HDL-C. B was also related to lower ApoB and LDL-C in both sexes together (effect sizes = − 0.036, − 0.029, respectively) and in women (effect sizes = − 0.048, − 0.037, respectively), but not in men (effect sizes = − 0.023, − 0.020, respectively), with a significantly larger effect size for ApoB among women than men (P = 0.015). B was also associated with higher levels of ALT, ALP, albumin, and creatinine (effect sizes = 0.037, 0.048, 0.028, 0.020, respectively), and lower levels of glycated hemoglobin (HbA1c) (effect size = − 0.039). A_2 was only negatively associated with ALP (effect size = − 0.141). Stronger associations in men than women were evident for the associations of ALP with A_2, B, and O, and in the negative association of B with HbA1c (P values < 0.05), whereas the association of A_1 with HDL-C with was stronger in women than men (P = 0.035).

ABO blood group was also associated with several measures related to the cardiovascular and respiratory systems, including blood pressure and spirometry. A_1 was associated with lower diastolic blood pressure (DBP), pulse rate, forced vital capacity (FVC), and forced expiratory volume in 1 s (FEV_1) (effect sizes = − 0.024, − 0.012, − 0.014, − 0.015, respectively), while O was positively associated with DBP, pulse rate, and FEV_1 (effect sizes = 0.016, 0.012, 0.010, respectively). Stronger associations were found in men than women for the negative association of blood group A_1 antigen with DBP (effect sizes = − 0.028, − 0.014, respectively), and the positive association of blood group O antigen with DBP (effect sizes = 0.022, 0.011, respectively). No significant differences by sex were found for pulse rate, FVC, or FEV_1.

ABO blood group was also associated with physical attributes (Fig. 4), including anthropometry. A_1 was associated with lower heel broadband ultrasound attenuation (BUA), heel quantitative ultrasound index (QUI), heel bone mineral density (BMD), and standing and sitting height (effect sizes = − 0.028, − 0.023, − 0.023, − 0.008, − 0.014, respectively), whereas O was positively associated with BUA, QUI, and BMD (effect sizes = 0.019, 0.016, 0.016, respectively). A_1 was related to higher arm fat percentage and impedance of the arm but lower arm fat-free mass (effect sizes = 0.013, 0.016, − 0.010, respectively), while O had opposite directions of association with these three traits (effect sizes = − 0.011, − 0.016, 0.008, respectively). In addition, B was associated with higher impedance of the arm and leg, and intra-ocular pressure (effect sizes = 0.022, 0.032, 0.044, respectively). Directions of associations were generally similar by sex, although stronger associations in men than women were evident for the negative associations of A_1 antigen with

arm fat-free mass, heel BUA, and BMD (men: − 0.024, − 0.045, − 0.043; women: − 0.009, − 0.018, − 0.010; P values < 0.05), and in the positive associations of O antigen with heel BUA and BMD (men: 0.030, 0.028; women: 0.013, 0.008; P values < 0.05), where associations were not evident among women.

Replications (Additional file 1: Table S5), where available, showed consistent directions of associations as the primary analysis with most replication results having similar effect sizes to the primary ones, particularly for the negative associations of A_1 with white blood cell traits, creatinine, and standing height; of A_2 with Hct and Hb; of B with ApoB, HbA1c, LDL-C, and monocyte traits; and of O with LDL-C, HDL-C, glucose, urea and C-reactive protein, and in the positive associations of A_1 with ApoB, HDL-C, and urea; of A_2 with MCV and MCH; of B with MCHC, Hct, albumin, and creatinine; and of O with Hct, creatinine, and white blood cell traits.
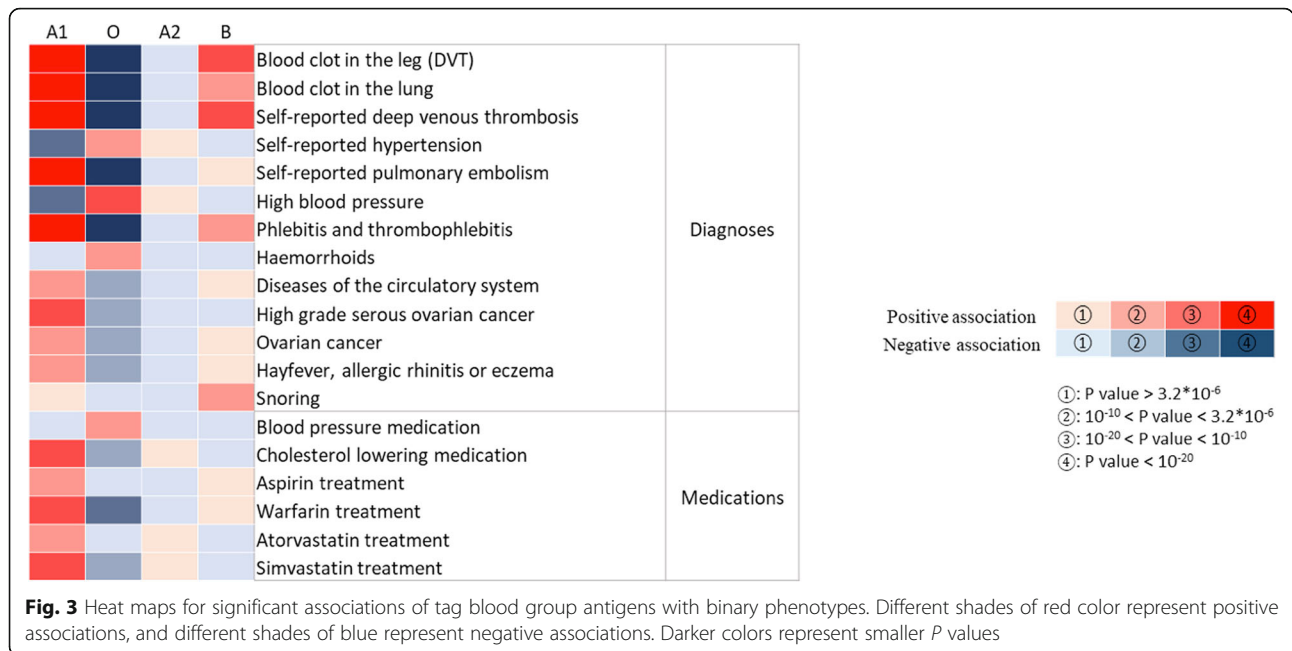
Discussion

Consistent with previous findings, O blood group antigen was associated with lower risk of circulatory diseases, particularly DVT [2], lower risk of ovarian cancer [6, 7], and higher levels of several red blood cell traits [11], whereas the direction of these associations is opposite for blood group A [2, 6, 7]. Our study adds by differentiating the effects of A_1 and A_2 antigens, by giving effects by sex and by subtype of some diseases, and by showing that these differences for O and A_1 antigens extend to white blood cell traits, specific blood biochemistry, and body composition.

As in the relevant existing literature, O blood group was protective and A was harmful for CVD [2, 4, 14]. In addition, only the A_1 subgroup rather than A_2 was related to higher risk of CVD. We also found some indications that men had greater risk than women, for example for the association of A_1 with DVT (P = 0.004).

The protective effects of blood group O and harmful effects of A_1 antigen on ovarian cancer, particularly for high-grade serous ovarian cancer classified as type 2 ovarian cancer clinically and genetically [36], are consistent with previous findings [6, 7]. We also found that associations of O with type 1 ovarian cancers [36] (low grade or low malignant serous/mucinous ovarian cancer, or clear cell ovarian cancer, or endometrioid ovarian cancer) were not significant (P values > 0.05).

In other relevant literature, ABO blood group has shown inconsistent associations with hypertension and blood pressure [37–39]. Here O was also positively associated with diastolic blood pressure in the UK Biobank, consistent with a population in central Asia [40], whereas the inverse association of A with DBP has also been reported in people of African ancestry [41].
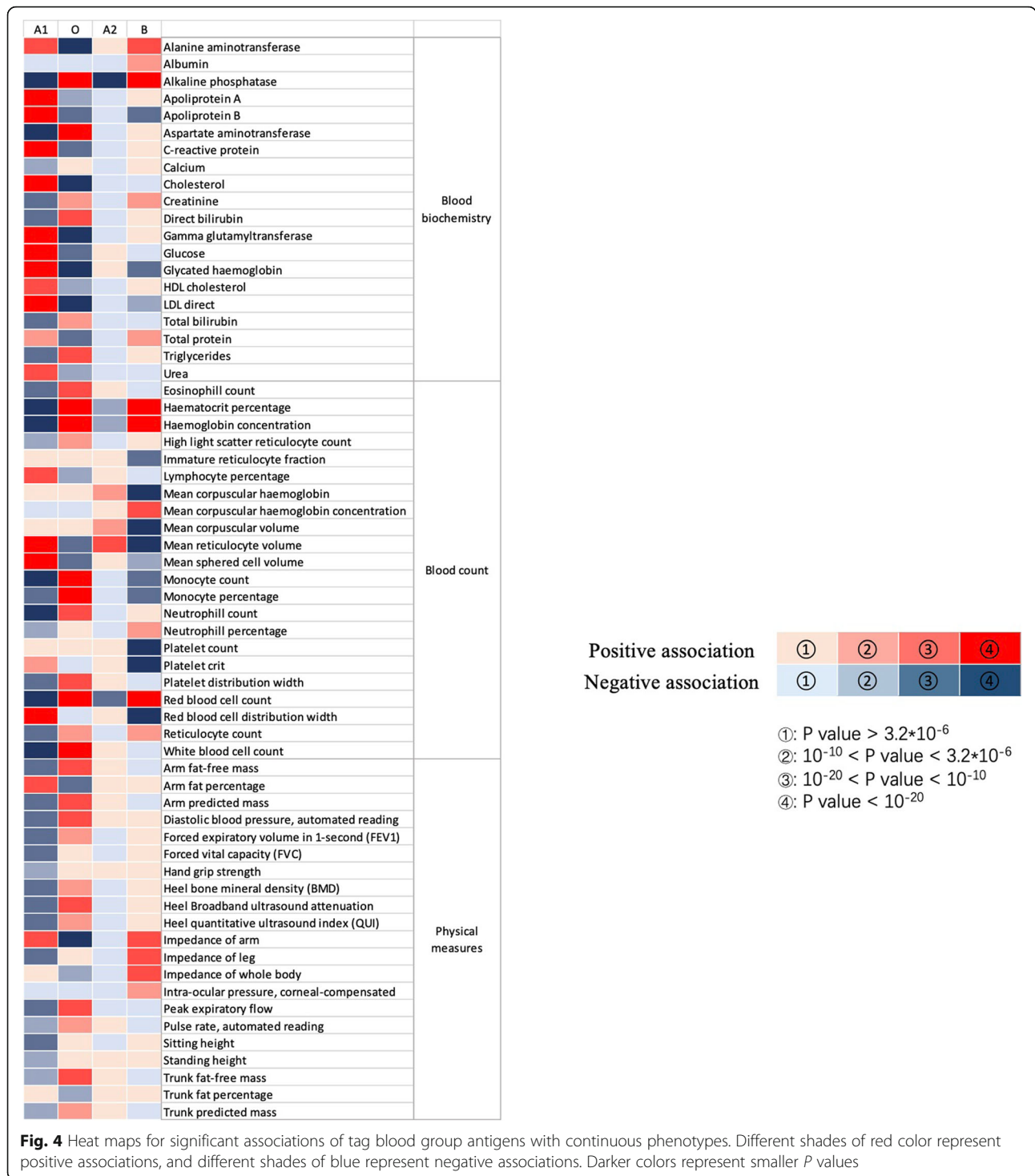
Fig. 3 Heat maps for significant associations of tag blood group antigens with binary phenotypes. Different shades of red color represent positive associations, and different shades of blue represent negative associations. Darker colors represent smaller *P* values

Consistent with previous findings, O was negatively associated with LDL-C [42] and A_1 positively with LDL-C [43, 44]; however, A_2 was not related to LDL-C. B was also related to lower LDL-C, although the magnitude was larger for O. In addition, O was also associated with lower levels of several other lipid traits, including HDL-C, total cholesterol, ApoA, and ApoB. Differences by sex were evident for LDL-C, where associations with A_1 and B antigens were stronger in women than men. Associations in the ABO gene with total cholesterol and LDL-C are consistent with GWAS [45, 46], so some effects of blood group on CVD may be due to LDL-C [47]. Meanwhile, the direction of associations with A_1 was opposite to O, which may be relevant to CVD [48, 49], and has some consistency with previous findings [43, 44, 50]. Blood group was not associated with type 2 diabetes, where previous findings have been inconsistent [51, 52], while we show that O and B antigens are protective for glucose metabolism. Although O was associated with higher levels of some liver biochemistry phenotypes (ALP, AST) but lower ALT, a Mendelian randomization study showed that ALP or AST are not relevant to coronary artery disease, but ALT might be a protective factor for coronary artery disease [53], which is consistent with our findings.

Despite few systematic studies of the relation of blood group with body composition, we found similar associations. Blood type O was associated with both lower LDL-c and higher bone mineral density, consistent with a Mendelian randomization study showing a negative causal association between them [54]. O was also associated with lower body fat and higher muscle mass and

vital capacity, similar to the short-term effects of testosterone [55].

The inverse associations of the tag SNP for A_1 (rs507666) with Hb concentration and Hct, and the positive associations of O antigen with these attributes are consistent with a meta-analysis in Europeans [11]. We provide new evidence that the B antigen is associated with lower platelet count and plateletcrit, and A_1 is related to lower platelet distribution width, possibly because of ABO modulation of platelet surface or platelet function [56, 57]. Few previous studies have considered the relation of ABO blood group antigens with platelets in detail [58].
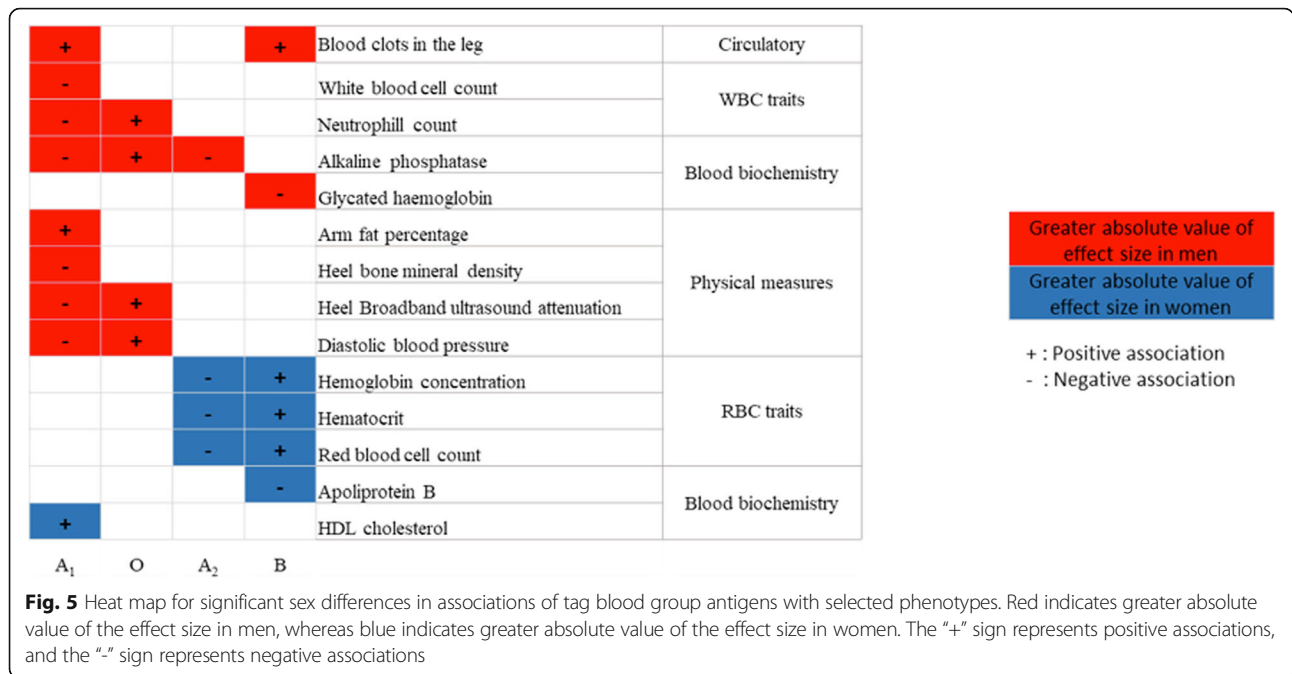
The negative association of O antigen with allergic rhinitis/eczema indicates effects of O antigen on the immune system. Opposite patterns of association for A_1 and O were also found for several white blood cell traits, particularly counts of monocytes, granulocytes, and neutrophils. O antigen is also associated with lower risk of severe acute respiratory syndrome coronavirus (SARS-CoV), where anti-A antibodies might block the receptor-binding process of SARS-CoV spike protein on angiotensin-converting enzyme 2 [59]. A recent study about coronavirus disease 2019 (COVID-19) from China also provides evidence of lower risk in the O blood group, and greater vulnerability to infection for blood group A [60], suggesting linkage between blood group and the immune system. Besides ABO blood group influencing inflammation, it also influences coagulation, for example vWF, although exactly how these factors together mediate effects of blood group on CVD is not well understood. It has been suggested that

Fig. 4 Heat maps for significant associations of tag blood group antigens with continuous phenotypes. Different shades of red color represent positive associations, and different shades of blue represent negative associations. Darker colors represent smaller *P* values

inflammatory processes may be more important for individuals with O blood group while coagulation may be more important for non-O individuals [61]. Elucidating the relative role of inflammation and coagulation in driving the effects of blood group could also inform the development of new interventions for CVD.

Despite conducting a systematic search in the largest available studies, this study has some limitations. First, not all phenotypes of interest were available, most notably vWF which is well-known to be associated with ABO blood group. Second, insufficient cases were available to test some associations, such as with malaria and

+			+	Blood clots in the leg	Circulatory
-				White blood cell count	WBC traits
-	+			Neutrophill count	
-	+	-		Alkaline phosphatase	Blood biochemistry
			-	Glycated haemoglobin	
+				Arm fat percentage	Physical measures
-				Heel bone mineral density	
-	+			Heel Broadband ultrasound attenuation	
-	+			Diastolic blood pressure	
		-	+	Hemoglobin concentration	RBC traits
		-	+	Hematocrit	
		-	+	Red blood cell count	
			-	Apoliprotein B	Blood biochemistry
+				HDL cholesterol	

A₁　　O　　A₂　　B

Fig. 5 Heat map for significant sex differences in associations of tag blood group antigens with selected phenotypes. Red indicates greater absolute value of the effect size in men, whereas blue indicates greater absolute value of the effect size in women. The "+" sign represents positive associations, and the "-" sign represents negative associations

pancreatic cancer. Some phenotypes, such as *H. pylori*-related phenotypes, were classified as secondary ICD codes and therefore were not included in the study. Third, by comparing the tag SNPs for the main ABO blood group antigens, we are providing effects of ABO antigens rather than ABO blood groups. However, presenting effects of ABO antigens provides more information than blood groups [24] and is more relevant to the key questions of how and why ABO blood groups matter, so as to inform the development of interventions. Fourth, the consortium used for most phenotypes was the UK Biobank, which is not representative of the UK population. However, the criteria for internal validity are no confounding and no selection bias. Like all studies in older adults, the UK Biobank may be open to selection bias, from death prior to recruitment, which is likely to generate most bias for conditions subject to competing risk from earlier onset diseases that share etiology with the disease of interest [62]. Such bias may explain why, unlike previous observational studies [2, 3], we did not find an association of ABO blood group with stroke. Fifth, because of data availability, the underlying studies largely relate to populations of European descent. It would be immensely valuable to repeat the analysis when suitable studies more representative of the global population become available, particularly as the geographical distribution of blood groups varies [63]. The underlying studies were adjusted for population stratification, as appropriate, but given much of the information came from the UK Biobank, it is possible that some

confounding due to the geographical distribution of blood groups in Great Britain remains.

Conclusion

After systematic examination, we found that tag SNPs for ABO blood group antigens are associated with CVD, red blood cell traits, white blood cell traits, lipid metabolism, and the musculoskeletal system. Some associations for CVD and the musculoskeletal system were stronger in men, while associations with blood traits and lipid metabolism were stronger in women. Lower LDL-C may underlie some of the benefits of blood group O, but the complexity of associations with blood group suggests overlooked drivers of common chronic diseases.

Supplementary information

Additional file 1: Table S1. ABO blood group antigens and corresponding tag SNPs. **Table S2.** The associations of four tag SNPs for ABO blood group antigens with 3873 phenotypes. **Table S3.** The number of binary phenotypes and characteristics of their cases by disease category. **Table S4.** The number of continuous and categorical ordered phenotypes included and characteristics of their sample sizes in each category. **Table S5.** Comparison between study results and replication results.

Abbreviations

ALP: Alkaline phosphatase; ALT: Alanine aminotransferase; ApoA: Apolipoprotein A; ApoB: Apolipoprotein B; AST: Aspartate aminotransferase; BMD: Bone mineral density; BUA: Broadband ultrasound attenuation; CARDIoGRAMplusC4D: Coronary Artery Disease Genome wide

References
1.   Franchini M, Bonfanti C. Evolutionary aspects of ABO blood group in humans. Clinica Chimica Acta Int J Clin Chem. 2015;444:66–71.
2.   Wu O, Bayoumi N, Vickers MA, Clark P. ABO(H) blood groups and vascular disease: a systematic review and meta-analysis. J Thromb Haemost. 2008; 6(1):62–9.
3.   Sabino Ade P, Ribeiro DD, Domingheti CP, Rios DR, Dusse LM, Carvalho M, Fernandes AP. ABO blood group polymorphisms and risk for ischemic stroke and peripheral arterial disease. Mol Biol Rep. 2014;41(3):1771–7.
4.   Vasan SK, Rostgaard K, Majeed A, Ullum H, Titlestad KE, Pedersen OB, Erikstrup C, Nielsen KR, Melbye M, Nyren O, et al. ABO blood group and risk of thromboembolic and arterial disease: a study of 1.5 million blood donors. Circulation. 2016;133(15):1449–57 discussion 1457.
5.   Zhang BL, He N, Huang YB, Song FJ, Chen KX. ABO blood groups and risk of cancer: a systematic review and meta-analysis. Asian Pacific J Cancer Prevention. 2014;15(11):4643–50.
6.   Gates MA, Wolpin BM, Cramer DW, Hankinson SE, Tworoger SS. ABO blood group and incidence of epithelial ovarian cancer. Int J Cancer. 2011;128(2):482–6.
7.   Henderson J, Seagroatt V, Goldacre M. Ovarian cancer and ABO blood groups. J Epidemiol Community Health. 1993;47(4):287–9.
8.   van Schie MC, van Loon JE, de Maat MP, Leebeek FW. Genetic determinants of von Willebrand factor levels and activity in relation to the risk of cardiovascular disease: a review. J Thromb Haemost. 2011;9(5):899–908.
9.   Ezzati M, Obermeyer Z, Tzoulaki I, Mayosi BM, Elliott P, Leon DA. Contributions of risk factors and medical care to cardiovascular mortality trends. Nat Rev Cardiol. 2015;12(9):508–30.
10. Marmot MG, Syme SL, Kagan A, Kato H, Cohen JB, Belsky J. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: prevalence of coronary and hypertensive heart disease and associated risk factors. Am J Epidemiol. 1975;102(6):514–25.
11. McLachlan S, Giambartolomei C, White J, Charoen P, Wong A, Finan C, Engmann J, Shah T, Hersch M, Podmore C, et al. Replication and characterization of association between ABO SNPs and red blood cell traits by meta-analysis in Europeans. PLoS One. 2016;11(6):e0156914.
12. Arguinano AA, Ndiaye NC, Masson C, Visvikis-Siest S. Pleiotropy of ABO gene: correlation of rs644234 with E-selectin and lipid levels. Clin Chem Lab Med. 2018;56(5):748–54.
13. Sliz E, Kalaoja M, Ahola-Olli A, Raitakari O, Perola M, Salomaa V, Lehtimaki T, Karhu T, Viinamaki H, Salmi M, et al. Genome-wide association study identifies seven novel loci associating with circulating cytokines and cell adhesion molecules in Finns. J Med Genet. 2019;56(9):607–16.
14. Chen Z, Yang SH, Xu H, Li JJ. ABO blood group system and the coronary artery disease: an updated systematic review and meta-analysis. Sci Rep. 2016;6:23250.
15. Liumbruno GM, Franchini M. Hemostasis, cancer, and ABO blood group: the most recent evidence of association. J Thromb Thrombolysis. 2014;38(2): 160–6.
16. Markt SC, Shui IM, Unger RH, Urun Y, Berg CD, Black A, Brennan P, Bueno-de-Mesquita HB, Gapstur SM, Giovannucci E, et al. ABO blood group alleles and prostate cancer risk: results from the breast and prostate cancer cohort consortium (BPC3). Prostate. 2015;75(15):1677–81.
17. Sytkowski PA, D'Agostino RB, Belanger A, Kannel WB. Sex and time trends in cardiovascular disease incidence and mortality: the Framingham Heart Study, 1950-1989. Am J Epidemiol. 1996;143(4):338–50.
18. Timberg G, Rahu M, Gornoi K, Aareleid T, Baburin A. Bladder cancer in Estonia, 1968-1992: incidence, mortality, prevalence and survival. Scand J Urol Nephrol. 1997;31(4):337–42.
19. NIH-Wide Strategic Plan, Fiscal Years 2016–2020: Turning Discovery Into Health [http://www.nih.gov/sites/default/files/about-nih/strategic-plan-fy2016-2020-508.pdf].
20. Ye Z, Mayer J, Ivacic L, Zhou Z, He M, Schrodi SJ, Page D, Brilliant MH, Hebbring SJ. Phenome-wide association studies (PheWASs) for functional variants. Eur J Human Genetics. 2015;23(4):523–9.
21. Gibran Hemani JZ, Kaitlin H Wade, Charles Laurin, Benjamin Elsworth, Stephen Burgess, Jack Bowden, Ryan Langdon, Vanessa Tan, James Yarmolinsky, Hashem A.: The MR-Base platform supports systematic causal inference across the human phenome. In: eLife 2018. 2018.
22. Pare G, Chasman DI, Kellogg M, Zee RY, Rifai N, Badola S, Miletich JP, Ridker PM. Novel association of ABO histo-blood group antigen with soluble ICAM-1: results of a genome-wide association study of 6,578 women. PLoS Genet. 2008;4(7):e1000118.
23. Kiechl S, Paré G, Barbalic M, Qi L, Dupuis J, Dehghan A, Bis JC, Laxton RC, Xiao Q, Bonora E, et al. Association of variation at the ABO locus with circulating levels of soluble intercellular adhesion molecule-1, soluble P-selectin, and soluble E-selectin: a meta-analysis. Circ Cardiovasc Genet. 2011;4(6):681–6.
24. Wolpin BM, Kraft P, Xu M, Steplowski E, Olsson ML, Arslan AA, Bueno-de-Mesquita HB, Gross M, Helzlsouer K, Jacobs EJ, et al. Variant ABO blood group alleles, secretor status, and risk of pancreatic cancer: results from the pancreatic cancer cohort consortium. Cancer Epidemiol Biomarkers Prevention. 2010;19(12):3140–9.
25. Trégouët DA, Heath S, Saut N, Biron-Andreani C, Schved JF, Pernod G, Galan P, Drouet L, Zelenika D, Juhan-Vague I, et al. Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach. Blood. 2009;113(21):5298–303.
26. Rizzato C, Campa D, Pezzilli R, Soucek P, Greenhalf W, Capurso G, Talar-Wojnarowska R, Heller A, Jamroziak K, Khaw KT, et al. ABO blood groups and pancreatic cancer risk and survival: results from the PANcreatic Disease ReseArch (PANDoRA) consortium. Oncol Rep. 2013; 29(4):1637–44.
27. Yamamoto F, Clausen H, White T, Marken J, Hakomori S. Molecular genetic basis of the histo-blood group ABO system. Nature. 1990;345(6272):229–33.
28. Zhang M, Lykke-Andersen S, Zhu B, Xiao W, Hoskins JW, Zhang X, Rost LM, Collins I, Bunt MV, Jia J, et al. Characterising cis-regulatory variation in the transcriptome of histologically normal and tumour-derived pancreatic tissues. Gut. 2018;67(3):521–33.
29. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, 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;12(3):e1001779.

30. GWAS RESULTS [http://www.nealelab.is/uk-biobank/].

31. Category Listings [http://biobank.ndph.ox.ac.uk/showcase/cats.cgi].

32. Delongchamp R, Faramawi MF, Feingold E, Chung D, Abouelenein S. The Association between SNPs and a Quantitative Trait: Power Calculation. Eur J Environ Public Health. 2018;2(2):10.

33. Read RW, Schlauch KA, Elhanan G, Metcalf WJ, Slonim AD, Aweti R, Borkowski R, Grzymski JJ. GWAS and PheWAS of red blood cell components in a Northern Nevadan cohort. PLoS One. 2019;14(6):e0218078.

34. Altman DG, Bland JM. Interaction revisited: the difference between two estimates. BMJ (Clinical research ed). 2003;326(7382):219.

35. Lloyd-Jones LR, Robinson MR, Yang J, Visscher PM. Transformation of summary statistics from linear mixed model association on all-or-none traits to odds ratio. Genetics. 2018;208(4):1397–408.

36. Kurman RJ, Shih I-M. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. Am J Surg Pathol. 2010;34(3):433–43.

37. Chandra T, Gupta A. Association and distribution of hypertension, obesity and ABO blood groups in blood donors. Iran J Ped Hematol Oncol. 2012; 2(4):140–5.

38. Maxwell RD, Maxwell KN. ABO blood groups and hypertension. Br Med J. 1955;2(4932):179–80.

39. 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(3):830–8.

40. Jawed S, Zia S, Tariq S. Frequency of different blood groups and its association with BMI and blood pressure among the female medical students of Faisalabad. JPMA J Pakistan Med Assoc. 2017;67(8):1132–7.

41. Nemesure B, Wu SY, Hennis A, Leske MC. Hypertension, type 2 diabetes, and blood groups in a population of African ancestry. Ethnicity Disease. 2006;16(4):822–9.

42. Li S, Xu RX, Guo YL, Zhang Y, Zhu CG, Sun J, Li JJ. ABO blood group in relation to plasma lipids and proprotein convertase subtilisin/kexin type 9. Nutrition Metabolism Cardiovascular Dis. 2015;25(4):411–7.

43. Etemadi A, Kamangar F, Islami F, Poustchi H, Pourshams A, Brennan P, Boffetta P, Malekzadeh R, Dawsey SM, Abnet CC, et al. Mortality and cancer in relation to ABO blood group phenotypes in the Golestan Cohort Study. BMC Med. 2015;13:8–8.

44. Gong P, Li S, Hu L, Luo S, Li J, Jiang H. Total cholesterol mediates the effect of ABO blood group on coronary heart disease. Zhonghua Xin Xue Guan Bing Za Zhi. 2015;43(5):404–7.

45. Zhou L, He M, Mo Z, Wu C, Yang H, Yu D, Yang X, Zhang X, Wang Y, Sun J, et al. A genome wide association study identifies common variants associated with lipid levels in the Chinese population. PLoS One. 2013;8(12): e82420.

46. Hoffmann TJ, Theusch E, Haldar T, Ranatunga DK, Jorgenson E, Medina MW, Kvale MN, Kwok PY, Schaefer C, Krauss RM, et al. A large electronic-health-record-based genome-wide study of serum lipids. Nat Genet. 2018;50(3): 401–13.

47. Chen Y, Chen C, Ke X, Xiong L, Shi Y, Li J, Tan X, Ye S. Analysis of circulating cholesterol levels as a mediator of an association between ABO blood group and coronary heart disease. Circ Cardiovasc Genet. 2014;7(1):43–8.

48. Campos H, Genest JJ Jr, Blijlevens E, McNamara JR, Jenner JL, Ordovas JM, Wilson PW, Schaefer EJ. Low density lipoprotein particle size and coronary artery disease. Arteriosclerosis Thrombosis. 1992;12(2):187–95.

49. Lawler PR, Akinkuolie AO, Harada P, Glynn RJ, Chasman DI, Ridker PM, Mora S. Residual Risk of Atherosclerotic Cardiovascular Events in Relation to Reductions in Very-Low-Density Lipoproteins. J Am Heart Assoc. 2017;6(12): e007402.

50. 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. 2016; 42(1):75–86.

51. Meo SA, Rouq FA, Suraya F, Zaidi SZ. Association of ABO and Rh blood groups with type 2 diabetes mellitus. Eur Rev Med Pharmacol Sci. 2016; 20(2):237–42.

52. Fagherazzi G, Gusto G, Clavel-Chapelon F, Balkau B, Bonnet F. ABO and Rhesus blood groups and risk of type 2 diabetes: evidence from the large E3N cohort study. Diabetologia. 2015;58(3):519–22.

53. Liu J, Au Yeung SL, Lin SL, Leung GM, Schooling CM. Liver enzymes and risk of ischemic heart disease and type 2 diabetes mellitus: a Mendelian randomization study. Sci Rep. 2016;6:38813.

54. Li GH, Cheung CL, Au PC, Tan KC, Wong IC, Sham PC. Positive effects of low LDL-C and statins on bone mineral density: an integrated epidemiological observation analysis and Mendelian randomization study. Int J Epidemiol. 2019:dyz145.

55. Mohammadi-Shemirani P, Chong M, Pigeyre M, Morton RW, Gerstein HC, Pare G: Clinical benefits and adverse effects of genetically-elevated free testosterone levels: a Mendelian randomization analysis. medRxiv 2019: 19005132. (Preprint).

56. Favaloro EJ, Soltani S, McDonald J, Grezchnik E, Easton L, Favaloro JW. Reassessment of ABO blood group, sex, and age on laboratory parameters used to diagnose von Willebrand disorder: potential influence on the diagnosis vs the potential association with risk of thrombosis. Am J Clin Pathol. 2005;124(6):910–7.

57. Zhong M, Zhang H, Reilly JP, Chrisitie JD, Ishihara M, Kumagai T, Azadi P, Reilly MP. ABO blood group as a model for platelet glycan modification in arterial thrombosis. Arterioscler Thromb Vasc Biol. 2015;35(7):1570–8.

58. Shahbazi S, Mashayekhi A, Fatahi N, Mahdavi MR. Association of ABO and Colton blood group gene polymorphisms with hematological traits variation. Medicine. 2015;94(48):e2144.

59. Guillon P, Clément M, Sébille V, Rivain JG, Chou CF, Ruvoën-Clouet N, Le Pendu J. Inhibition of the interaction between the SARS-CoV spike protein and its cellular receptor by anti-histo-blood group antibodies. Glycobiology. 2008;18(12):1085–93.

60. Zhao J, Yang Y, Huang H-P, Li D, Gu D-F, Lu X-F, Zhang Z, Liu L, Liu T, Liu Y-K, et al. Relationship between the ABO blood group and the COVID-19 susceptibility. medRxiv 2020:2020.2003.2011.20031096.

61. Johansson Å, Alfredsson J, Eriksson N, Wallentin L, Siegbahn A. Genome-wide association study identifies that the ABO blood group system influences Interleukin-10 levels and the risk of clinical events in patients with acute coronary syndrome. PLoS One. 2015;10(11):e0142518.

62. Schooling CM, Lopez PM, Yang Z, Zhao JV, Au Yeung SL, Huang JV. Use of multivariable Mendelian randomization to address biases due to competing risk before recruitment. bioRxiv. 2020:716621. (Preprint).

63. Dean L: ABO Blood Group. In: Medical genetics summaries. edn. Edited by Pratt VM, McLeod HL, Rubinstein WS, Scott SA, Dean LC, Kattman BL, Malheiro AJ. Bethesda: National Center for Biotechnology Information (US); 2012.