openheart

Using genetics to detangle the relationships between red cell distribution width and cardiovascular diseases: a unique role for body mass index

Timothy E Thayer ![ORCID],[1] Shi Huang,[2] Eric Farber-Eger ![ORCID],[3] Joshua A Beckman,[1] Evan L Brittain,[1] Jonathan D Mosley,[1,4] Quinn S Wells[1]

![Check for updates]

[1]Department of Medicine, Division of Cardiovascular Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA
[2]Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee, USA
[3]VICTR, Vanderbilt University Medical Center, Nashville, Tennessee, USA
[4]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA

Correspondence to
Dr Timothy E Thayer; timothy. thayer@vumc.org

ABSTRACT

Objective Red cell distribution width (RDW) is an enigmatic biomarker associated with the presence and severity of multiple cardiovascular diseases (CVDs). It is unclear whether elevated RDW contributes to, results from, or is pleiotropically related to CVDs. We used contemporary genetic techniques to probe for evidence of aetiological associations between RDW, CVDs, and CVD risk factors.

Methods Using an electronic health record (EHR)-based cohort, we built and deployed a genetic risk score (GRS) for RDW to test for shared genetic architecture between RDW and the cardiovascular phenome. We also created GRSs for common CVDs (coronary artery disease, heart failure, atrial fibrillation, peripheral arterial disease, venous thromboembolism) and CVD risk factors (body mass index (BMI), low-density lipoprotein, high-density lipoprotein, systolic blood pressure, diastolic blood pressure, serum triglycerides, estimated glomerular filtration rate, diabetes mellitus) to test each for association with RDW. Significant GRS associations were further interrogated by two-sample Mendelian randomisation (MR). In a separate EHR-based cohort, RDW values from 1-year pre-gastric bypass surgery and 1–2 years post-gastric bypass surgery were compared.

Results In a cohort of 17 937 subjects, there were no significant associations between the RDW GRS and CVDs. Of the CVDs and CVD risk factors, only genetically predicted BMI was associated with RDW. In subsequent analyses, BMI was associated with RDW by multiple MR methods. In subjects undergoing bariatric surgery, RDW decreased postsurgery and followed a linear relationship with BMI change.

Conclusions RDW is unlikely to be aetiologically upstream or downstream of CVDs or CVD risk factors except for BMI. Genetic and clinical association analyses support an aetiological relationship between BMI and RDW.

INTRODUCTION

Red cell distribution width (RDW), a measure of blood cell (RBC) volume heterogeneity, is consistently associated with the presence and severity of diverse cardiovascular diseases (CVDs) including coronary artery disease (CAD), heart failure (HF) and atrial fibrillation (AF).[1–5] However, it is unclear if RBC

Key questions

What is already known about this subject?
► Red cell distribution width (RDW) is associated with the presence and severity of all major cardiovascular diseases via epidemiological studies. Epidemiological studies are subject to confounding whereas genetic approaches provide means to analyse associations independently of many potential confounders.

What does this study add?
► Our findings provide evidence for the currently unsupported (but widely held) belief that RDW is not aetiologically associated with cardiovascular diseases. Also, we found that RDW did not share a genetic architecture with any tested cardiovascular disease nor any tested cardiovascular disease risk factor other than body mass index (BMI). Furthermore, this study provides evidence for an aetiological relationship between BMI and RDW.

How might this impact on clinical practice?
► Through this study RDW can be better interpreted by clinicians as an integrative biomarker with both benign and pathological causes of RDW variation.

volume heterogeneity is mechanistically involved in the pathogenesis of CVDs and, elevated RDW is often presumed to be an epiphenomenon that integrates disparate underlying pathophysiological processes rather than an aetiological factor in disease, per se. However, there have not been formal analyses using genetic associations to assess for evidence for aetiological relationships between RDW, CVDs, and CVD risk factors.

We sought to use genetic risk score (GRS) and Mendelian randomisation (MR) analyses to disentangle the complex relationships between RDW, CVDs, and CVD risk factors. These methods can detect shared genetic architecture and provide evidence for aetiological relationships between traits.

If RBC volume heterogeneity is an aetiological factor in CVDs, genetically predicted RDW should associate with CVDs. Therefore, we constructed GRSs for RDW, CVDs, and CVD risk factors in a well-phenotyped cohort and conducted analyses to test for shared genetic architecture between traits. Among the traits with evidence of shared genetic architecture, we used MR to assess whether there was evidence for an aetiological relationship.

METHODS

Population

The genetic study population included individuals derived from BioVU, a Vanderbilt University Medical Center (VUMC) resource linking deidentified electronic health records to DNA samples, which requires informed consent for subject enrolment. Informed consent was waived for this study given all data had been deidentified. All subjects had previously been genotyped using the Illumina Infinium Multi-Ethnic Genotyping Array (MEGA) platform. The population was limited to subjects of European ancestry because there are few subjects of different continental ancestry in BioVU with the relevant phenotypes. Continental ancestry was determined using principal components. Patients and the public were not involved in the design, conduct, reporting or dissemination plans of this research project.

Clinical data

All clinical data, including demographics, diagnosis codes, procedure codes, body mass index (BMI), and laboratory values, were extracted from the VUMC deidentified electronic health record. Similar to previous studies, included individuals were required to have ≥2 RDW measurements separated by at least 2 years.[2] These criteria were applied to account for single, high-acuity interactions with the healthcare system that could disproportionately contribute outlier RDW measurements. RDW values were extracted from complete blood counts obtained during routine clinical care at VUMC or affiliated outpatient clinics on Sysmex haematology analysers. For each individual, RDW was summarised as median value and unrealistic outliers were excluded as previously done.[2]

Genetic data

All subjects were genotyped using the Illumina MEGA platform. Quality control was conducted using PLINK V.1.9 and included reconciling strand flips, identifying duplicate and related individuals (one of each pair of individual with pi-hat >0.05 was randomly excluded), and applying filters at the sample (missingness rate <1%) and SNP (missingness rate <1%; Hardy-Weinberg $p<10^{-6}$) level.[6] Data were imputed using the Michigan Imputation Server along with the 10/2014 release of the 1000 Genomes phase 3 cosmopolitan reference haplotypes.[7] SNPs with a minor allele frequency >1% after imputation were retained for analyses. Principal components were generated using the SNPRelate package.[8]

GRS construction

We used publicly available summary statistics of published large-scale genome-wide association studies (GWAS) to construct GRSs for RDW, selected CVDs, and CVD risk factors. In the case that there were multiple GWASs available for a given trait, we used the study with the largest number of participants. All GWASs populations were majority, if not exclusively, of European continental ancestry. Specifically, summary statistics were obtained for: RDW,[9] CAD,[10] HF,[11] peripheral arterial disease,[12] venous thromboembolism[13] and AF.[14] Similarly, GRSs were also produced for CVD risk factors including blood pressure (systolic and diastolic),[15] lipids (low-density lipoprotein, high-density lipoprotein, triglycerides),[16] BMI,[17] type 2 diabetes mellitus,[18] and estimated glomerular filtration rate.[19] For clinical disease GWASs, collected summary statistics were genetic variant, effect allele, OR, and p value. For laboratory and blood pressure GWASs, collected summary statistics were genetic variant, effect allele, beta value, and p value. Variants included in the GRSs were required to have a strength of association p value in the original GWAS that was at or below the cutoff for genome-wide significance (5×10^{-8}). Variants were excluded from analyses if in linkage disequilibrium with a lead variant (PLINK clumping performed using cutoffs of $R^2 > 0.1$ or within 250 kb).[20] For each subject, the GRS value was calculated using PRSice2.[21] The RDW GRS was validated against measured median RDW values by ordinal regression. The CVD and CVD risk factor GRSs were each validated against their respective phenotype as determined by phecodes (hierarchical groupings of ICD9 and ICD10 codes) using logistic regression adjusted for age, sex, and principal components 1–3.[22]

GRS and MR analyses

We used a two-stage strategy to detect and characterise relationships between RDW, CVDs, and CVD risk factors. In the first stage, we conducted a series of analyses using a GRS for the predictor of the dependent variable adjusted for age, sex, and principal components 1–3. Significant associations (after Bonferroni correction) in these analyses were interpreted as indicating shared genetic architecture between the two traits. In the second stage, significant associations detected via GRS analyses were selected for formal MR. In MR analyses, each genetic variant is used as an instrumental variable rather than a GRS (which is the sum of the effects of multiple genetic variants). Using each genetic variant as an instrument variable allows for testing for heterogeneity of effect and horizontal pleiotropy.[23]

To evaluate for associations between RDW and the full range of CVDs, we used the validated RDW GRS as the predictor in a targeted phenome scan. The analysis was implemented using the R package 'PheWAS' which creates clinical phenotypes based on hierarchical groupings of ICD9 and ICD10 codes ('phecodes') as previously described.[22] Analyses were limited to 171 phenotypes in the 'circulatory system' group and with at least 60 cases.

The Bonferroni adjusted p value for cardiovascular phenotypes was 0.05/(cardiovascular phenotypes with at least 60 cases). In secondary phenome-wide analyses (PheWAS) significance was considered 0.05/(phenotypes with at least 60 cases). Our PheWAS analyses used logistic regression adjusted for age at time of last ICD code, sex, and principal components 1–3.

The relationship of common CVDs and conventional CVD risk factors to RDW was assessed by testing the associations between a suite of GRSs for CVDs and CVD risk factors with measured RDW. The p value for significance was set to a Bonferroni corrected p value of 0.004 since 13 GRSs were tested against median RDW. Because RDW has a non-normal distribution, we used ordinal regression adjusting for age at time of last RDW measurement, sex, and principal components 1–3.

Significant associations from GRS analyses were carried forward for MR using two-sample inverse-weighted random-effects modelling. The genetic variants used in MR were the same variants used to build the GRSs, but more stringently filtered by distance (1000 kb) and R^2 (>0.01). More stringent filtering was used since MR assumes complete independence of the included genetic variants. Additional MR methods that are more robust to pleiotropy, including simple median, weighted median, and MR-Egger were used to confirm associations. A non-zero intercept in MR-Egger analysis with a p<0.1 was considered evidence of horizontal pleiotropy. Heterogeneity was assessed using Cochran's Q statistic. MR analyses were performed using the MR R package.[24]

RDW pregastric and postgastric bypass

In follow-up analyses, we sought to assess whether changes in BMI were associated with changes in RDW using an orthogonal, complementary approach. We searched Vanderbilt's deidentified electronic health record for individuals that had a BMI of at least $35\,kg/m^2$, underwent gastric bypass surgery (CPT code 43644), and had RDW and BMI measurements in the year preceding surgery and 1–2 years postsurgery. For each individual, median RDW in the year preceding surgery was compared with median RDW 1–2 years postsurgery. The Friedman test was used for comparison of pregastric and postgastric bypass surgery RDW measurements. The relationship between change in BMI and change in RDW was assessed using ordinal regression in subjects who lost between 40 and $0\,kg/m^2$.

RESULTS

Study cohort for genetic analyses

The cohort used for targeted and phenome-wide association analyses consisted of 17 937 subjects (from 33 031 available subjects: 527 removed for non-European ancestry, 2923 removed for relatedness, 11 644 removed for missing RDW values). Clinical characteristics of the cohort are presented in table 1.

Table 1 Genetic cohort characteristics

n	17 937
Sex, male (%)	8091 (45.1)
Age in years, mean (SD)	61.4 (15.7)
Hypertension, n (%)	10 098 (62.4)
Ischaemic heart disease, n (%)	5128 (39.7)
Atrial fibrillation, n (%)	3049 (26.3)
Congestive heart failure, n (%)	2671 (17.6)
Peripheral vascular disease, n (%)	1042 (7.5)
Other venous embolism and thrombosis, n (%)	1076 (8.3)
Hyperlipidaemia, n (%)	8572 (52.5)
Diabetes mellitus, n (%)	3895 (24.7)
Chronic renal failure, n (%)	2323 (16.7)
Obesity, n (%)	2775 (17.6)

RDW GRS validation and associations with the cardiovascular phenome

The RDW GRS was associated with median RDW (R=0.13, p<0.0001) in validation analyses. Targeted cardiovascular PheWAS analyses (including 141 phenotypes with at least 60 cases) no associations between the RDW GRS and any CVD (figure 1, all p>0.01; Bonferroni p value for significance: 4×10^{-4}). In secondary phenome-wide analyses of 1206 phenotypes with at least 60 cases, there were no non-haematological phenotypes associated with the RDW GRS (online supplemental figure 2), including obesity (phecode 278.1, OR 1.01, p=0.65), hyperlipidaemia (phecode 272.1, OR 0.99, p=0.26), diabetes mellitus (phecode 250, OR 1.02, p=0.30), or chronic renal failure (phecode 585.3, OR 1.02, p=0.3). Full tabular results available in online supplemental table 1.

Associations between GRS for CVD and CVD risk factors and RDW

We next tested GRSs for CVD and CVD risk factors for association with median RDW (figure 2). Each GRS was first validated by testing for its association with the phenotype it was constructed to predict (online supplemental table 2), p<0.0001 for all. After adjusting for multiple comparisons, only the BMI GRS was associated with RDW (beta ±SEM: 0.08±0.01 arbitrary units (given ordinal regression was used) per kg/m^2, p<0.0001).

MR analysis of BMI effects on RDW

Two sample inverse-variance weighted random-effects modelling demonstrated that genetically predicted BMI was positively associated with genetically predicted median RDW (figure 3). Similar results were seen in sensitivity analyses using other MR methods (table 2). There was no evidence of heterogeneity (Cochran's Q statistic p=0.23) nor horizontal pleiotropy (MR-Egger intercept 0.001, p=0.91).
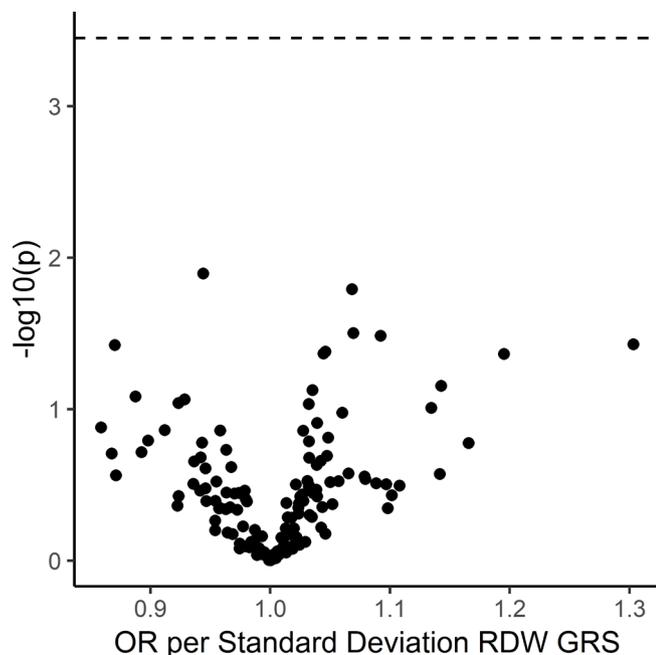
Figure 1 Red cell distribution width (RDW) genetic risk score (GRS) in a targeted cardiovascular phenome-wide association study. Each dot represents a cardiovascular phenotype plotted at the intersection of magnitude of effect of RDW GRS (x-axis) and strength of association (y-axis). No phenotypes approached Bonferroni corrected p value (represented by dashed line) for significant association by logistic regression adjusted for age, sex and principal components 1–3.
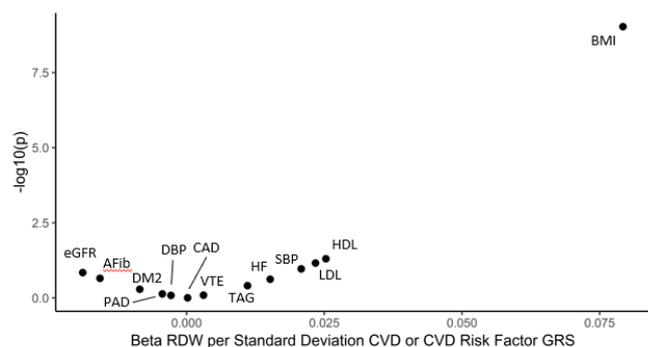

Figure 2 The association of cardiovascular disease (CVD) and CVD risk factor genetic risk scores with median red cell distribution width (RDW) values. Each Genetic Risk Score (GRS) was tested by ordinal regression for association with median RDW values adjusted for age at last RDW measurement, sex and principal components 1–3. Because ordinal regression was used, beta values do not represent actual units. RDW, red cell distribution width. AFib, atrial fibrillation, CAD, coronary artery disease; DBP, diastolic blood pressure; DM, diabetes mellitus; eGFR, estimated glomerular filtration rate; HDL, high density lipoprotein; HF, heart failure; LDL, low density lipoprotein; PAD, peripheral arterial disease; SBP, systolic blood pressure; VTE, venous thromboembolism.
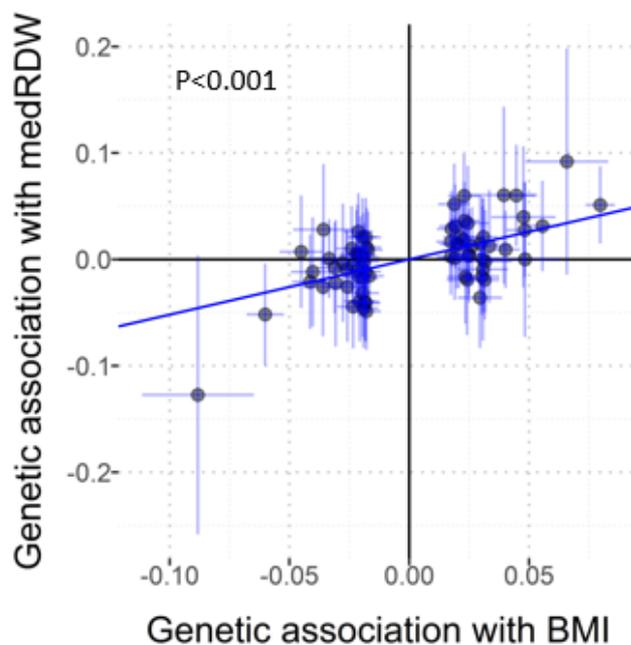

Figure 3 Mendelian randomisation of body mass index (BMI) supports that BMI is aetiologically associated with higher median lifetime RDW value (medRDW). Each point represents a single genetic variant plotted at intersection of its beta value for association with BMI and RDW with SEs. Inverse-weighted regression modelling p value and fit line displayed. Note, because ordinal regression was used to establish RDW ~BMI genetic variant relationship, Y axis does not represent actual units.

RDW decreases post gastric bypass surgery

We identified 1442 subjects who underwent gastric bypass surgery (92% female, 81% white, 17% black, median 47 years old (IQR: 39–55 years old)) and had RDW values available in the year proceeding surgery and 1–2 years postsurgery. RDW decreased from a mean±SD of 14.0%±1.2% presurgery to 13.5%±1.0% postsurgery (figure 4). BMI decreased in this cohort by a mean±SD of 15.9±5.2 kg/m^2. The magnitude of change in BMI after surgery correlated with and the magnitude RDW change (R=0.09, p=0.0007).

Table 2 Outputs from multiple Mendelian randomisation (MR) methods using genetic variants associated with body mass index (BMI) as instrument variables to test for genetic evidence of an aetiological relationship between BMI and red cell distribution width

Method	Estimate	SE	P value
Simple median	0.5	0.13	<0.001
Weighted median	0.55	0.14	<0.001
Inverse weighted median	0.52	0.09	<0.001
MR-egger	0.5	0.22	0.025
(MR-egger intercept)	0.001	0.006	0.91

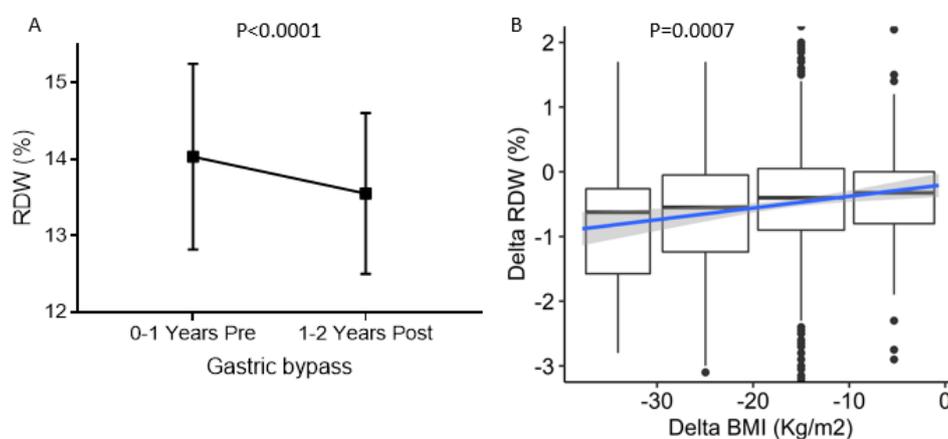All models used random-effects modelling.

Figure 4 Red cell distribution width (RDW) decreases post bypass surgery. (A) Median RDW values obtained in the year preceding surgery were compared with median RDW values 1–2 years after surgery. Comparison made using Friedman test (n=1574). (B) Dose response of delta RDW from delta BMI compared using ordinal regression in subjects who lost between –40 and 0 kg/m^2 (n=1439). BMI, body mass index.

DISCUSSION

We used GRS and MR analyses in a deeply phenotyped cohort to probe for shared genetic architecture and evidence of aetiological relationships between RDW, CVDs, and CVD risk factors. A GRS for RDW was not associated with any CVD or CVD risk factor. Analyses using GRSs for CVDs and common CVD risk factors demonstrated that only BMI shared genetic architecture with RDW. MR analyses provided evidence supporting an aetiological relationship between BMI and RDW. That relationship was further supported by an orthogonal clinical analysis among subjects who underwent gastric bypass surgery that revealed a linear relationship between change in BMI and change in RDW presurgery and postsurgery. Together, these analyses support that RDW is unlikely to be aetiologically upstream or downstream from CVDs or common CVD risk factors with the exception of BMI.

A previous study investigated the aetiological relationships between RDW and disease; Ulrich *et al* used MR to demonstrate that there was no evidence to suggest a role for RDW in the aetiology of pulmonary arterial hypertension.[25] That study used genetically instrumented RDW as a surrogate for iron status and concluded that although iron deficiency is epidemiologically associated with pulmonary arterial hypertension, iron deficiency is unlikely to contribute to the development of the disease. Like pulmonary arterial hypertension, multiple CVDs and CVD risk factors tested in our analyses are also epidemiologically associated with iron deficiency.[26] Similarly to the Ulrich *et al* study, our targeted and phenome-wide PheWAS analyses of the RDW GRS demonstrate that RDW is unlikely to share significant genetic architecture (and by extension, is unlikely to have aetiological relationships with) any non-haematological clinical phenotypes.

RDW and BMI have been linked by multiple studies, though none have provided evidence for a potential aetiological relationship between the two.[27–29] We used the two orthogonal methods of MR and a retrospective clinical analysis of preweight and postweight loss intervention to assess whether BMI may be aetiologically related to elevated RDW. MR analyses implicated BMI with elevated RDW, and this conclusion was further supported by the observed average decrease of 0.5% in RDW 1–2 years postgastric bypass. While both methods have limitations, our confidence in MR analyses is increased by the use of well-validated genetic variants associated with BMI as instruments, concordance of results across multiple MR methods, and lack of evidence for horizontal pleiotropy. Moreover, the gastric bypass analysis was biased toward the null since nearly half of gastric bypass subjects are expected to develop RDW-raising vitamin deficiency anaemia following surgery.[30] Of course, many physiological changes occur postgastricbypass surgery including changes in insulin sensitivity, blood pressure, etc, so BMI change itself cannot be isolated as the mechanism leading to RDW change in this analysis.

There are multiple potential mechanistic links between RDW and BMI. One is that both BMI and RDW have been linked to red blood rheology and rigidity.[31 32] Elevated BMI is known to induce changes in RBC flexibility, bone marrow adipocyte content, and erythropoietin signalling, all of which may affect RDW.[32 33] Additionally, BMI-associated chronic inflammation has been hypothesised to be a mechanistic link between BMI and RBC health; however, recent studies have detected no significant correlation between RDW and inflammatory markers in obese patients.[28 29] Thus, future studies are needed to explore the mechanism(s) of BMI-induced RBC volume variance.

Our study has some limitations. The relative weak and variable predictive power of genetic predictors used could have led to failure to detect true associations due to power. Additionally, all assumptions of MR experiments cannot be empirically tested and can therefore provide evidence for, but not definitive proof of, causality between phenotypes. Given the subjects studied was of European descent, our findings may not be generalisable

to other populations. Smoking, a strong risk factor for CVDs, is a poorly ascertained risk factor via the electronic medical record and thus was not included in our analyses. By not being able to include smoking in our models, we may have missed interactions between smoking and RDW, CVDs, and CVD risk factors.

In conclusion, we found no genetic evidence for aetiological relationships between RDW and CVDs. This finding demonstrates that the aetiology of RDW variation is what is important to CVD biology: not the state of RBC population volume variability itself. Genetic and clinical evidence supports an aetiological relationship between increased BMI and elevated RDW. Further studies are needed to elucidate the mechanism(s) underlying the BMI-RDW relationship and the relevance of elevated RDW to obesity-related CVD risk.

ORCID iDs
Timothy E Thayer http://orcid.org/0000-0001-5317-6171
Eric Farber-Eger http://orcid.org/0000-0003-0281-3796

REFERENCES

1 Perlstein TS, Weuve J, Pfeffer MA, *et al*. Red blood cell distribution width and mortality risk in a community-based prospective cohort. *Arch Intern Med* 2009;169:588–94.
2 Thayer TE, Huang S, Levinson RT. Unbiased phenome-wide association studies of red cell distribution width identifies key associations with pulmonary hypertension. *Ann Am Thorac Soc* 2019;16:589–98.
3 Tonelli M, Sacks F, Arnold M, *et al*. Relation between red blood cell distribution width and cardiovascular event rate in people with coronary disease. *Circulation* 2008;117:163–8.
4 Felker GM, Allen LA, Pocock SJ, *et al*. Red cell distribution width as a novel prognostic marker in heart failure: data from the CHARM program and the duke databank. *J Am Coll Cardiol* 2007;50:40–7.
5 Adamsson Eryd S, Borné Y, Melander O, *et al*. Red blood cell distribution width is associated with incidence of atrial fibrillation. *J Intern Med* 2014;275:84–92.
6 Purcell S, Neale B, Todd-Brown K, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–75.
7 Das S, Forer L, Schönherr S, *et al*. Next-generation genotype imputation service and methods. *Nat Genet* 2016;48:1284–7.
8 Zheng X, Levine D, Shen J, *et al*. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 2012;28:3326–8.
9 Pilling LC, Atkins JL, Duff MO, *et al*. Red blood cell distribution width: genetic evidence for aging pathways in 116,666 volunteers. *PLoS One* 2017;12:e0185083.
10 Nikpay M, Goel A, Won H-H, *et al*. A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;47:1121–30.
11 Shah S, Henry A, Roselli C, *et al*. Genome-wide association and mendelian randomisation analysis provide insights into the pathogenesis of heart failure. *Nat Commun* 2020;11:163.
12 Klarin D, Lynch J, Aragam K, *et al*. Genome-wide association study of peripheral artery disease in the million veteran program. *Nat Med* 2019;25:1274–9.
13 Klarin D, Busenkell E, Judy R, *et al*. Genome-wide association analysis of venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular disease. *Nat Genet* 2019;51:1574–9.
14 Roselli C, Chaffin MD, Weng L-C, *et al*. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet* 2018;50:1225–33.
15 Evangelou E, Warren HR, Mosen-Ansorena D, *et al*. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet* 2018;50:1412–25.
16 Klarin D, Damrauer SM, Cho K, *et al*. Genetics of blood lipids among ~300,000 multi-ethnic participants of the million veteran program. *Nat Genet* 2018;50:1514–23.
17 Yengo L, Sidorenko J, Kemper KE, *et al*. Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum Mol Genet* 2018;27:3641–9.
18 Xue A, Wu Y, Zhu Z, *et al*. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nat Commun* 2018;9:2941.
19 Hellwege JN, Velez Edwards DR, Giri A, *et al*. Mapping eGFR loci to the renal transcriptome and phenome in the Va million veteran program. *Nat Commun* 2019;10:3842.
20 Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc* 2020;15:2759–72.
21 Choi SW, O'Reilly PF. PRSice-2: polygenic risk score software for biobank-scale data. *Gigascience* 2019;8. doi:10.1093/gigascience/giz082. [Epub ahead of print: 01 07 2019].
22 Denny JC, Ritchie MD, Basford MA, *et al*. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics* 2010;26:1205–10.
23 Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in mendelian randomization studies. *Hum Mol Genet* 2018;27:R195–208.
24 Yavorska OO, Burgess S. Mendelian randomization: an R package for performing mendelian randomization analyses using summarized data. *Int J Epidemiol* 2017;46:1734–9.
25 Ulrich A, Wharton J, Thayer TE, *et al*. Mendelian randomisation analysis of red cell distribution width in pulmonary arterial hypertension. *Eur Respir J* 2020;55. doi:10.1183/13993003.01486-2019. [Epub ahead of print: 12 02 2020].
26 von Haehling S, Jankowska EA, van Veldhuisen DJ, *et al*. Iron deficiency and cardiovascular disease. *Nat Rev Cardiol* 2015;12:659–69.
27 Kim J, Im J-S, Choi CH, *et al*. The association between red blood cell distribution width and sarcopenia in U.S. adults. *Sci Rep* 2018;8:11484–84.
28 Fujita B, Strodthoff D, Fritzenwanger M, *et al*. Altered red blood cell distribution width in overweight adolescents and its association with markers of inflammation. *Pediatr Obes* 2013;8:385–91.
29 Vayá A, Alis R, Hernandez-Mijares A, *et al*. Red blood cell distribution width is not related with inflammatory parameters in morbidly obese patients. *Clin Biochem* 2014;47:464–6.
30 Vargas-Ruiz AG, Hernández-Rivera G, Herrera MF. Prevalence of iron, folate, and vitamin B12 deficiency anemia after laparoscopic Roux-en-Y gastric bypass. *Obes Surg* 2008;18:288–93.
31 Valadão Cardoso A. An experimental erythrocyte rigidity index (RI) and its correlations with transcranial doppler velocities (TAMMV), gosling pulsatility index Pi, hematocrit, hemoglobin concentration and red cell distribution width (RDW). *PLoS One* 2020;15:e0229105–e05.
32 Wiewiora M, Piecuch J, Glück M, *et al*. The effects of weight loss surgery on blood rheology in severely obese patients. *Surg Obes Relat Dis* 2015;11:1307–14.
33 Suresh S, Alvarez JC, Dey S, *et al*. Erythropoietin-induced changes in bone and bone marrow in mouse models of diet-induced obesity. *Int J Mol Sci* 2020;21:1657.