ARTICLE

A Mendelian randomization-based exploration of red blood cell distribution width and mean corpuscular volume with risk of hemorrhagic strokes

Jundong Liu,[1] Elizabeth L Chou,[2] Kui Kai Lau,[3,4] Peter Yat Ming Woo,[5] Tsz Kin Wan,[6] Ruixuan Huang,[6] and Kei Hang Katie Chan[1,6,7,8,*]

Summary

Red blood cell distribution width (RCDW) and mean corpuscular volume (MCV) are associated with different risk factors for hemorrhagic stroke. However, whether RCDW and MCV are causally related to hemorrhagic stroke remains poorly understood. Therefore, we explored the causality between RCDW/MCV and nontraumatic hemorrhagic strokes using Mendelian randomization (MR) methods. We extracted exposure and outcome summary statistics from the UK Biobank and FinnGen. We evaluated the causality of RCDW/MCV on four outcomes (subarachnoid hemorrhage [SAH], intracerebral hemorrhage [ICH], nontraumatic intracranial hemorrhage [nITH], and a combination of SAH, cerebral aneurysm, and aneurysm operations) using univariable MR (UMR) and multivariable MR (MVMR). We further performed colocalization and mediation analyses. UMR and MVMR revealed that higher genetically predicted MCV is protective of ICH (UMR: odds ratio [OR] = 0.89 [0.8–0.99], p = 0.036; MVMR: OR = 0.87 [0.78–0.98], p = 0.021) and nITH (UMR: OR = 0.89 [0.82–0.97], p = 0.005; MVMR: OR = 0.88 [0.8–0.96], p = 0.004). There were no strong causal associations between RCDW/MCV and any other outcome. Colocalization analysis revealed a shared causal variant between MCV and ICH; it was not reported to be associated with ICH. Proportion mediated via diastolic blood pressure was 3.1% (0.1%,14.3%) in ICH and 3.4% (0.2%,15.8%) in nITH. The study constitutes the first MR analysis on whether genetically elevated RCDW and MCV affect the risk of hemorrhagic strokes. UMR, MVMR, and mediation analysis revealed that MCV is a protective factor for ICH and nITH, which may inform new insights into the treatments for hemorrhagic strokes.

Introduction

Hemorrhagic stroke results from bleeding into/around the brain when a weakened vessel ruptures.[1,2] The blood accumulates to compress the surrounding brain cells, leading to brain tissue damage and neurological deficits.[1,3] According to American Stroke Association, hemorrhagic strokes occupy about 13% of all strokes consisting of intracerebral hemorrhage (ICH)—bleeding into the brain parenchyma—and subarachnoid hemorrhage (SAH)—bleeding into the subarachnoid space.[1] Mortality and morbidity are high for hemorrhagic stroke while identifying risk factors of the hemorrhagic stroke helps detect and modify the risk, potentially reducing the odds of death or disability of the fatal disease.[4,5] Identifying risk factors of the hemorrhagic stroke helps detect and modify the risk, potentially reducing the odds of death or disability of the fatal disease.[6]

Mean corpuscular volume (MCV) describes the average size and volume of erythrocytes circulating in the bloodstream.[7] It is calculated by dividing the hematocrit by the concentration of erythrocytes (hematocrit [%] × 10/ red blood cell count [millions/mm^3 blood]). Red blood cell distribution width (RCDW) is defined as one standard deviation (SD) of erythrocyte volume divided by the MCV (SD/MCV × 100) and is used to measure the heterogeneity of red blood cell size.[8] Therefore, RCDW and MCV are correlated biologically. RCDW could be a strong predictor of the risk and prognosis of ischemic stroke.[8] Similarly, MCV can be used to calculate the mortality and morbidity rates of ischemic stroke.[9,10] RCDW and MCV are often used to diagnose hematological system diseases, such as iron-deficiency anemia and bone marrow dysfunction.[11,12] They help determine the risk of a cardiovascular event occurring after surgery or blood transfusions.[13] Increases in RCDW can be observed more easily as MCV tends to remain at low levels[8]; RCDW and stroke are generally associated with a group of patients with low MCV.[11]

However, studies on the causal effects of RCDW and MCV on stroke are limited. Recently, a Mendelian randomization (MR) study proposed that RCDW and MCV are causally associated with small-vessel and cardioembolic stroke, respectively.[14] Whether these two metrics are causally related to hemorrhagic stroke remains to be poorly understood. Numerous studies suggest RCDW and MCV are associated with the risk factors for

[1]Department of Biomedical Sciences, City University of Hong Kong, Hong Kong SAR, China; [2]Massachusetts General Hospital, Boston, MA, USA; [3]Division of Neurology, Department of Medicine, The University of Hong Kong, Hong Kong SAR, China; [4]The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong SAR, China; [5]Department of Neurosurgery, Kwong Wah Hospital, Hong Kong SAR, China; [6]Department of Electrical Engineering, City University of Hong Kong, Hong Kong SAR, China; [7]Department of Epidemiology, Centre for Global Cardiometabolic Health, Brown University, RI, USA
[8]Lead contact
*Correspondence: katie.kh.chan@cityu.edu.hk
https://doi.org/10.1016/j.xhgg.2022.100135.

hemorrhagic stroke.[15–27] Higher MCV levels had lower body mass index, hemoglobin A1c, cholesterol, triglyceride, uric acid levels, diabetes mellitus, dyslipidemia, hypertension, and metabolic syndrome.[28–31] Risk factors for hemorrhagic stroke include hypertension, smoking, sleep apnea, cocaine use, alcohol abuse, and clotting problems due to blood disorders or medicine (from Winchester Hospital). For hypertension, clinical studies propose that hypertensives have lower MCV than normotensives, while epidemiological studies suggest that they have higher MCVs or no relation exists.[16,20] Among chronic kidney disease patients, diabetes mellitus and hypertension were more common in the low-MCV group.[28] A study on erythrocytotic transgenic mice demonstrated by experiment that the blood viscosity regulation is an important adaptive mechanism to excessive erythrocytosis.[32] Blood decreases its viscosity by greater shear stress primarily from the high deformability of red blood cells.[33] This can be achieved by increasing the proportion of juvenile flexible erythrocytes, which is more deformable with higher MCV than older ones.[32,33] Therefore, higher MCV correlates with lower blood viscosity, while blood viscosity was reported to be positively associated with systolic blood pressure, diastolic blood pressure, and mean arterial pressure.[34] Also, RCDW was reported to increase venous thrombosis and diabetes mellitus,[23,35] and be positively associated with blood pressure.[24,25] Hypertensives are prone to carotid artery atherosclerosis if they have high RCDW.[26] For smoking, continuous cigarette smoking has severe undesirable influences on hematological indexes, including MCV, hemoglobin, and red blood cells count, leading to the development of cardiovascular diseases. The MCV level in smokers is higher than in nonsmokers.[27] For clotting disorder, blood clot risk can result from low iron levels,[17] while iron deficiency, best characterized by anisocytosis (high RCDW), tends to cause microcytic anemia (low MCV anemia) in children.[18] For alcohol use disorder, a study also found MCV increases in blood after 4–8 weeks of excessive alcohol intake.[19] A previous study also reported that RCDW and D-dimer were associated with cerebral venous thrombosis, a determinant of ICH.[15] Thus, RCDW and MCV have been shown to be associated with different risk factors of hemorrhagic stroke.

Based on the above associations, we postulated that the RCDW/MCV is causally associated with hemorrhagic strokes, meditated by some causal pathways. Therefore, this study explored the causal link between RCDW/MCV and hemorrhagic stroke using MR methods with a focus on nontraumatic hemorrhagic strokes.

Materials and methods

Data source

All summary-level data can be obtained from the publicly available Medical Research Council Integrative Epidemiology Unit OpenGWAS database.[36] The exposure and outcome data are from non-overlapping European populations. The two exposures (MCV and RCDW) are derived from the UK Biobank database, where MCV is extracted from the second wave of a genome-wide association study (GWAS) by the Neale Lab and RCDW from another GWAS.[37] The four hemorrhagic stroke outcomes are from the FinnGen consortium; we chose the R5 release of GWAS results. These outcomes are SAH, ICH, nontraumatic intracranial hemorrhage (nITH), and a combination of SAH, unruptured cerebral aneurysm, and aneurysm operations SAH (AOS). Aneurysm operations in AOS are endovascular or surgical operations to intracerebral aneurysms, coded with Nordic Medico-Statistical Committee Operations classification of surgical procedures[38]: AAC0 [0–5], AAC1[0–5], AAC99, and AAL00. We involved AOS because SAH is one of the endpoints. Based on the definition of the endpoint in FinnGen, we considered that nITH excludes other brain hemorrhages—nITH is equivalent to hemorrhagic stroke. To test the mediation effect, we included blood disorders (elevated erythrocyte sedimentation rate and abnormality of plasma viscosity) from FinnGen and blood pressure (SBP, systolic blood pressure; DBP, diastolic blood pressure) from a combined GWAS from the UK Biobank and the International Consortium of Blood Pressure.[39] Table 1 presents the description of GWAS summary data used, the name of each trait, the specific endpoints coded with the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10), the sample size, and the mean age of the first event; more information is recorded across Tables S1, S2, and S3. No ethical approval was required for this study as we used publicly available data. The MR study was approved by the City University of Hong Kong Research Committee (JCC approval no.: jcc2122ay006).

Phenotypic correlations

First, we estimated the phenotypic correlations between the traits using the GWAS summary statistics with the PhenoSpD R toolkit.[40] However, PhenoSpD assumes the GWAS samples must overlap substantially to compute the phenotypic correlation effectively; otherwise, the correlation will attenuate toward zero as overlap decreases.[40] Our exposures and outcomes are from two cohorts (UK Biobank and FinnGen). Therefore, the correlations between the exposure and outcome might unreliably approach zero. This may be why we observed the correlations between the exposures or between the outcomes, respectively.

Univariable and multivariable MR

The principles of two-sample univariable MR (UMR) and multivariable MR (MVMR) have been described elsewhere.[41–44] We adopted single-nucleotide polymorphisms (SNPs) as instruments from non-overlapping individuals of the same ethnicity for RCDW/ MCV and outcomes. As RCDW and MCV are correlated biologically, they might share genetic variants associated with the risk of stroke, that is, they are pleiotropic.[45] We adjusted MCV and RCDW in MVMR to control the pleiotropic SNPs and assess the causality between exposures and outcomes. In general, the UMR is performed under three assumptions (Figure 1A): (1) SNPs correlated with the exposure (RCDW or MCV); (2) lack of correlated horizontal pleiotropy (e.g., SNPs are not associated with other confounders between MCV/RCDW and stroke); and (3) lack of uncorrelated horizontal pleiotropy (e.g., SNPs only affect stroke through the exposure of interest).[45,46] Similar assumptions for MVMR are proposed,[43] as listed: (1) each SNP is associated with RCDW

Table 1. Description of genome-wide association study summary data for exposures and outcomes

Trait	ICD-10	Sample size	Event age (years)
Mean corpuscular volume (fL)	–	350,473	57
Red blood cell distribution width (10^{12} cells/L)	–	116,666	56.5[a]
Nontraumatic intracranial hemorrhage	I6[0–1]	2,794 cases	61.0
		203,068 controls	
Intracerebral hemorrhage	I61	1,687 cases	65.6
		201,146 controls	
Subarachnoid hemorrhage	I60	1,338 cases	54.9
		201,230 controls	
Subarachnoid hemorrhage, unruptured cerebral aneurysm, and aneurysm operations	I60, I671	2,127 cases	56.0
		203,068 controls	
Elevated erythrocyte sedimentation rate and abnormality of plasma viscosity	R70	1,093	66.5
		212,004	
Systolic blood pressure (mmHg)	–	757,601	N//A
Diastolic blood pressure (mmHg)	–	757,601	NA

[a]Mean age is inferred from Hewitt et al.[89] ICD-10, International Statistical Classification of Diseases and Related Health Problems, 10th Revision. Event age is the mean age of the first event.

or/and MCV; (2) each variant is not associated with a confounder other than RCDW or MCV; (3) the variant is conditionally independent of stroke given the exposures (Figure 1B).

We preprocessed the data to perform two-sample UMR and MVMR, respectively. For UMR, we performed linkage disequilibrium clumping (p = 5×10^{-8}, $r^2 = 0.01$ and kb = 10,000), allowed for proxy SNPs, corrected non-palindromic strands, and excluded palindromic and ambiguous SNPs via the TwoSampleMR R package.[47] Then, we excluded outliers, computed MR estimates, and performed the sensitivity analyses. For MVMR, we also extracted the SNPs using the TwoSampleMR package.[47] In clumping, we first kept the SNPs that were both present in two traits. Then, we selected those with p < 5×10^{-8} in either of the exposures. We performed clumping and looked for proxies by $r^2 = 0.01$ and kb = 10,000, and harmonized the SNPs on the same strand. Finally, we extracted outcome SNPs by a 0.01 MAF threshold to infer palindromic SNPs; we used default settings for the other parameters.

Model selection framework

According to the Rücker model selection framework, we selected the inverse-variance weighted (IVW) analysis under a fixed effect model as the primary MR estimation when MR assumptions are satisfied without pleiotropy or heterogeneity effects.[48] According to the MVMR package framework,[49] we prioritized IVW in the absence of weak instrument bias and pleiotropy. Multiple MR methods have been developed to account for MR estimates under different violations of assumptions or the existence of heterogeneity, for example, MR-Egger[48] weighted median, and MR-pleiotropy residual sum and outlier.[50] In this study, we computed UMR estimates using models as listed below: Radial IVW fixed effect (RadialMR IVW [FE]), Radial IVW random effect (RadialMR IVW [RE]), and Radial MR-Egger (RadialMR Egger) in the RadialMR R package[49]; weighted median and weighted mode in the TwoSampleMR R package[47]; and MR-pleiotropy residual sum

and outlier (MR-PRESSO) in the MR-PRESSO.[51] Also, we obtained MVMR estimates using the following models: MVMR IVW in the MVMR package[52]; multivariable IVW fixed, random effect models, and Egger (namely, MR IVW [FE], MR IVW [RE], and MR Egger) in the MendelianRandomization package[53]; and multivariable MR-PRESSO in the MR-PRESSO.[51] Moreover, we computed MVMR estimates under multivariable Robust models (Robust), which are unbiased and more efficient than IVW with a low proportion of invalid instruments.[49] All of the causal estimates for RCDW and MCV are represented as odds ratios (ORs) per SD increase in genetically predicted levels of the exposures. The significance level of p value is 0.05.

Outliner exclusion, heterogeneity, and weak instrument assessment

We applied the following heterogeneity test statistics for outlier exclusion and heterogeneity assessment. For UMR, we excluded outliners and tested the heterogeneity and pleiotropy using the Cochran's Q (assuming balanced pleiotropy),[48] the Rücker's Q (assuming unbalanced pleiotropy) statistics,[48] and the MR-PRESSO global test.[51] We also performed MR under the Wald model to exclude influential SNPs.[54] Other heterogeneity tests based on the funnel plot and Higgins I^2 index were provided—0% and negative I^2 indicates no observed heterogeneity, and increasing values show higher heterogeneity.[55,56] We examined weak instrument bias according to the F statistic formula, where a rule of thumb of F > 10 is considered strong for instruments.[55] Finally, we performed the leave-one-out sensitivity test as sensitivity analysis.

We still removed pleiotropic SNPs in MVMR using MR-Lasso and MR-PRESSO as there might be some exposures acting as confounders not controlled in our model.[49] For example, hematocrit might be a confounder between MCV and hemorrhagic stroke.[21,57] So, there might be some pleiotropic instruments associated with MCV and hematocrit. Besides Cochran's Q and
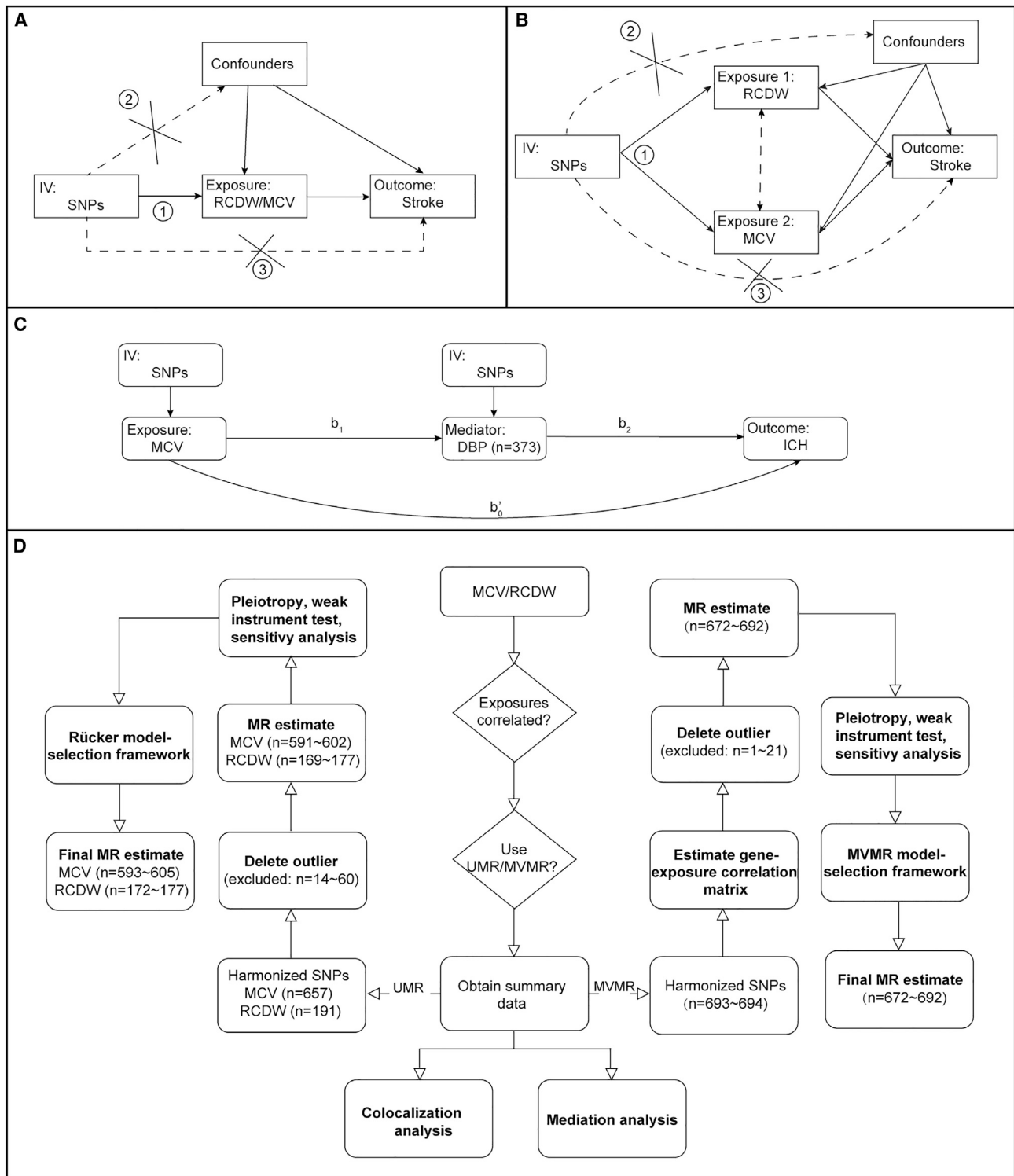
Figure 1. Flowchart of univariable and multivariable Mendelian randomization
UMR, univariable Mendelian randomization; MVMR, multivariable Mendelian randomization; IV, instrumental variable; SNPs, single-nucleotide polymorphisms; RCDW, red blood cell distribution width; MCV, mean corpuscular volume; ICH, intracerebral hemorrhage; nITH, nontraumatic intracranial hemorrhage; SBP, systolic blood pressure; DBP, diastolic blood pressure; BD, blood disorders.

MR-PRESSO, we computed Q_A (robust to weak instruments with appropriate type 1 error rate) to evaluate the heterogeneity.[58] Finally, we calculated the conditional F statistic to assess weak instrument bias by the MVMR package.[49]

Colocalization analysis
Apart from MR and genetic correlations, colocalization analysis is another widely used approach that leverages genetic data to

obtain insights into the genetic relationship between two traits.[59] We performed colocalization analysis only between traits that had a significant causal effect in MR to supplement the association between the two traits.

Description of colonization analysis

The colocalization analysis used is based on a novel Bayesian statistical test, which focuses on a single genomic region at a time to identify shared genes. This method explores whether the data support a shared causal variant for both traits. If there is more than one independent association at a locus for a trait of interest, only the strongest of these distinct association signals is considered by the algorithms.[60] Therefore, using the Bayesian-based colocalization analysis, we only have at most a single locus in a region at a time, which is the shared causal variant for the two traits.

At first, two datasets for two traits are from unrelated individuals with the same ethnicity. We relate genotypes X to phenotypes Y in a linear regression $Y = \mu + \beta X + \varepsilon$, assuming the causal variant is included in the set of Q variants and there is at most one association for each trait in the genomic region of interest. Then, we have a $2 \times Q$ matrix where, in each row, each element $\in \{0, 1\}$, indicates whether the variant is causally associated with the trait (1 means association). At most one element is non-zero for each trait within the region of Q variants. The two pairs of vectors generate five conditions as five hypotheses:

H_0: No association.
H_1: Association between variant and trait 1, but not trait 2.
H_2: Association between variant and trait 2, but not trait 1.
H_3: Association between variant and trait 1 and between trait 2, and two distinct variants.
H_4: Association between one shared variant and trait 1 and trait 2.

Sets of all variants for each hypothesis are S_0 S_1, S_2, S_3, S_4

We would like to calculate $P(H_i|D)$, $i \in \{1, 2, 3, 4\}$ from posterior computation.

$$P(H_i|D) \propto \sum_{S \in S_i} P(D|S)P(S) \Rightarrow \frac{P(H_i|D)}{P(H_0|D)} \propto \sum_{S \in S_i} \frac{P(D|S) \times P(S)}{P(D|S_0) \times P(S_0)}$$

$\frac{P(D|S)}{P(D|S_0)}$ in the above equation can be computed by summary statistics from GWAS and $\frac{P(S_i)}{P(S_0)}$ and $P(S_0)$ are assigned by the study.[60,61]

Finally, we denoted the posterior probabilities with PP0, PP1, PP2, PP3, and PP4 for SNP causality relating two traits for the above five mutually exclusive hypotheses. A large PP4 (PP4 > 75%) is evidence of colocalization,[60] suggesting a variant shared between RCDW/MCV and the outcome. Although there are other colocalization methods for multiple causal variants, for example, through Sum of Single Effects regression framework,[62] our focus was on the causal inference of MCV on hemorrhagic stroke. Colocalization analysis can supplement the relationship between MCV and hemorrhagic strokes but cannot infer causal effect between the two traits. We did not assess the number of the shared variants and used the more efficient Bayesian statistical test for colocalization.

Selection of locus for colocalization analysis

Before colocalization, we filtered minor allele frequency and allele frequency, and p values out of range of $(0, 1)$ and deleted missing SNP ID. The searching regions of colocalization are not consistent across studies. One study defined the regions within 500 kb of the lead SNPs,[63] another study searched for significant SNPs at ± 200 kb of the transcribed region of interest,[64] and another study

defined test regions from MR-independent SNPs within 200 kb distance to define 118 unique regions.[65] In this study, we searched for the lead SNP within a region of ± 1 MB and repeated the colocalization analysis within ± 3 MB using all the GWAS data. We did not screen the lead SNPs from UMR or MVMR data, because a shared SNP might not be in the causal path of exposure and outcome. The most associated SNPs might be excluded during MR pre-processing, such as during clumping or harmonization. We used the default setting of prior probabilities.

Mediation analysis

A mediator is a variable on the causal path between an exposure and an outcome variable; it is vertically pleiotropic. We would use mediation analysis to account for the mechanism whereby our exposure could affect the outcome. The rationale of mediation analysis, assumptions, estimation of mediation by MR, and implementation challenges have been discussed.[66,67] In mediation analysis, assuming no interaction, we defined and estimated three parameters for the effects of the exposure on the outcome: (1) the total effect b_0 is the effect of exposure on the outcome via all pathways; (2) the direct effect b'_0—controlled or natural—is the remaining effect beyond mediator pathways; and (3) the indirect effect (or mediation effect) is what acts through the mediator(s). The indirect effect can be calculated by $b_0 - b'_0$ with the difference method or $b_1 \times b_2$ with the product of coefficients method.[67,68] As shown in Figure 1C, suppose we only involve MCV as the exposure and ICH as the outcome in the mediation analysis (this is shown in the results). As suggested, we could apply the two-step MR (akin to the product of coefficients method), which assumes no interactions between exposures and mediators, for mediation analysis.[67,69] We computed b_0 and b_1 for MCV on ICH and on the mediator in MVMR, controlling for RCDW. We also evaluated b_2 in MVMR, controlling for both the exposures, to ensure any effect of the mediator on the outcome is independent of the exposure.[67] Considering the binary outcome with low prevalence (<10%), the proportion-mediated measure is approximately $1 + \frac{e^{b_0}+1}{e^{b_0+b_1b_2}-1}$ per SD increase in genetically instrumented MCV.[69] We applied bootstrap for resampling to calculate the confidence intervals.[70]

Results

Phenotypic correlation analysis

Figure 2 visualizes the phenotypic correlation between every trait after correction for multiple testing, where green indicates a positive correlation and red indicates a negative correlation. As MCV and RCDW are biologically correlated, they appear negatively associated ($r^2 = -0.10$; Figure 2A). We can witness high positive associations between SAH and AOS ($r^2 = 0.79$), between nITH and ICH ($r^2 = 0.78$), nITH and SAH ($r^2 = 0.69$), and nITH and AOS ($r^2 = 0.57$); they cluster around in the hierarchically clustered heatmap (Figure 2B), possibly due to the overlapped endpoints from SAH/ICH. Moreover, SAH is more closely related to AOS than to nITH.

UMR results

MR estimates for UMR are presented in the forest plots across Figures 3A–3D. Each forest plot presents the MR
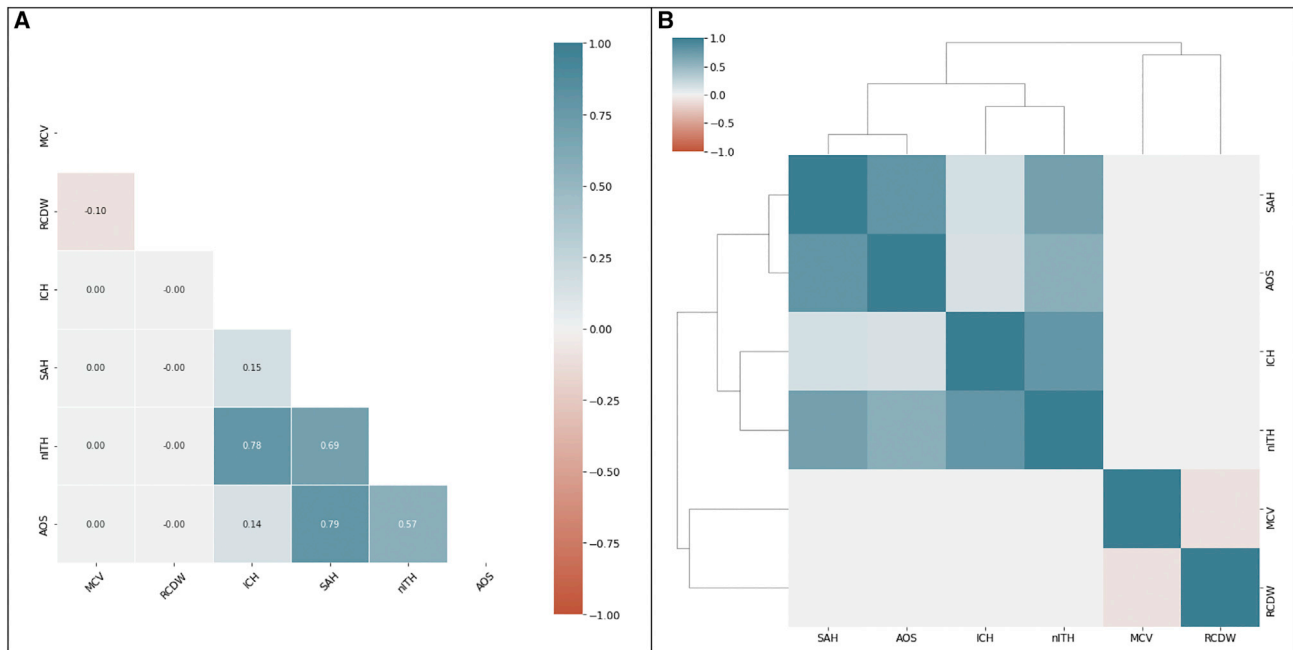
Figure 2. Phenotypic correlation after applying multiple testing corrections
(A) Phenotypic correlation matrix.
(B) Hierarchically clustered heatmap. MCV, mean corpuscular volume; RCDW, red blood cell distribution width; SAH, subarachnoid hemorrhage; ICH, intracerebral hemorrhage; nITH, nontraumatic intracranial hemorrhage; AOS, cerebral aneurysm, aneurysm operations, and SAH.

estimates between MCV/RCDW and the outcome, including the method, number of SNPs, OR (95% CI), and p values. The outcome title is presented at the bottom middle part. A pair of arrows above the outcome indicates the direction of risk that exposure results in for the outcomes. When the OR and its CI are below one toward the "lower risk" arrow ($p < 0.05$), this means genetically determined higher MCV/RCDW lowers the risk of the outcome; high MCV/RCDW is protective from the outcome. On the other hand, if the OR and its CI are over one toward the "higher risk" arrow ($p < 0.05$), it means genetically determined higher MCV/RCDW increases the risk of the outcome; high MCV/RCDW is a risk factor of the outcome. If the CI crosses one ($p \geq 0.05$), the causal effect of MCV/RCDW on the outcome is non-significant.

IVW fixed effect model RadialMR (FE) showed that the genetic risk of MCV decreased the odds of ICH (OR = 0.89 [0.8–0.99], p = 0.036) and nITH (OR = 0.89 [0.82–0.97], p = 0.005) per SD increase in genetically predicted levels of MCV (Figures 3A and 3C). No evidence of causality was found between MCV and SAH (OR = 0.88 [0.78–1], p = 0.051), and MCV and AOS (OR = 1.01 [0.91–1.11], p = 0.89); the same null results were observed between RCDW and any other outcome: $OR_{RCDW-SAH} = 1$ (0.85–1.19), p = 0.96; $OR_{RCDW-ICH} = 0.96$ (0.83–1.11), p = 0.57; $OR_{RCDW-nITH} = 0.97$ (0.87–1.09), p = 0.62; and $OR_{RCDW-AOS} = 0.96$ (0.84–1.1), p = 0.6; Figure 3A–3D. Heterogeneity test statistics presented a lack of heterogeneity and pleiotropy bias between every exposure-outcome pair (Cochran's

Q: p > 0.93; Rücker's Q: p > 0.06; MR-PRESSO global test: p > 0.95; $I^2 = 0$; Table S1). The F statistics were over 84, indicating no weak instruments bias (Table S1). The leave-one-out analyses showed the lack of influential instruments in any MR estimates (Figures S1–S8). The funnel plots are relatively symmetrical (Figures S9A–S10D), suggesting a lack of directional pleiotropy. These test indexes confirmed the use of RadialMR (FE) as the primary analysis for UMR estimates.

The IVW under the fixed effect model is most powerful when all IVs are valid.[71] Therefore, for other models (RadialMR Egger, weighted median, and weighted mode), ICH and nITH showed nonsignificant relation with hemorrhagic stroke, possibly due to limited power. MR-Egger estimators are less powerful and less efficient than IVW estimators due to the need to estimate both the slope and intercept parameters.[72] Compared with IVW, the wider confidence interval for MR-Egger is because the standard error of the MR-Egger estimate tends to be larger than that of the IVW method.[73] Moreover, instrument strength independent of direct effect violation introduces more bias in the MR-Egger estimate than in the IVW estimate.[74] Similarly, weighted median and weighted mode estimate the median or mode of the ratio estimate distribution as a causal effect. They could be less efficient than other estimators that combine multiple genetic variants more directly,[45] e.g., IVW.

For SAH, we observed a marginal significance for MCV (OR = 0.88 [0.78–1], p = 0.051) in RadialMR IVW (FE) while we observed significant effects in RadialMR IVW

A — Subarachnoid hemorrhage

Method	N SNPs	OR (95% CI)	P value
RadialMR IVW(FE)			
RCDW	174	1 (0.85,1.19)	9.6e−01
MCV	597	0.88 (0.78,1)	5.1e−02
RadialMR IVW(RE)			
RCDW	174	1 (0.86,1.17)	9.6e−01
MCV	597	0.88 (0.79,0.99)	3.5e−02
RadialMR Egger			
RCDW	174	0.87 (0.63,1.22)	4.2e−01
MCV	597	1 (0.82,1.23)	9.6e−01
Weighted median			
RCDW	143	0.93 (0.7,1.24)	6.2e−01
MCV	521	0.86 (0.69,1.08)	1.9e−01
Weighted mode			
RCDW	143	0.82 (0.55,1.23)	3.5e−01
MCV	521	0.85 (0.62,1.15)	2.9e−01
MR−PRESSO			
RCDW	143	1 (0.87,1.16)	9.6e−01
MCV	521	0.88 (0.79,0.98)	2.5e−02
MR−PRESSO(outlier corrected)			
RCDW	NA	NA	NA
MCV	NA	NA	NA

Scale: 0.50 1.0 1.29 Lower Risk Higher Risk

B — Intracerebral hemorrhage

Method	N SNPs	OR (95% CI)	P value
RadialMR IVW(FE)			
RCDW	177	0.96 (0.83,1.11)	5.7e−01
MCV	602	0.89 (0.8,0.99)	3.6e−02
RadialMR IVW(RE)			
RCDW	177	0.96 (0.83,1.11)	5.7e−01
MCV	602	0.89 (0.8,0.99)	3.0e−02
RadialMR Egger			
RCDW	177	1.17 (0.86,1.57)	3.2e−01
MCV	602	0.89 (0.74,1.06)	1.9e−01
Weighted median			
RCDW	145	1.12 (0.86,1.46)	3.9e−01
MCV	528	1 (0.81,1.22)	9.6e−01
Weighted mode			
RCDW	145	1.19 (0.83,1.69)	3.5e−01
MCV	528	0.97 (0.76,1.23)	8.1e−01
MR−PRESSO			
RCDW	145	0.96 (0.84,1.1)	5.4e−01
MCV	528	0.89 (0.81,0.99)	2.5e−02
MR−PRESSO(outlier corrected)			
RCDW	NA	NA	NA
MCV	NA	NA	NA

Scale: 0.69 1.0 1.74 Lower Risk Higher Risk

C — Nontraumatic intracranial hemorrhage

Method	N SNPs	OR (95% CI)	P value
RadialMR IVW(FE)			
RCDW	175	0.97 (0.87,1.09)	6.2e−01
MCV	605	0.89 (0.82,0.97)	5.4e−03
RadialMR IVW(RE)			
RCDW	175	0.97 (0.89,1.07)	5.4e−01
MCV	605	0.89 (0.82,0.96)	2.3e−03
RadialMR Egger			
RCDW	175	1.04 (0.85,1.28)	6.8e−01
MCV	605	0.91 (0.8,1.05)	2.0e−01
Weighted median			
RCDW	144	0.98 (0.8,1.19)	8.1e−01
MCV	527	0.92 (0.78,1.08)	3.2e−01
Weighted mode			
RCDW	144	1.05 (0.81,1.38)	7.0e−01
MCV	527	0.93 (0.78,1.1)	4.0e−01
MR−PRESSO			
RCDW	144	0.97 (0.89,1.07)	5.5e−01
MCV	527	0.89 (0.82,0.96)	2.3e−03
MR−PRESSO(outlier corrected)			
RCDW	NA	NA	NA
MCV	NA	NA	NA

Scale: 0.73 1.0 1.43 Lower Risk Higher Risk

D — Aneurysms, operation, subarachnoid hemorrhage

Method	N SNPs	OR (95% CI)	P value
RadialMR IVW(FE)			
RCDW	172	0.96 (0.84,1.1)	6.0e−01
MCV	593	1.01 (0.91,1.11)	8.9e−01
RadialMR IVW(RE)			
RCDW	172	0.96 (0.86,1.08)	5.3e−01
MCV	593	1.01 (0.92,1.11)	8.8e−01
RadialMR Egger			
RCDW	172	0.88 (0.67,1.15)	3.4e−01
MCV	593	1.14 (0.97,1.34)	1.1e−01
Weighted median			
RCDW	141	0.92 (0.72,1.17)	5.1e−01
MCV	519	1.07 (0.89,1.29)	4.8e−01
Weighted mode			
RCDW	141	0.86 (0.63,1.16)	3.2e−01
MCV	519	1.09 (0.89,1.33)	4.0e−01
MR−PRESSO			
RCDW	141	0.96 (0.85,1.09)	5.7e−01
MCV	519	1.01 (0.92,1.1)	8.8e−01
MR−PRESSO(outlier corrected)			
RCDW	NA	NA	NA
MCV	NA	NA	NA

Scale: 0.58 1.0 1.39 Lower Risk Higher Risk

Figure 3. Univariable Mendelian randomization analysis results
OR (95% CI) estimates in odds ratio and the 95% confidence intervals for outcome risk per 1 SD increase in exposure genetically predicted levels. Significance level for p is 0.05. IVW, inverse variance weighted; FE, fixed effect; RE, random effect; RCDW, red blood cell distribution width; MCV, mean corpuscular volume; MR-PRESSO, Mendelian randomization pleiotropy residual sum and outlier; SNP, single-nucleotide polymorphism; CIs, confidence intervals.

(RE) (OR = 0.88 [0.79,0.99], p = 0.035) and MR-PRESSO (OR = 0.88 [0.79,0.98], p = 0.025); non-significance existed in the remaining models. MR-PRESSO also adopted IVW before and after removal of outliers.[51] However, we still considered random effects, and fixed effect models resulted in rather consistent results because the point estimators were the same and the latter just produced a 0.01 wider confidence interval than the former. The random effects p value of 0.051 is marginal and may be clinically significant.[75] We would further compare them when adjusting RCDW in MVMR.

To conclude, UMR suggested that genetically predicted MCV levels are associated with lower ICH and nITH risk. MCV and SAH association needed further exploration in MVMR. No substantial evidence was found for an association between MCV/RCDW and other outcomes.

MVMR results

MR estimates for MVMR were presented in the forest plots across Figures 4A–4D, which could be interpreted similarly to Figures 3A–3D. After adjusting for RCDW, we found MCV still reduced the risk of ICH (OR = 0.87 [0.78–0.98], p = 0.021) and nITH (OR = 0.88 [0.8–0.96], p = 0.004) under the IVW fixed effect model—MR IVW (FE); other models also showed similar MR estimates ($OR_{MCV-ICH}$ ranged from 0.87 to 0.88; $OR_{MCV-nITH}$ were all 0.88) with consistent significance (p < 0.05) (Figures 4B and 4C). There was no support for causal effects in the remaining exposure-outcome estimates in all models. For example, under MR IVW (FE), $OR_{MCV-SAH}$ = 0.91 (0.8–1.03), p = 0.14; $OR_{MCV-AOS}$ = 1.02 (0.92–1.13), p = 0.73; $OR_{RCDW-SAH}$ = 0.94 (0.79–1.12), p = 0.49; $OR_{RCDW-ICH}$ = 0.88 (0.76–1.03), p = 0.1; $OR_{RCDW-nITH}$ = 0.91 (0.81–1.03), p = 0.14; $OR_{RCDW-AOS}$ = 1.02 (0.89–1.17), p = 0.79 (Figures 4A–4D). Heterogeneity test statistics were all strong indications of a lack of heterogeneity and pleiotropy (Cochran's Q p > 0.44, Q_A p > 0.44, and MR-PRESSO global test p > 0.46 in Table S2). Moreover, for weak instrument assessment, the conditional F statistics (Table S2) were over 18; the instruments were adequately powerful for MVMR. Besides the Rücker model-selection framework, following the MVMR model-selection framework, we could also choose MVMR IVW as our preferred model because of the lack of evidence of heterogeneity and weak instruments.[76] Thus, we considered the MVMR IVW as the sensitivity analysis and alternative MVMR estimation. We could still observe similar results and draw the same casual inference for MCV and ICH (OR = 0.87 [0.78–0.98], p = 0.02), and MCV and nITH (OR = 0.88 [0.8–0.96], p = 0.004) (Figures 4A and 4C); we found no evidence to support causation between any other exposure and outcome (Figures 4A–4D).

MVMR produced more consistent MR estimates than UMR, justifying the control of pleiotropic instruments from RCDW and MCV in our analysis. MR estimates for MCV with SAH in UMR all became nonsignificant in MVMR, indicating that the adjustment strengthened the

stability in MR estimates. We considered MVMR, especially IVW models, as the final MR estimates for the interpretation of our study. Therefore, genetically determined MCV was significantly related to ICH and nITH. The ORs of ICH and nITH were 0.87 (0.78–0.98), p = 0.021 and 0.88 (0.8–0.96), p = 0.004, respectively, per genetically predicted 1 SD increase in the MCV level.
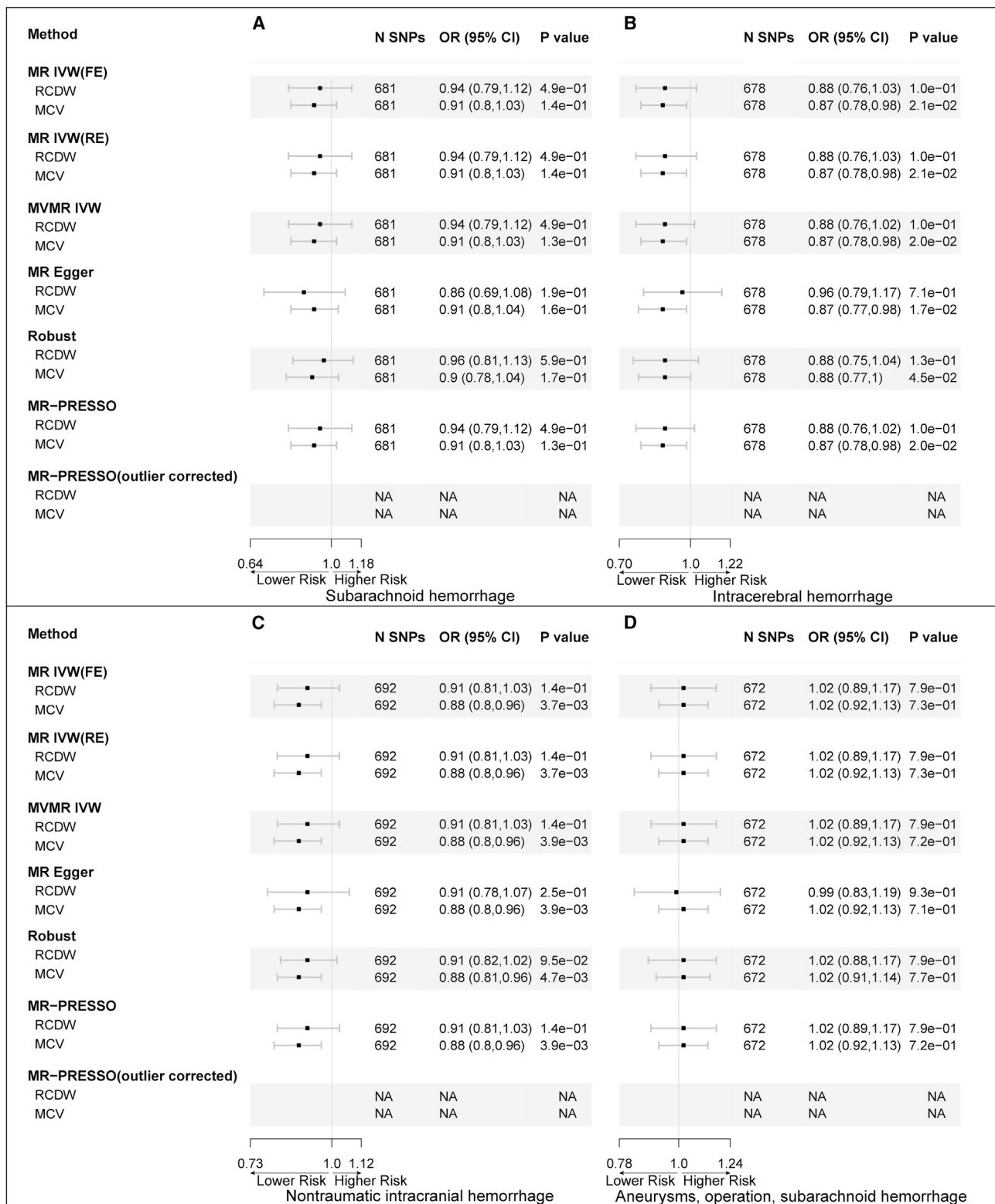
Colocalization analysis

We conducted the colocalization analyses between MCV and ICH, and MCV and nITH. The PP4 for these two pairs of traits—within the ± 1 MB region of the top hit—were 75.1% and 26.4%; similar values were witnessed within the ± 3 MB region (75.6% and 25.9%) (Table 2). Using colocalization analysis based on Bayesian statistical test, we finally resulted in a shared SNP (rs62393683) within two regions for MCV and ICH that both had PP4 over 75% (Table 2), indicating a strong support for colocalization between two traits; an association was found between MCV and ICH61. While for MCV and nITH, the two PP4 did not support colocalization, possibly due to the result of limited power61. This SNP (rs62393683) was not involved in our UMR or MVMR data. We queried PhenoScanner[77]—a database of human genotype-phenotype associations—for this SNP and found that it is associated with MCV, RCDW, hemoglobin concentration, hematocrit, diastolic blood pressure, hypertension, hereditary/genetic hematological disorder, etc. However, hemorrhagic stroke is not recorded in these associations. Our colocalization result might fill in the gap of the association of MCV and ICH and contribute to the large-scale genetic association studies.

Mediation results

We applied the IVW fixed effect model to compute the MVMR estimates of MCV on the mediators and outcomes (results are shown in Figure 5. The interpretation of Figure 5 was different from Figures 3 and 4. As mentioned in the Materials and methods, b_0 is the MR estimate of MCV adjusted by RCDW for outcome, i.e., $b_0 = b_{MCV|RCDW-out}$; b_1 is the MR estimate of MCV adjusted by RCDW for meditator, i.e., $b_1 = b_{MCV|RCDW-med}$; b'_0 is the estimate of MCV adjusted by RCDW and mediator for outcome, i.e., $b'_0 = b_{MCV|RCDWmed-out}$; and b_2 is the estimate of the mediator adjusted by MCV and RCDW for outcome, i.e., $b_2 = b_{med|exp-out}$. Figure 5 includes the names of outcomes or mediators of estimates, number of SNPs, MR estimates (95% CI), and p values.

In two-step MVMR, for step one b_1 (MCV on mediator) we observed that only MCV on DBP, conditional on RCVDW, showed significant estimates; we still had similar findings when we used MVMR IVW −0.085 (−0.166 to −0.003, p = 0.031; Table S3). Therefore, we selected DBP as the mediator in the following analysis. For step two b_2, effect of DBP on ICH/nITH, total effect b_0, and direct effect b'_0, we also had significant estimates. The heterogeneity tests (Cochran's Q p > 0.45 and Q_A p > 0.45), MR-PRESSO global test (p > 0.46), and conditional F

A — Subarachnoid hemorrhage

Method		N SNPs	OR (95% CI)	P value
MR IVW(FE)	RCDW	681	0.94 (0.79,1.12)	4.9e−01
	MCV	681	0.91 (0.8,1.03)	1.4e−01
MR IVW(RE)	RCDW	681	0.94 (0.79,1.12)	4.9e−01
	MCV	681	0.91 (0.8,1.03)	1.4e−01
MVMR IVW	RCDW	681	0.94 (0.79,1.12)	4.9e−01
	MCV	681	0.91 (0.8,1.03)	1.3e−01
MR Egger	RCDW	681	0.86 (0.69,1.08)	1.9e−01
	MCV	681	0.91 (0.8,1.04)	1.6e−01
Robust	RCDW	681	0.96 (0.81,1.13)	5.9e−01
	MCV	681	0.9 (0.78,1.04)	1.7e−01
MR−PRESSO	RCDW	681	0.94 (0.79,1.12)	4.9e−01
	MCV	681	0.91 (0.8,1.03)	1.3e−01
MR−PRESSO(outlier corrected)	RCDW	NA	NA	NA
	MCV	NA	NA	NA

0.64 1.0 1.18
← Lower Risk Higher Risk →

B — Intracerebral hemorrhage

Method		N SNPs	OR (95% CI)	P value
MR IVW(FE)	RCDW	678	0.88 (0.76,1.03)	1.0e−01
	MCV	678	0.87 (0.78,0.98)	2.1e−02
MR IVW(RE)	RCDW	678	0.88 (0.76,1.03)	1.0e−01
	MCV	678	0.87 (0.78,0.98)	2.1e−02
MVMR IVW	RCDW	678	0.88 (0.76,1.02)	1.0e−01
	MCV	678	0.87 (0.78,0.98)	2.0e−02
MR Egger	RCDW	678	0.96 (0.79,1.17)	7.1e−01
	MCV	678	0.87 (0.77,0.98)	1.7e−02
Robust	RCDW	678	0.88 (0.75,1.04)	1.3e−01
	MCV	678	0.88 (0.77,1)	4.5e−02
MR−PRESSO	RCDW	678	0.88 (0.76,1.02)	1.0e−01
	MCV	678	0.87 (0.78,0.98)	2.0e−02
MR−PRESSO(outlier corrected)	RCDW	NA	NA	NA
	MCV	NA	NA	NA

0.70 1.0 1.22
← Lower Risk Higher Risk →

C — Nontraumatic intracranial hemorrhage

Method		N SNPs	OR (95% CI)	P value
MR IVW(FE)	RCDW	692	0.91 (0.81,1.03)	1.4e−01
	MCV	692	0.88 (0.8,0.96)	3.7e−03
MR IVW(RE)	RCDW	692	0.91 (0.81,1.03)	1.4e−01
	MCV	692	0.88 (0.8,0.96)	3.7e−03
MVMR IVW	RCDW	692	0.91 (0.81,1.03)	1.4e−01
	MCV	692	0.88 (0.8,0.96)	3.9e−03
MR Egger	RCDW	692	0.91 (0.78,1.07)	2.5e−01
	MCV	692	0.88 (0.8,0.96)	3.9e−03
Robust	RCDW	692	0.91 (0.82,1.02)	9.5e−02
	MCV	692	0.88 (0.81,0.96)	4.7e−03
MR−PRESSO	RCDW	692	0.91 (0.81,1.03)	1.4e−01
	MCV	692	0.88 (0.8,0.96)	3.9e−03
MR−PRESSO(outlier corrected)	RCDW	NA	NA	NA
	MCV	NA	NA	NA

0.73 1.0 1.12
← Lower Risk Higher Risk →

D — Aneurysms, operation, subarachnoid hemorrhage

Method		N SNPs	OR (95% CI)	P value
MR IVW(FE)	RCDW	672	1.02 (0.89,1.17)	7.9e−01
	MCV	672	1.02 (0.92,1.13)	7.3e−01
MR IVW(RE)	RCDW	672	1.02 (0.89,1.17)	7.9e−01
	MCV	672	1.02 (0.92,1.13)	7.3e−01
MVMR IVW	RCDW	672	1.02 (0.89,1.17)	7.9e−01
	MCV	672	1.02 (0.92,1.13)	7.2e−01
MR Egger	RCDW	672	0.99 (0.83,1.19)	9.3e−01
	MCV	672	1.02 (0.92,1.13)	7.1e−01
Robust	RCDW	672	1.02 (0.88,1.17)	7.9e−01
	MCV	672	1.02 (0.91,1.14)	7.7e−01
MR−PRESSO	RCDW	672	1.02 (0.89,1.17)	7.9e−01
	MCV	672	1.02 (0.92,1.13)	7.2e−01
MR−PRESSO(outlier corrected)	RCDW	NA	NA	NA
	MCV	NA	NA	NA

0.78 1.0 1.24
← Lower Risk Higher Risk →

Figure 4. MVMR analysis results
OR (95% CI) estimates in odds ratio and the 95% confidence intervals for outcome risk per 1 SD increase in exposure genetically predicted levels. Significance level for p is 0.05. RCDW, red blood cell distribution width; MCV, mean corpuscular volume; IVW, inverse variance weighted; FE, fixed effect; RE, random effect; MR-PRESSO, Mendelian randomization pleiotropy residual sum and outlier; SNP, single-nucleotide polymorphism; OR, odds ratio.

Table 2. Colocalization analysis results under significant MR estimates

Region within position ± 1 MB

Chr	Position	Genes	Traits	Candidate SNP	NSNP	PP0	PP1	PP2	PP3	PP4
6	**25,633,204**	***ZFP57***	**MCV-ICH**	**rs62393683**	**138,664**	**0.000**	**0.134**	**0.000**	**0.115**	**0.751**
6	25,633,204	*ZFP57*	MCV-nITH	rs62393683	138,664	0.000	0.425	0.000	0.311	0.264

Region within position ± 3 MB

Chr	Position	Genes	Traits	Candidate SNP	NSNP	PP0	PP1	PP2	PP3	PP4
6	**25,633,204**	***ZFP57***	**MCV-ICH**	**rs62393683**	**413,360**	**0.000**	**0.135**	**0.000**	**0.109**	**0.756**
6	25,633,204	*ZFP57*	MCV-nITH	rs62393683	413,359	0.000	0.417	0.000	0.323	0.259

PP0-PP4: five posterior probabilities for SNP causality relating two traits for five mutually exclusive hypotheses: (1) H0, association with neither trait; (2) H1, association with trait 1 but not trait 2; (3) H2, association with trait 2 but not trait 1; (4) H3, association with both traits, with two distinct SNPs; and (5) H4, association with both traits, with one shared SNP. SNP, single nucleotide polymorphism; candidate SNP, candidate causal SNP with maximum PP4 in a test region; NSNP, the number of SNPs included in the colocalization analysis; MCV, mean corpuscular volume; RCDW, red blood cell distribution width; ICH, intracerebral hemorrhage; SAH, subarachnoid hemorrhage; nITH, nontraumatic intracranial hemorrhage.

statistics (>10), as calculated in the MVMR analysis, showed a lack of heterogeneity effects or weak instrument bias (Table S3). However, as DBP consists of data from the International Consortium for Blood Pressure and UK Biobank, the study participants with DBP and MCV data overlapped partly. We assessed the sample overlap bias and type 1 error inflation.[78] Across the whole range of possible sample overlap, it suggested no increase of bias (Table S4). Therefore, we could infer the statistics for mediation effect from MR should be robust to unmeasured confounding.

Table 3 presents the OR for total effect, indirect effect, direct effect, and proportion-mediated measure for MCV on ICH via DBP and MCV on nITH via DBP. Significant results for these statistics indicated diastolic blood pressure mediates between MCV and hemorrhagic strokes. Indirect effects of MCV on hemorrhagic via DBP existed, although they were close to null (OR = 0.995 [0.99,0.999]). The proportion mediated was also small, occupying 3.1% (0.1%,14.3%) in ICH and 3.4% (0.2%,15.8%) in nITH; we might infer that the causal effect of MCV on hemorrhagic strokes could result from the independent causal mechanism of MCV or other mediators.
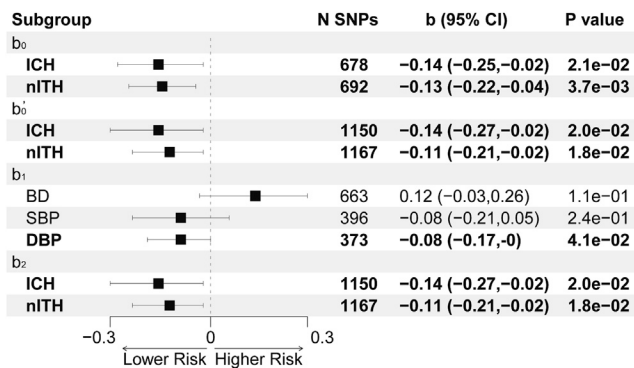
Discussion

This analysis explores the causal effects of MCV and RCDW on hemorrhagic strokes, including SAH, ICH, nITH, and AOS. First, our phenotypic correlation analysis suggested a negative correlation between RCDW and MCV. Second, UMR and MVMR revealed that only higher genetically predicted MCV is protective of ICH and nITH. There is no strong evidence of a causal inference between RCDW/MCV and any other outcome. Third, the results of colocalization analysis proposed a single variant on the causal pathway of MCV and ICH. Finally, mediation analysis suggested that diastolic blood pressure mediates between MCV and ICH/nITH. To the best of our knowledge, this is the first study to report the causal effects of RCDW/MCV on the risk of hemorrhagic strokes.

Although the Rücker model selection framework suggested that the IVW under the fixed effect model be selected for interpretation of our results, the IVW fixed-model or multiplicative random effect model should produce a consistent MR estimate when there is either a lack of pleiotropy (intercept = 0 for all variants) or a balanced pleiotropy, but the residual standard error of the random effects model can be greater.[48,79] However, in the UMR estimate between MCV and SAH, our standard error was smaller in IVW (RE) than in IVW (FE), and the causal effect shown by UMR IVW (RE) was found toward the null in the MVMR. We speculated that there might be pleiotropic SNPs in UMR and, subsequently, they might be controlled in MVMR. We also observed that all estimates were consistent between each exposure-outcome pair. Therefore, our MVMR should produce more robust estimates than the UMR.

Our results are consistent with previous studies. Hypertension was more common in the low-MCV group, although among chronic kidney disease patients.[29] We could also infer that genetically determined lower MCV increased the risk of DBP in our study. Our phenotypic correlation analysis presented a negative association between RCDW and MCV, consistent with previous studies.[25,80] A meta-analysis and a longitudinal study found that elevated RCDW was a risk factor for general stroke and ischemic stroke but not for SAH or ICH.[23,81] Our findings showed no supportive evidence for MCV on SBP, controlling for RCDW, which was consistent with the previous report about the lack of a significant relationship between SBP and MCV after adjusting for red blood cell count.[21]

The underlying mechanism behind MCV and ICH and nITH may be mediated by blood pressure. Studies showed that there is an important adaptive mechanism to regulate blood viscosity to excessive erythrocytosis by increasing more juvenile flexible erythrocytes of higher MCV.[32,33] We found a mediation effect from DBP instead of SBP. This might be because DBP is more specific to measure overall resistance to blood flow than SBP and thus is more strongly associated with blood characteristics that influence viscosity, i.e., MCV.[21] When whole blood viscosity

Subgroup		N SNPs	b (95% CI)	P value
b_0				
ICH		678	−0.14 (−0.25,−0.02)	2.1e−02
nITH		692	−0.13 (−0.22,−0.04)	3.7e−03
b'_0				
ICH		1150	−0.14 (−0.27,−0.02)	2.0e−02
nITH		1167	−0.11 (−0.21,−0.02)	1.8e−02
b_1				
BD		663	0.12 (−0.03,0.26)	1.1e−01
SBP		396	−0.08 (−0.21,0.05)	2.4e−01
DBP		373	−0.08 (−0.17,−0)	4.1e−02
b_2				
ICH		1150	−0.14 (−0.27,−0.02)	2.0e−02
nITH		1167	−0.11 (−0.21,−0.02)	1.8e−02

−0.3   0   0.3
Lower Risk Higher Risk

Figure 5. Two-step MVMR analysis results
b_0, total effect of mean corpuscular volume (MCV) on outcomes conditional on red blood cell distribution width (RCDW); b'_0, direct effect of MCV on outcomes conditional on RCDW; b_1, effect of MCV on mediators conditional on RCDW; b_2, effect of MCV on outcomes conditional on RCDW and diastolic blood pressure; ICH, intracerebral hemorrhage; nITH, nontraumatic intracranial hemorrhage; BD, blood disorders; SBP, systolic blood pressure; DBP, diastolic blood pressure.

increases, peripheral resistance to blood flow also increases, and then blood pressure increases to maintain blood flow in the face of increased peripheral resistance.[21] Blood viscosity was positively associated with systolic blood pressure, diastolic blood pressure, and mean arterial pressure.[34] Therefore, we could infer that decreased whole blood viscosity would result in decreased blood pressure.

How MCV is applied to the prevention of hemorrhagic strokes in medical practice, i.e., the external validity, is still an open question. First, our result was based on the European population. It could not be extended to other populations with ethnic backgrounds different from our study. In the Korean elderly population, MCV values were significantly greater than in young adults,[82] while our mean age of first-ever hemorrhagic stroke was between 50 and 70 years. There might be heterogeneity to interpret the results of MCV associated with diseases under different conditions. For example, some studies proposed that elevated MCV is associated with a reduced incidence of metabolic syndrome and predicts a lower risk of in-stent restenosis.[31,83] There are multiple studies indicating that subjects with higher MCV levels have lower body mass index, hemoglobin A1c, cholesterol, triglyceride, uric acid levels, diabetes mellitus, dyslipidemia, hypertension, and (for men only) metabolic syndrome[28–31] However, higher MCV levels are also related to an increased risk of all-cause mortality and cognitive impairment.[28,29,84,85] Athletes with adequate and regular training have higher MCV levels than sedentary subjects.[86] Taken together, it seems that elevated MCV level reduces a series of health problems, although this is not always the case. Heterogeneity exists in subjects with different genders, patients with anemia or chronic diseases, and athletes without a normal MCV. Our study conclusion can only be generalized to the European population. There was not much other phenotypic information in our data. Second, MCV may be altered by

changes in crystalline osmotic pressure.[31] However, according to one of our co-authors as a neurologist, currently there is still a lack of prescribed medications that can modulate MCV. As mentioned that trained athletes have higher MCV levels than the sedentary subjects,[86] it is likely that MCV can be altered by physical activity. Still, the efficacy of short-term targeted interventions on MCV cannot be assessed. But these do not preclude the efficacy of a therapeutic intervention on levels of MCV. Our study still provides a cost-effective target for drug development for the prevention of hemorrhagic strokes.

Our study may provide some clinical implications. We have identified novel potential targets for pharmacological therapy (RCDW/MCV) to reduce the risk of ICH and nITH. As higher genetically predicted MCV is considered protective of ICH and nITH, experiment-based research is warranted to explore further whether this finding can be exploited to develop new pharmaceutical treatments for hemorrhagic strokes. However, as we found supportive evidence of causality between MCV and ICH or nTH (combination of ICH and SAH), but not between MCV and SAH, the causation of MCV and nTH might result from that of MCV and ICH. Further investigations are needed to confirm this inference and clarify the association between MCV and nITH.

Our study has some limitations. First, we did not involve and adjust other red blood cell indices in MVMR, such as hematocrit and mean corpuscular hemoglobin, which are correlated with MCV and RCDW and might affect stroke[87]; they might share pleiotropic SNPs and should be further controlled in MVMR. Second, we did not conduct a reverse casual study for hemorrhagic strokes on MCV/RCDW. It is unknown whether reverse causation exists. Third, our data are collected from the European population. Our conclusion could not be extended to other races. Fourth, we did not perform casual inference using non-linear models. We cannot detect the causal relationship that might be non-linear for our null results under linear models. Despite these limitations, we found no heterogeneity or weak instrument bias, adopted multiple rigorous MR estimates, and followed documented approaches to interpret our results and present convincing findings on the causal effect of RCDW and MCV on hemorrhagic strokes.

In conclusion, we conducted the first MR analysis to investigate the role of genetically elevated RCDW and MCV in the risk of hemorrhagic strokes. We found that MCV might reduce the risk of ICH and nITH. These findings may provide new insights into potential drug targets for hemorrhagic stroke.

Data and code availability
This research used secondary and publicly available GWAS data from studies with all participants' informed consent and review boards and/or ethics committees' ethical approval.

Data are available directly from Medical Research Council Integrative Epidemiology Unit OpenGWAS database (ID

Table 3. Mediation effect of MCV on ICH and nITH via DBP

Exposure	Mediator	Outcome	Total effect (OR)	Indirect effect (OR)	Direct effect (OR)	Proportion mediated (%)
MCV	DBP	ICH	0.873 (0.778,0.98)	0.995 (0.99,0.999)	0.866 (0.767,0.977)	3.1% (0.1%,14.3%)
			p = 0.021	p < 0.001	p = 0.02	
MCV	DBP	nITH	0.876 (0.801,0.958)	0.995 (0.99,0.999)	0.866 (0.767,0.977)	3.4% (0.2%,15.8%)
			p = 0.004	p < 0.001	p = 0.02	

as shown across Tables S1, S2, and S3), the Neale Lab – UK Biobank, and the FinnGen consortium (see Web resources). We also provide data processed across Tables S5, S6, S7, and S8. Codes for the analysis are available at Github (https://github.com/bbb801/MR).

Supplemental information

Acknowledgments

Declaration of interests

Web resources

American Stroke Association: https://www.stroke.org/en/about-stroke/types-of-stroke/hemorrhagic-strokes-bleeds

FinnGenEN: https://www.finngen.fi/fi

Medical Research Council Integrative Epidemiology Unit OpenGWAS: https://gwas.mrcieu.ac.uk/

UK Biobank database: https://www.ukbiobank.ac.uk/

UK Biobank - Neale Lab: http://www.nealelab.is/uk-biobank

Winchester Hospital: https://www.winchesterhospital.org/

References

1. Unnithan, A.K.A., Das, J.M., and Mehta, P. (2022). Hemorrhagic Stroke (StatPearls).
2. Torpy, J.M., Burke, A.E., and Glass, R.M. (2010). Hemorrhagic stroke. JAMA *303*, 2312.
3. Madangarli, N., Bonsack, F., Dasari, R., and Sukumari–Ramesh, S. (2019). Intracerebral hemorrhage: blood components and neurotoxicity. Brain Sci. *9*, 316.
4. Ariesen, M.J., Claus, S.P., Rinkel, G.J.E., and Algra, A. (2003). Risk factors for intracerebral hemorrhage in the general population: a systematic review. Stroke *34*, 2060–2065.
5. Grysiewicz, R.A., Thomas, K., and Pandey, D.K. (2008). Epidemiology of ischemic and hemorrhagic stroke: incidence, prevalence, mortality, and risk factors. Neurol. Clin. *26*, 871–895.
6. Boehme, A.K., Esenwa, C., and Elkind, M.S.V. (2017). Stroke risk factors, genetics, and prevention. Circ. Res. *120*, 472–495.
7. Maner, B.S., and Moosavi, L. (2019). Mean Corpuscular Volume.
8. Feng, G.H., Li, H.P., Li, Q.L., Fu, Y., and Huang, R.B. (2017). Red blood cell distribution width and ischaemic stroke. Stroke Vasc. Neurol. *2*, 172–175.
9. Hatamian, H., Saberi, A., and Pourghasem, M. (2014). The relationship between stroke mortality and red blood cell parameters. Iran. J. Neurol. *13*, 237–240.
10. Aksoy, D., İnanir, A., Ayan, M., Çevik, B., Kurt, S., and Karaer Ünaldi, H. (2013). Predictors of mortality and morbidity in acute ischemic stroke. npa. *50*, 40–44.
11. Li, N., Zhou, H., and Tang, Q. (2017). Red blood cell distribution width: a novel predictive indicator for cardiovascular and cerebrovascular diseases. Dis. Markers *2017*.
12. Kor, C.T., Hsieh, Y.P., Chang, C.C., and Chiu, P.F. (2018). The prognostic value of interaction between mean corpuscular volume and red cell distribution width in mortality in chronic kidney disease. Sci. Rep. *8*, 11870–11878.
13. Maner, B.S., and Moosavi, L. (2019). Mean Corpuscular Volume (StatPearls).
14. Harshfield, E.L., Sims, M.C., Traylor, M., Ouwehand, W.H., and Markus, H.S. (2020). The role of haematological traits in risk of ischaemic stroke and its subtypes. Brain *143*, 210–221.
15. Ling, J., Fang, M., and Wu, Y. (2022). Association of red cell distribution width and D-dimer levels with intracranial hemorrhage in patients with cerebral venous thrombosis. Clin. Neurol. Neurosurg. *214*, 107178.
16. Bruschi, G., Minari, M., Bruschi, M.E., Tacinelli, L., Milani, B., Cavatorta, A., and Borghetti, A. (1986). Similarities of essential and spontaneous hypertension. Volume and number of blood cells. Hypertension *8*, 983–989.
17. Livesey, J.A., Manning, R.A., Meek, J.H., Jackson, J.E., Kulinskaya, E., Laffan, M.A., and Shovlin, C.L. (2012). Low serum iron levels are associated with elevated plasma levels of coagulation factor VIII and pulmonary emboli/deep venous thromboses in replicate cohorts of patients with hereditary haemorrhagic telangiectasia. Thorax *67*, 328–333.
18. Powers, J.M., Sandoval, C., and Lorin, M.I. (2021). Approach to the Child with Anemia.
19. Allen, J.P., and Litten, R.Z. (2001). The role of laboratory tests in alcoholism treatment. J. Subst. Abuse Treat. *20*, 81–85.

20. Postnov, Y.V., Kravtsov, G.M., Orlov, S.N., Pokudin, N.I., Postnov, I.Y., and Kotelevtsev, Y.V. (1988). Effect of protein kinase C activation on cytoskeleton and cation transport in human erythrocytes. Reproduction of some membrane abnormalities revealed in essential hypertension. Hypertension *12*, 267–273.

21. Sharp, D.S., Curb, J.D., Schatz, I.J., Meiselman, H.J., Fisher, T.C., Burchfiel, C.M., Rodriguez, B.L., and Yano, K. (1996). Mean red cell volume as a correlate of blood pressure. Circulation *93*, 1677–1684.

22. Rezende, S.M., Lijfering, W.M., Rosendaal, F.R., and Cannegieter, S.C. (2014). Hematologic variables and venous thrombosis: red cell distribution width and bloodmonocyte count are associated with an increased risk. Haematologica *99*, 194–200.

23. Song, S.Y., Hua, C., Dornbors, D., Kang, R.J., Zhao, X.X., Du, X., He, W., Ding, Y.C., and Meng, R. (2019). Baseline red blood cell distribution width as a predictor of stroke occurrence and outcome: a comprehensive meta-analysis of 31 studies. Front. Neurol. *10*, 1237.

24. Wang, H., Yang, G., Zhao, J., and Wang, M. (2020). Association between mean corpuscular volume and severity of coronary artery disease in the Northern Chinese population: a cross-sectional study. J. Int. Med. Res. *48*. 300060519896713.

25. Seo, S.-G., Lee, M.-Y., Park, S.H., Han, J.M., Lee, K.-B., Kim, H., and Hyun, Y.Y. (2020). The association between red cell distribution width and incident hypertension in Korean adults. Hypertens. Res. *43*, 55–61.

26. Wen, Y. (2010). High red blood cell distribution width is closely associated with risk of carotid artery atherosclerosis in patients with hypertension. Exp. Clin. Cardiol. *15*, 37–40.

27. Malenica, M., Prnjavorac, B., Bego, T., Dujic, T., Semiz, S., Skrbo, S., Gusic, A., Hadzic, A., and Causevic, A. (2017). Effect of cigarette smoking on haematological parameters in healthy population. Med. Arch. *71*, 132–136.

28. Hsieh, Y.-P., Chang, C.-C., Kor, C.-T., Yang, Y., Wen, Y.-K., and Chiu, P.-F. (2017). Mean corpuscular volume and mortality in patients with CKD. Clin. J. Am. Soc. Nephrol. *12*, 237–244.

29. Ueda, T., Kawakami, R., Horii, M., Sugawara, Y., Matsumoto, T., Okada, S., Nishida, T., Soeda, T., Okayama, S., Somekawa, S., et al. (2013). High mean corpuscular volume is a new indicator of prognosis in acute decompensated heart failure. Circ. J. *77*, 2766–2771.

30. Seppä, K., Heinilä, K., Sillanaukee, P., and Saarni, M. (1996). Evaluation of macrocytosis by general practitioners. J. Stud. Alcohol *57*, 97–100.

31. Tanaka, M., Okada, H., Hashimoto, Y., Kumagai, M., Nishimura, H., and Fukui, M. (2020). Combined effect of hemoglobin and mean corpuscular volume levels on incident metabolic syndrome: a population-based cohort study. Clin. Nutr. ESPEN *40*, 314–319.

32. Vogel, J., Kiessling, I., Heinicke, K., Stallmach, T., Ossent, P., Vogel, O., Aulmann, M., Frietsch, T., Schmid-Schönbein, H., Kuschinsky, W., and Gassmann, M. (2003). Transgenic mice overexpressing erythropoietin adapt to excessive erythrocytosis by regulating blood viscosity. Blood *102*, 2278–2284.

33. Shiga, T., Maeda, N., and Kon, K. (1990). Erythrocyte rheology. Crit. Rev. Oncol. Hematol. *10*, 9–48.

34. Dhas, Y., Banerjee, J., and Mishra, N. (2020). Blood viscosity, glycemic markers and blood pressure: a study in middle-aged normotensive and hypertensive type 2 diabetics. Indian J. Clin. Biochem. *35*, 102–108.

35. Rezende, S.M., Lijfering, W.M., Rosendaal, F.R., and Cannegieter, S.C. (2014). Hematologic variables and venous thrombosis: red cell distribution width and blood monocyte count are associated with an increased risk. Haematologica *99*, 194–200.

36. Elsworth, B.L., Lyon, M.S., Alexander, T., Liu, Y., Matthews, P., Hallett, J., Bates, P., Palmer, T., Haberland, V., Smith, G.D., et al. (2020). The MRC IEU OpenGWAS data infrastructure. Preprint at bioRxiv. https://doi.org/10.1101/2020.08.10.244293.

37. Pilling, L.C., Atkins, J.L., Duff, M.O., Beaumont, R.N., Jones, S.E., Tyrrell, J., Kuo, C.-L., Ruth, K.S., Tuke, M.A., Yaghootkar, H., et al. (2017). Red blood cell distribution width: genetic evidence for aging pathways in 116, 666 volunteers. PLoS One *12*, e0185083.

38. NOMESCO, N. (2009). NOMESCO Classification of Surgical Procedures (NCSP), Version 1.14.

39. Evangelou, E., Warren, H.R., Mosen-Ansorena, D., Mifsud, B., Pazoki, R., Gao, H., Ntritsos, G., Dimou, N., Cabrera, C.P., Karaman, I., et al. (2018). Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nat. Genet. *50*, 1412–1425.

40. Zheng, J., Richardson, T.G., Millard, L.A.C., Hemani, G., Elsworth, B.L., Raistrick, C.A., Vilhjalmsson, B., Neale, B.M., Haycock, P.C., Smith, G.D., and Gaunt, T.R. (2018). PhenoSpD: an integrated toolkit for phenotypic correlation estimation and multiple testing correction using GWAS summary statistics. GigaScience *7*, 1–10.

41. Burgess, S., Butterworth, A., and Thompson, S.G. (2013). Mendelian randomization analysis with multiple genetic variants using summarized data. Genet. Epidemiol. *37*, 658–665.

42. Sanderson, E., Davey Smith, G., Windmeijer, F., and Bowden, J. (2019). An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int. J. Epidemiol. *48*, 713–727.

43. Burgess, S., and Thompson, S.G. (2015). Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. Am. J. Epidemiol. *181*, 251–260.

44. Georgi, B., Mielke, J., Chaffin, M., Khera, A.V., Gelis, L., Mundl, H., van Giezen, J.J.J., Ellinor, P., Kathiresan, S., Ziegelbauer, K., and Freitag, D.F. (2019). Leveraging human genetics to estimate clinical risk reductions achievable by inhibiting factor XI. Stroke *50*, 3004–3012.

45. Burgess, S., and Thompson, S.G. (2021). Mendelian Randomization: Methods for Causal Inference Using Genetic Variants (CRC Press).

46. Morrison, J., Knoblauch, N., Marcus, J.H., Stephens, M., and He, X. (2020). Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. Nat. Genet. *52*, 740–747.

47. Hemani, G., Zheng, J., Elsworth, B., Wade, K.H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., et al. (2018). The MR-base platform supports systematic causal inference across the human phenome. Elife *7*. e34408–29.

48. Bowden, J., Spiller, W., Del Greco M, F., Sheehan, N., Thompson, J., Minelli, C., and Davey Smith, G. (2018). Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. Int. J. Epidemiol. *47*, 2100.

49. Grant, A.J., and Burgess, S. (2021). Pleiotropy robust methods for multivariable Mendelian randomization. Stat. Med. *40*, 5813–5830.

50. Slob, E.A.W., and Burgess, S. (2020). A comparison of robust Mendelian randomization methods using summary data. Genet. Epidemiol. *44*, 313–329.

51. Verbanck, M., Chen, C.Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat. Genet. *50*, 693–698.

52. Sanderson, E., Spiller, W., and Bowden, J. (2021). Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. Stat. Med. *40*, 5434–5452.

53. Burgess, S. MendelianRandomization V0 . 2 . 2 : An R Package for Performing Mendelian Randomization Analyses Using Summarized Data Created and Maintained by Olena Yavorska and Stephen Burgess.

54. Song, W., Qian, W., Wang, W., Yu, S., and Lin, G.N. (2021). Mendelian randomization studies of brain MRI yield insights into the pathogenesis of neuropsychiatric disorders. BMC Genom. *22*, 342–411.

55. Greco M, F.D., Minelli, C., Sheehan, N.A., Thompson, J.R., and Thompson, J.R. (2015). Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. Stat. Med. *34*, 2926–2940.

56. Higgins, J.P.T., Thompson, S.G., Deeks, J.J., and Altman, D.G. (2003). Measuring inconsistency in meta-analyses. Bmj *327*, 557–560.

57. Gotoh, S., Hata, J., Ninomiya, T., Hirakawa, Y., Nagata, M., Mukai, N., Fukuhara, M., Ikeda, F., Ago, T., Kitazono, T., and Kiyohara, Y. (2015). Hematocrit and the risk of cardiovascular disease in a Japanese community: the Hisayama Study. Atherosclerosis *242*, 199–204.

58. Bowden, J., Del Greco M, F., Minelli, C., Davey Smith, G., Sheehan, N., and Thompson, J. (2017). A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat. Med. *36*, 1783–1802.

59. Burgess, S., Foley, C.N., and Zuber, V. (2018). Inferring causal relationships between risk factors and outcomes from genome-wide association study data. Annu. Rev. Genom. Hum. Genet. *19*, 303–327.

60. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., and Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. *10*, e1004383.

61. Wakefield, J. (2009). Bayes factors for genome-wide association studies: comparison with P-values. Genet. Epidemiol. *33*, 79–86.

62. Wallace, C. (2021). A more accurate method for colocalisation analysis allowing for multiple causal variants. PLoS Genet. *17*, e1009440.

63. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. Nature *466*, 707–713.

64. Al-Soufi, L., and Costas, J. (2021). Colocalization of association signals at nicotinic acetylcholine receptor genes between schizophrenia and smoking traits. Drug Alcohol Depend. *220*, 108517.

65. Liu, Q., Pan, J., Berzuini, C., Rutter, M.K., and Guo, H. (2020). Integrative analysis of Mendelian randomization and Bayesian colocalization highlights four genes with putative BMI-mediated causal pathways to diabetes. Sci. Rep. *10*, 7476–7479.

66. Richmond, R.C., Hemani, G., Tilling, K., Davey Smith, G., and Relton, C.L. (2016). Challenges and novel approaches for investigating molecular mediation. Hum. Mol. Genet. *25*, R149–R156.

67. Carter, A.R., Sanderson, E., Hammerton, G., Richmond, R.C., Davey Smith, G., Heron, J., Taylor, A.E., Davies, N.M., and Howe, L.D. (2021). Mendelian randomisation for mediation analysis: current methods and challenges for implementation. Eur. J. Epidemiol. *36*, 465–478.

68. Sanderson, E. (2021). Multivariable Mendelian randomization and mediation. Cold Spring Harb. Perspect. Med. *11*, a038984.

69. VanderWeele, T.J. (2016). Mediation analysis: a practitioner's guide. Annu. Rev. Publ. Health *37*, 17–32.

70. Hesterberg, T. (2011). Bootstrap. WIREs. Comp. Stat. *3*, 497–526.

71. Lin, Z., Deng, Y., and Pan, W. (2021). Combining the strengths of inverse-variance weighting and Egger regression in Mendelian randomization using a mixture of regressions model. PLoS Genet. *17*, e1009922.

72. Dixon, P., Hollingworth, W., Harrison, S., Davies, N.M., and Davey Smith, G. (2020). Mendelian Randomization analysis of the causal effect of adiposity on hospital costs. J. Health Econ. *70*, 102300.

73. Burgess, S., and Thompson, S.G. (2017). Interpreting findings from Mendelian randomization using the MR-Egger method. Eur. J. Epidemiol. *32*, 377–389.

74. Bowden, J. (2017). Misconceptions on the use of MR-Egger regression and the evaluation of the InSIDE assumption. Int. J. Epidemiol. *46*, 2097–2099.

75. Thiese, M.S., Ronna, B., and Ott, U. (2016). P value interpretations and considerations. J. Thorac. Dis. *8*, E928–E931.

76. By, C.C. (2021). Correcting for Weak and Pleiotropic Instruments in Two-Sample University of Bristol - Explore Bristol Research Testing and Correcting for Weak and Pleiotropic Instruments in Two-Sample Multivariable Mendelian Randomisation, *40*, pp. 5434–5452.

77. Kamat, M.A., Blackshaw, J.A., Young, R., Surendran, P., Burgess, S., Danesh, J., Butterworth, A.S., and Staley, J.R. (2019). PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics *35*, 4851–4853.

78. Burgess, S., Davies, N.M., and Thompson, S.G. (2016). Bias due to participant overlap in two-sample Mendelian randomization. Genet. Epidemiol. *40*, 597–608.

79. Rees, J.M.B., Wood, A.M., Dudbridge, F., and Burgess, S. (2019). Robust methods in Mendelian randomization via penalization of heterogeneous causal estimates. PLoS One *14*. e0222362–24.

80. Yan, Z., Fan, Y., Meng, Z., Huang, C., Liu, M., Zhang, Q., Song, K., and Jia, Q. (2019). The relationship between red blood cell distribution width and metabolic syndrome in elderly Chinese: a cross-sectional study. Lipids Health Dis. *18*, 34–39.

81. Söderholm, M., Borné, Y., Hedblad, B., Persson, M., and Engström, G. (2015). Red cell distribution width in relation to incidence of stroke and carotid atherosclerosis: a population-based cohort study. PLoS One *10*, 1–14.

82. Je, S.J., Suh, C., and Chi, H.S. (1992). Increase of red blood cell's mean corpuscular volume in the elderly normal population of Korea. Korean J. Hematol. *27*, 227–231.

83. Sun, L., Zhang, C., Ju, Y., Tang, B., Gu, M., Pan, B., Guo, W., and Wang, B. (2019). Mean corpuscular volume predicts in-stent restenosis risk for stable coronary artery disease patients receiving elective percutaneous coronary intervention. Med. Sci. Mon. Int. Med. J. Exp. Clin. Res. *25*, 3976–3982.

84. Yoon, H.-J., Kim, K., Nam, Y.-S., Yun, J.-M., and Park, M. (2016). Mean corpuscular volume levels and all-cause and liver cancer mortality. Clin. Chem. Lab. Med. *54*, 1247–1257.

85. Li, Y., Chen, Y., Chen, J., Chiou, J., Chen, T., and Lai, L. (2020). Association between mean corpuscular volume and cognitive impairment in an 8-year cohort study in the community-dwelling elderly: Epidemiology/Risk and protective factors in MCI and dementia. Alzheimer's. Dementia *16*, e039280.

86. Cazzola, R., Russo-Volpe, S., Cervato, G., and Cestaro, B. (2003). Biochemical assessments of oxidative stress, erythrocyte membrane fluidity and antioxidant status in professional soccer players and sedentary controls. Eur. J. Clin. Invest. *33*, 924–930.

87. Yang, R., Wang, A., Ma, L., Su, Z., Chen, S., Wang, Y., Wu, S., and Wang, C. (2018). Hematocrit and the incidence of stroke: a prospective, population-based cohort study. Therapeut. Clin. Risk Manag. *14*, 2081–2088.

88. Chan, K.H.K., Huang, R., Liu, J., Wan, T.K., and Woo, Y.M.P. (2022). Method and system for machine learning and deep learning based assessment of stroke.

89. Hewitt, J., Walters, M., Padmanabhan, S., and Dawson, J. (2016). Cohort profile of the UK Biobank: diagnosis and characteristics of cerebrovascular disease. BMJ Open *6*. e009161–6.