Genetically Elevated LDL Associates with Lower Risk of Intracerebral Hemorrhage

Guido J. Falcone, MD, ScD, MPH,[1†] Elayna Kirsch, BA,[1†] Julian N. Acosta, MD,[1] Rommell B. Noche, MS,[1] Audrey Leasure, BS,[1] Sandro Marini, MD,[2] Jaeyoon Chung, PhD,[2] Magdy Selim, MD, PhD,[3] James F. Meschia, MD,[4] Devin L. Brown, MD, MS,[5] Bradford B. Worrall, MD, MSc,[6] David L. Tirschwell, MD, MSc,[7] Jeremiasz M. Jagiella, MD, PhD,[8] Helena Schmidt, MD,[9] Jordi Jimenez-Conde, MD, PhD,[10,11] Israel Fernandez-Cadenas, PhD,[12] Arne Lindgren, MD,[13,14] Agnieszka Slowik, MD, PhD,[8] Dipender Gill, MD,[15] Michael Holmes, MBBS, PhD,[16,17] Chia-Ling Phuah, MD, MMSc,[18] Nils H. Petersen, MD, MSc,[1] Charles N. Matouk, MD,[19] Murat Gunel, MD,[19] Lauren Sansing, MD, MSc,[20] Derrick Bennett, PhD, CStat,[17] Zhengming Chen, DPhil,[17] Luan L. Sun, DPhil,[21] Robert Clarke, MD,[1] Robin G. Walters, PhD,[16,17] Thomas M. Gill, MD,[22] Alessandro Biffi, MD,[2,23,24,25] Sekar Kathiresan, MD,[2,23,27] Carl D. Langefeld, PhD,[28] Daniel Woo, MD, MSc,[29] Jonathan Rosand, MD, MSc,[2,23,26,30] Kevin N. Sheth, MD,[1†] and Christopher D. Anderson, MD, MMSc,[2,23,26,30†]

For the International Stroke Genetics Consortium

View this article online at wileyonlinelibrary.com. DOI: 10.1002/ana.25740

Received Sep 27, 2019, and in revised form Apr 1, 2020. Accepted for publication Apr 3, 2020.

Address correspondence to Guido J. Falcone, 15 York Street, LLCI Room 1004D, P.O. Box 208018, New Haven, CT 06510. E-mail: guido.falcone@yale.edu; Christopher D. Anderson, 185 Cambridge Street, CPZN 6818, Boston, MA 02114. E-mail: cdanderson@mgh.harvard.edu

[†]G.J.F., E.K., K.N.S., and C.D.A. contributed equally.

From the [1]Division of Neurocritical Care & Emergency Neurology, Department of Neurology, Yale School of Medicine, New Haven, CT; [2]Center for Genomic Medicine, Massachusetts General Hospital (MGH), Boston, MA; [3]Department of Neurology, Harvard Medical School, Beth Israel Deaconess Medical Center, Boston, MA; [4]Department of Neurology, Mayo Clinic, Jacksonville, FL; [5]Stroke Program, Department of Neurology, University of Michigan Health System, Ann Arbor, MI; [6]Department of Neurology and Public Health Sciences, University of Virginia Health System, Charlottesville, VA; [7]Stroke Center, Harborview Medical Center, University of Washington, Seattle, WA; [8]Department of Neurology, Jagiellonian University Medical College, Kraków, Poland; [9]Institute of Molecular Biology and Medical Biochemistry, Medical University Graz, Graz, Austria; [10]Neurovascular Research Unit, Department of Neurology, Institut Municipal d'Investigacio' Medica-Hospital del Mar, Universitat Autonoma de Barcelona, Barcelona, Spain; [11]Program in Inflammation and Cardiovascular Disorders, Institut Municipal d'Investigacio' Medica-Hospital del Mar, Universitat Autonoma de Barcelona, Barcelona, Spain; [12]Neurovascular Research Laboratory and Neurovascular Unit, Institut de Recerca, Hospital Vall d'Hebron, Universitat Autonoma de Barcelona, Barcelona, Spain; [13]Department of Clinical Sciences Lund, Neurology, Lund University, Lund, Sweden; [14]Department of Neurology, Skåne University Hospital, Lund, Sweden; [15]Department of Epidemiology and Biostatistics and Department of Stroke Medicine, Imperial College London, London, UK; [16]Medical Research Council Population Health Research Unit, University of Oxford, Oxford, UK; [17]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, Medical Research Council Population Health Research Unit, University of Oxford, Oxford, UK; [18]Department of Neurology, Washington University School of Medicine in St. Louis, St. Louis, MO; [19]Department of Neurosurgery, Yale School of Medicine, New Haven, CT; [20]Division of Vascular Neurology and Stroke, Department of Neurology, Yale School of Medicine, New Haven, CT; [21]Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK; [22]Department of Internal Medicine, Geriatric Medicine, Yale School of Medicine, New Haven, CT; [23]Program in Medical and Population Genetics, Broad Institute, Cambridge, MA; [24]Division of Behavioral Neurology, Department of Neurology, MGH, Boston, MA; [25]Division of Psychiatry, Department of Psychiatry, MGH, Boston, MA; [26]Department of Neurology, MGH, Boston, MA; [27]Cardiovascular Disease Prevention Center, MGH, Boston, MA; [28]Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC; [29]Department of Neurology, University of Cincinnati College of Medicine, Cincinnati, OH; and [30]Henry and Allison McCance Center for Brain Health, MGH, Boston, MA

Additional supporting information can be found in the online version of this article.

Objective: Observational studies point to an inverse correlation between low-density lipoprotein (LDL) cholesterol levels and risk of intracerebral hemorrhage (ICH), but it remains unclear whether this association is causal. We tested the hypothesis that genetically elevated LDL is associated with reduced risk of ICH.
Methods: We constructed one polygenic risk score (PRS) per lipid trait (total cholesterol, LDL, high-density lipoprotein [HDL], and triglycerides) using independent genomewide significant single nucleotide polymorphisms (SNPs) for each trait. We used data from 316,428 individuals enrolled in the UK Biobank to estimate the effect of each PRS on its corresponding trait, and data from 1,286 ICH cases and 1,261 matched controls to estimate the effect of each PRS on ICH risk. We used these estimates to conduct Mendelian Randomization (MR) analyses.
Results: We identified 410, 339, 393, and 317 lipid-related SNPs for total cholesterol, LDL, HDL, and triglycerides, respectively. All four PRSs were strongly associated with their corresponding trait (all $p < 1.00 \times 10^{-100}$). While one SD increase in the PRSs for total cholesterol (odds ratio [OR] = 0.92; 95% confidence interval [CI] = 0.85–0.99; $p = 0.03$) and LDL cholesterol (OR = 0.88; 95% CI = 0.81–0.95; $p = 0.002$) were inversely associated with ICH risk, no significant associations were found for HDL and triglycerides (both $p > 0.05$). MR analyses indicated that 1mmol/L (38.67mg/dL) increase of genetically instrumented total and LDL cholesterol were associated with 23% (OR = 0.77; 95% CI = 0.65–0.98; $p = 0.03$) and 41% lower risks of ICH (OR = 0.59; 95% CI = 0.42–0.82; $p = 0.002$), respectively.
Interpretation: Genetically elevated LDL levels were associated with lower risk of ICH, providing support for a potential causal role of LDL cholesterol in ICH.

ANN NEUROL 2020;88:56–66

Novel therapies are needed for spontaneous, nontraumatic intracerebral hemorrhage (ICH), a disease responsible for 50% of stroke-related deaths and disability with no proven acute treatments.[1] Several complementary lines of evidence highlight the importance of lipid metabolism as a promising pathophysiological pathway for risk prediction and therapeutic strategies. The Stroke Prevention by Aggressive Reduction in Cholesterol Levels (SPARCL) trial reported that statin treatment in surviving adults with a first ischemic stroke reduced the risk of recurrent ischemic stroke but increased the risk of ICH.[2] However, this trial could not exclude possible pleiotropic effects of statins, whereby the effects on ICH may be independent of the effects of statins on lipids.[3] In addition, clinical trials investigating statins as the primary cardiovascular disease prevention strategy have yielded inconsistent results for ICH risk.[4–6]

Results from several observational studies evaluating data from a large number of ICH cases also reported inverse associations between lipid levels and ICH risk. The Genetic and Environmental Risk Factors for Hemorrhagic Stroke (GERFHS) study reported a one third lower ICH risk among study participants with a medical history of hypercholesterolemia.[7] Other observational studies reported that higher low-density lipoprotein cholesterol (LDL-C) levels were correlated with ICH severity and 3-month clinical outcome.[8] Although promising, these studies are limited by the observational nature of the underlying design, which preclude the possibility of establishing causality.

Population genetics provides powerful tools to overcome such limitations in causal inference. Genetic variants known to associate with lipid levels can be used as instruments to evaluate the causal relationship between different lipid fractions and risk of ICH.[9] These genetic variants are randomly distributed during meiosis and are ought to be exempt from confounding by environmental exposures.[9–11] A recent report from the China Kadoorie Biobank involving ~5,000 ICH cases demonstrated concordant effects estimates between the observational and genetic analyses for LDL-C and ICH risk in Chinese adults, thereby providing strong support for the causal relevance for this association in this ethnic group. However, in contrast with highly significant inverse observational associations of directly measured LDL-C and ICH risks, the associations of genetically instrumented LDL-C and ICH risk were not statistically significant in this study.[12]

To overcome the limitations of previous randomized trials, observational studies, and genetic analyses, we conducted a multistage genetic association study that combined polygenic risk score[13] (PRS) and Mendelian Randomization[9] (MR) analyses to test the hypothesis that genetically elevated lipid levels are associated with a lower risk of ICH. We separately evaluated the associations of genetically instrumented differences in total cholesterol (TC), LDL-C, high-density lipoprotein cholesterol (HDL-C), and triglycerides. Because of the known differences in underlying biology according to the location of ICH within the brain, we conducted stratified analyses based on hemorrhage location.[14]

Methods

Study Design

We conducted a 3-stage genetic study in participants from European ancestry. All study stages utilized publicly available individual-level data accessible through the National Institute of Health database of Genotypes and Phenotypes and UK Biobank. All studies had approval from the local institutional review board or ethics committee at each participating

institution. Informed consent was obtained from all study participants or their legally authorized representatives, or consent was waived via protocol-specific allowance. All study participants had available genome-wide genotyping data, allowing the implementation of principal component analysis to confirm ancestry and account for population stratification. In stage 1, we constructed 4 PRSs for TC, LCL-C, HDL-C, and triglycerides using genetic variants known to associate with each of these traits in previous studies.[15–17] We estimated the effect of each PRS on its corresponding lipid trait using data from the UK Biobank. In stage 2, we conducted an individual participant data meta-analysis of genetic studies of ICH to evaluate the association among the 4 PRSs created in stage 1 and risk of ICH. This stage included individual-level data from 3 case-control genetic studies: the Genetics of Cerebral Hemorrhage on Anticoagulation (GOCHA) study,[18] International Stroke Genetics Consortium ICH (ISGC-ICH) study,[19] and the GERFHS study.[7] In stage 3, we combined the effect estimates from stage 1 (associations between each PRS and its corresponding lipid fraction) and stage 2 (associations between each PRS and ICH risk) to conduct MR analyses of genetically instrumented lipid levels and ICH risk.

Blood Lipids in the UK Biobank

Stage 1 utilized data from the UK Biobank, a prospective population-based cohort study that recruited 500,000 community-dwelling individuals aged 40 to 69 years between 2006 and 2010 from across the United Kingdom. Study participants underwent multiple baseline physical measures, provided blood, urine and saliva samples for different analysis, provided detailed information about themselves, and agreed to have their health followed.[20] We used recently released values for lipid traits, which were measured using a Beckman Colter AU5800 clinical chemistry analyzer. Analyzer performance was verified continually throughout the project.

Ascertainment of ICH Cases and ICH-Free Controls

ICH cases included in stage 2 were defined as new and acute (<24 hours) neurological deficits with consistent findings in neuroimaging. Cases were aged >55 years in GOCHA, and >18 years in GERFHS and the ISGC-ICH. Patients were excluded if they were taking anticoagulants (anti-aggregants were permitted) or had head trauma, hemorrhagic conversion of an ischemic stroke, intracerebral tumor, intracerebral vascular malformation, vasculitis, or any other cause of secondary ICH. Controls included in stage 2 were ICH-free individuals enrolled at the same study sites as the cases and followed the same age and ethnicity criteria. Controls were sampled by random digit dialing in GERFHS and by random selection from ambulatory clinics in GOCHA and ISGC-ICH.

Neuroimaging Analysis

For ICH cases included in stage 2, stroke neurologists or neuroradiologists at each participating site confirmed the diagnosis and, following the known differences in underlying biology, classified each case as lobar or nonlobar according to location.[21] ICH originating in the corticosubcortical junction was defined as lobar, whereas ICH selectively involving the thalamus, internal capsule, basal ganglia, brainstem, or cerebellum was defined as nonlobar.

Genetic Data

Study participants were genotyped using the UK Biobank Axiom Array (UK Biobank Study), Illumina Human Hap610-Quad (GOCHA and ISGC-ICH), and Affymetrix 6.0 (GERFHS). Standard quality control procedures[22] were implemented separately for each participating study. Single nucleotide polymorphisms (SNPs) with palindromic alleles (A/T or C/G), a genotype call rate <95%, significant difference in missingness between cases and controls ($p < 0.05$), deviation from Hardy–Weinberg Equilibrium ($p < 1 \times 10^{-6}$), or minor allele frequency (MAF) <1% were removed. Individuals with a genotype call rate <95%, inconsistency between self-reported and genotyped sex, an inferred first or second degree relative in the sample, and a genomewide heterozygosity F-statistic >5 times the SD were filtered out from the analysis. Principal component analysis was implemented to account for population structure.[23] After quality control and principal component analyses, genetic data were prephased and imputed to 1000 Genomes integrated reference panels (phase 3 integrated variant set release in National Center for Biotechnology Information [NCBI] build 37).[24] Postimputation filters included MAF <1%, an information score <0.7, and missing estimates in one or more studies.

Statistical Analyses

We present discrete variables as counts (percentage) and continuous variables as mean (SD) or median (interquartile range [IQR]), as appropriate.

Stage 1. Derivation of Lipid-Related PRSs. We used PRSs to model each individual's genetic load of lipid-related risk alleles. To build these PRSs, we queried the Genomewide Association Study (GWAS) catalog and reviewed published GWAS of lipids.[15–17] Following similar recent analyses, we selected independent ($r^2 < 0.3$) and common (MAF >5%) SNPs associated with at least one lipid trait at $p < 5 \times 10^{-8}$.[15] All selected SNPs were aligned to the GRCh37 assembly of the human genome and, for each

SNP, the allele associated with an increase in lipid levels was identified and utilized as the tested allele in downstream analyses. The epsilon variants within apolipoprotein E (*APOE*) were not included in these PRSs, as they are not captured by commercially available genotyping arrays and there is a plausible alternative pathway via cerebral amyloid angiopathy that could mediate its association with ICH.[25] The PRS for each individual is the sum of the product of the risk allele counts for each locus multiplied by the allele's reported effect on the corresponding lipid level. To assure common directionality of effects, the allele associated with higher lipid levels was selected as the effect allele during scoring. One PRS per lipid trait was generated (TC, LDL-C, HDL-C, and triglycerides). All 4 PRSs were standardized (by subtracting the mean and dividing by the SD) and entered as continuous predictors into regression models. With this approach, the beta for the PRS can be interpreted as the change in the dependent variable per 1 SD increase of the PRS. The association between each PRS and its corresponding lipid trait was evaluated using linear regression, adjusting for age, sex, and principal components 1 to 4. The primary analysis was restricted to unrelated study participants of genetically determined European ancestry who were not taking lipid-lowering medications. In secondary analyses, we followed a less restrictive approach and included all study participants of self-reported European ancestry.

Stage 2. Association Between each PRS and ICH Risk. We evaluated the association between each PRS and risk of ICH via logistic regression adjusting for age, sex, and principal components 1 to 4. Analyses were completed separately in each genetic study and pooled in meta-analysis using fixed-effects (primary analysis) and random effects (secondary analysis) approaches with evaluation of heterogeneity via Cochrane's Q (with corresponding p) and I^2. In sensitivity analyses, we excluded genetic variants within *CETP*, a locus that powerfully modifies lipid levels previously shown to be associated with ICH risk. To account for recognized differences in underlying causative ICH mechanism by location within the brain, lipid traits with significant associations were taken forward to stratified analyses based on hemorrhage location (lobar or nonlobar).

Stage 3. Mendelian Randomization Analyses. The causal relationship between genetically determined lipid levels and ICH risk was evaluated via MR analyses using each PRS as an instrument. The primary MR analysis utilized the ratio method combining the point estimates and standards errors from stage 1 (denominator) and stage 2 (numerator). In secondary analyses, we implemented other MR methods usually used with summary level data, including inverse-variance weighted, weighted median, MR-Egger, and Mendelian Randomization Pleiotropy Residual Sum and Outlier (MR-PRESSO) analyses. To confirm the validity of our results, we implemented MR analyses of genetically instrumented cholesterol levels and risk of ischemic stroke using estimates for lipids levels from the UK Biobank and estimates for ischemic stroke risk from the MEGASTROKE[26] consortium; these results were compared with previously reported MR studies for the same analysis.

Software. We used the GWAS Catalog to identify genetic variants related to lipid levels, PLINK for quality control procedures and generation of PRSs,[27] EIGENSTRAT for principal component analysis,[28] SHAPEIT for genotype prephasing,[29] IMPUTE2 for imputation,[30] and Rstudio (version 1.1.453) for association testing, meta-analysis, and MR analysis.[31]

Results

Selected population characteristics are presented in Table 1.

Stage 1: Derivation of Cholesterol-Related PRSs

We identified 1,459 common (MAF >5%) genetic variants reported by prior studies as strongly associated with one or more of the lipid traits of interest. These common variants included 410 SNPs for TC, 339 for LDL-C, 393 for HDL-C, and 317 for triglycerides (Tables S1–S4). We built 4 different PRSs, one for each lipid trait, and evaluated their associations with their corresponding trait in the UK Biobank. All 4 PRSs showed highly significant associations with their corresponding lipid trait, both in the primary analysis considering 316,428 (mean age 68 years [SD = 8], and 170,871 women [54%]) unrelated individuals of European ancestry not on lipid-lowering medications (all $p < 1 \times 10^{-100}$), and in the secondary analysis not applying any exclusion criteria (all $p < 1 \times 10^{-100}$; Table 2).

Stage 2: Association Between each PRS and ICH Risk

A total of 1,286 ICH cases (mean age 71 years [SD = 13], and 593 women [46%]) and 1,261 ICH-free controls (mean age 68 years [SD = 14], and 613 women [49%]) from the GOCHA, ISGC ICH, and GERFHS studies were included in association testing (see Table 1). For TC, each additional SD increase of the corresponding PRS was associated with an 8% lower ICH risk (odds ratio [OR] = 0.92; 95% confidence interval [CI] = 0.85–0.99; p = 0.03). When evaluating specific lipid fractions, we found that each additional SD increase of the LDL-C-based PRS was

TABLE 1. Studies Included in this Analysis

Characteristic	UK Biobank	GOCHA	ISGC-ICH study	GERFHS
Analytical stage	Association cholesterol level ~ PRS	Association ICH risk ~ PRS	Association ICH risk ~ PRS	Association ICH risk ~ PRS
Study design	Cohort	Case / control	Case / control	Case / control
Study participants	316,428	277 / 248	563 / 523	446 / 490
Age, mean (SD)	68 (8)	73 (10) / 72 (8)	71 (14) / 66 (16)	70 (14) / 68 (13)
Female sex n, %	170,871 (54)	130 (47) / 123 (50)	252 (45) / 255 (49)	211 (47) / 235 (48)
Genotyping platform	Affymetrix UK Biobank array	Illumina HumanHap550	Illumina HumanHap550	Affymetrix 6.0
Genotyped SNPs	820,967	527,508	527,508	580,491
Imputed SNPs	73,355,667	7,965,700	7,965,700	7,967,430

GERFHS = Genetic and Environmental Risk Factors for Hemorrhagic Stroke; GOCHA = Genetics of Cerebral Hemorrhage with Anticoagulation; ICH = intracerebral hemorrhage; ISGC = International Stroke Genetics Consortium; PRS = polygenic risk score; SNPs = single nucleotide polymorphisms.

associated with a 12% lower risk of ICH (OR = 0.88; 95% CI = 0.81–0.95; p = 0.002; Table 3). Similar results were obtained when utilizing random-effects meta-analyses. These associations remained significant after removing *CETP* (Table 4). We did not find significant associations for the PRSs based on HDL-C or triglycerides (both $p > 0.05$).

TABLE 2. Association Results for 4 PRS with their Corresponding Trait in the UK Biobank

Lipid trait PRS	Independent SNPs in PRS	UK Biobank Effective sample size	Mean increase in cholesterol trait per 1-SD increase in PRS	Standard error	Variance explained	p
Primary analysis[a]						
Total cholesterol	410	316,428	0.33mmol/L (12.76mg/dL)	0.0018	9.33%	$<1 \times 10^{-100}$
LDL cholesterol	339	315,841	0.24mmol/L (9.28mg/dL)	0.0014	8.38%	$<1 \times 10^{-100}$
HDL cholesterol	393	289,349	0.11mmol/L (4.25mg/dL)	0.0006	8.17%	$<1 \times 10^{-100}$
Triglycerides	317	316,174	0.22mmol/L (19.49mg/dL)	0.0017	4.8%	$<1 \times 10^{-100}$
Secondary analysis[b]						
Total cholesterol	410	437,676	0.26mmol/L (10.05mg/dL)	0.0017	5.21%	$<1 \times 10^{-100}$
LDL cholesterol	339	436,867	0.19mmol/L (7.35mg/dL)	0.0013	4.72%	$<1 \times 10^{-100}$
HDL cholesterol	393	400,579	0.11mmol/L (4.25mg/dL)	0.0005	8.04%	$<1 \times 10^{-100}$
Triglycerides	317	437,331	0.23mmol/L (20.37mg/dL)	0.0015	5%	$<1 \times 10^{-100}$

[a]The primary analysis was restricted to unrelated study participant of genetically determined European ancestry who were not taking lipid-lowering medications.

[b]The secondary analysis included all study participants of self-reported European ancestry without any other filters.

HDL = high-density lipoprotein; LDL = low-density lipoprotein; PRS = polygenic risk score; SNP = single nucleotide polymorphism; UK = United Kingdom.

Conversion of mmol/L to mg/dL = for total cholesterol, LDL, and HDL 1mmol/L = 38.67mg/dL; for triglycerides, and 1mmol/L = 88.57mg/dL.

TABLE 3. Study-Specific and Meta-Analysis of Logistic Regression Results Modeling ICH Risk as a Function of Different PRS

	Total cholesterol		LDL cholesterol		HDL cholesterol		Triglycerides	
Study	OR (95% CI)	*p*	OR (95% CI)	*p*	OR (95% CI)	*p*	OR (95% CI)	*p*
GOCHA	0.95 (0.80–1.14)	0.59	0.93 (0.78–1.11)	0.41	1.12 (0.94–1.34)	0.20	0.95 (0.80–1.14)	0.59
ISGC-ICH	0.93 (0.82–1.05)	0.24	0.88 (0.77–0.99)	0.04	1.15 (1.01–1.30)	0.03	1.01 (0.89–1.14)	0.84
GERFHS	0.88 (0.77–1.01)	0.07	0.85 (0.75–0.97)	0.02	0.99 (0.86–1.12)	0.84	1.19 (1.04–1.36)	0.009
Fixed effects Meta-analysis	0.92 (0.85–0.99)	0.03	0.88 (0.81–0.95)	0.002	1.10 (1.00–1.21)	0.06	1.11 (0.98–1.23)	0.14
Random effects meta-analysis	0.92 (0.84–0.99)	0.03	0.88 (0.81–0.95)	0.002	1.08 (0.98–1.19)	0.13	1.05 (0.93–1.20)	0.42
Heterogeneity	I^2 = 0%	0.77	I^2 = 0%	0.75	I^2 = 32%	0.23	I^2 = 60%	0.08

CI = confidence interval; GERFHS = Genetic and Environmental Risk Factors for Hemorrhagic Stroke; GOCHA = Genetics of Cerebral Hemorrhage with Anticoagulation; HDL = high-density lipoprotein; ICH = intracerebral hemorrhage; ISGC = International Stroke Genetics Consortium; LDL = low-density lipoprotein; OR = odds ratio; PRS = polygenic risk score.

TABLE 4. Meta-Analysis of Logistic Regression Results Modeling ICH Risk as a Function of Different PRS, Excluding *CETP* Variants

Lipid PRS	OR (95% CI)	*p*
Total cholesterol	0.91 (0.84–0.99)	0.03
LDL cholesterol	0.88 (0.81–0.96)	0.003
HDL cholesterol	1.12 (0,99–1.21)	0.08
Triglycerides	1.11 (0.98–1.23)	0.12

CI = confidence interval; HDL = high-density lipoprotein; ICH = intracerebral hemorrhage; LDL = low-density lipoprotein; OR = odds ratio; PRS = polygenic risk score.

Stage 3: Mendelian Randomization Analysis

The primary MR analysis implemented the ratio method utilizing the effect estimates obtained in stages 1 and 2. As shown in Table 5, each 1mmol/L (or 38.67mg/dL) increase of genetically instrumented TC was associated with a 23% reduction of ICH risk (OR = 0.77; 95% CI = 0.60–0.98; p = 0.03), whereas a 1mmol/L (or 38.67mg/dL) increase of genetically instrumented LDL-C was associated with a 41% reduction in this risk (OR = 0.59; 95% CI = 0.42–0.82; p = 0.002). These results remained unaltered when the effect of each PRS on its corresponding lipid trait was estimated without excluding any individuals from the UK Biobank (data not shown). Secondary analyses utilizing an $r^2 < 0.1$ yielded comparable results for LDL-C (OR = 0.62; 95% CI = 0.41–0.94; p = 0.02) and confirmed the direction

TABLE 5. MR Analysis of Genetically Instrumented Lipid Levels and Risk of ICH

		Total cholesterol		LDL cholesterol	
MR method	Instrument	OR (95% CI)	*p*	OR (95% CI)	*p*
Ratio method	Polygenic risk score using on individual level data	0.77 (0.6–0.98)	0.03	0.59 (0.42–0.82)	0.002
IVW	Multiple SNPs using summary level data	0.84 (0.72–0.99)	0.04	0.65 (0.52–0.82)	<0.001
Weighted median	Multiple SNPs using summary level data	0.95 (0.72–1.30)	0.74	0.79 (0.56–1.10)	0.20
MR-Egger (causal estimates)	Multiple SNPs using summary level data	0.87 (0.66–1.20)	0.33	0.72 (0.48–1.10)	0.10
MR-Egger (intercept)	Multiple SNPs using summary level data	1.00 (0.99–1.01)	0.81	1.00 (0.98–1.01)	0.59

CI = confidence interval; ICH = intracerebral hemorrhage; IVW = inverse variance weighted; LDL = low-density lipoprotein; MR = Mendelian Randomization; OR = odds ratio; SNPs = single nucleotide polymorphisms.

TABLE 6. Sensitivity MR Analyses Evaluating the Effect of Genetically Determined LDL Cholesterol on Risk of Different Types of Ischemic Stroke: Comparison of Results using Effect Estimates for Lipids Using GLGC and the UK Biobank

Ischemic stroke subtype	Lipid estimates from the GLGC *2019 Valdes-Marquez et al*[a]	Lipid estimates from the GLGC *2018 Hindy et al*[b]	Lipid estimates from the UK Biobank (this study)
Cardioembolic	1.06 (0.84–1.33)	0.99 (0.84–1.16)	1.05 (0.97–1.13)
Large artery	1.10 (0.82–1.47)	1.28 (1.07–1.53)	1.37 (1.24–1.51)
Small vessel	1.14 (0.88–1.48)	1.09 (0.93–1.28)	1.12 (1.03–1.22)

[a]Neurology. Mar 12, 2019;92 (11):e1176–e1187.
[b]Stroke. 2018 Apr;49 (4):820–827.
GLGC = Global Lipids Genetics Consortium; LDL = low-density lipoprotein; MR = Mendelian Randomization.

of effect for TC without reaching statistical significance (OR = 0.94; 95% CI = 0.87–1.02; p = 0.17). Secondary analyses utilizing other MR methods confirmed the direction of effect, although not all yielded statistical significance (see Table 5). There was no indication of pleiotropy for either TC or LDL-C (MR-egger intercepts and MR-PRESSO global test $p > 0.05$). MR analyses for LDL-C and risk of ischemic stroke utilizing the estimates for LDL-C from the UK Biobank yielded similar results to those reported by prior publications based on lipid estimates from the Global Lipid Genetics Consortium (Table 6).

Stratification Based on Location of the ICH Within the Brain

A total of 1,243 ICH cases (96%) had available information about the location of the hematoma within the brain. Of these, 539 (43%) had lobar ICH and 704 (56%) had nonlobar ICH. Location-specific analyses indicated that the association between the LDL-C PRS and ICH risk remained significant for both lobar (OR = 0.81; 95% CI = 0.73–0.89; $p < 0.001$) and nonlobar ICH (OR = 0.90; 95% CI = 0.82–0.99; p = 0.04; Table 7), whereas the association between ICH risk and the TC PRS was significant for lobar

TABLE 7. Location-Specific Results for ICH risk

	Lobar ICH n = 539 cases			Nonlobar ICH n = 704 cases		
Lipid trait	OR (95% CI)	p	Meta-analysis heterogeneity p	OR (95% CI)	p	Meta-analysis heterogeneity p
Polygenic risk score analysis[a]						
Total cholesterol	0.89 (0.80–0.99)	0.03	0.42	0.94 (0.85–1.08)	0.20	0.96
LDL cholesterol	0.81 (0.73–0.89)	<0.001	0.96	0.90 (0.82–0.99)	0.04	0.99
Mendelian randomization analysis[b]						
Total cholesterol	0.70 (0.51–0.96)	0.03	–	0.73 (0.62–1.11)	0.20	–
LDL cholesterol	0.41 (0.27–0.64)	<0.001	–	0.66 (0.44–0.97)	0.04	–

[a]Inverse variance fixed effects meta-analysis of logistic regression results for intracerebral hemorrhage (ICH) across Genetics of Cerebral Hemorrhage with Anticoagulation (GOCHA), International Stroke Genetics Consortium ICH (ISGC-ICH) genomewide association study (GWAS), and Genetic and Environmental Risk Factors for Hemorrhagic Stroke (GERFHS). For each study, the logistic regression model used ICH risk as the dependent variable and a polygenic risk score as the independent variable, adjusting for age, sex, and 4 principal components. The PRS were normalized and entered to the model as a continuous predictor. The OR represents the change in the odds of ICH per each additional SD of the PRS.
[b]Mendelian randomization results of genetically instrumented cholesterol levels using a polygenic risk score as the instrument. Each lipid fraction-specific analysis utilized the ratio method, taking the effect estimates for ICH ~ PRS (numerator) and lipid level ~ PRS (denominator).
CI = confidence intervals; ICH = intracerebral hemorrhage; LDL = low-density lipoprotein.; OR = odds ratio; PRS = polygenic risk score.

(OR = 0.89; 95% CI = 0.80–0.99; p = 0.03) but not nonlobar bleeds (OR = 0.94; 95% CI = 0.85–1.08; p = 0.20). MR analyses implementing the ratio method using these location-specific estimates indicated that genetically elevated LDL-C was associated with a decreased risk of ICH for both lobar (OR = 0.41; 95% CI = 0.27–0.64; $p < 0.001$) and nonlobar ICH (OR = 0.66; 95% CI = 0.44–0.97; p = 0.04), whereas genetically elevated TC was associated with a decreased risk of lobar (OR = 0.70; 95% CI = 0.51–0.96; p = 0.03) but not nonlobar ICH (OR = 0.73; 95% CI = 0.62–1.11; p = 0.20).

Discussion

We report the results of a multistage genetic association study that evaluated whether genetically instrumented levels of different lipid traits influence the risk of spontaneous ICH. We constructed 4 PRSs to model the aggregate genetic load of risk alleles for TC, LDL-C, HDL-C, and triglycerides; assessed for association between each PRS and its corresponding lipid trait; assessed for association between each PRS and ICH risk; and utilized the estimates obtained in prior steps to conduct MR analyses. We found that all 4 PRSs were robustly associated with their corresponding lipid trait and that the PRSs for TC and LDL-C were inversely associated with ICH risk. Analyses stratified by location indicated that these associations remained significant for both lobar and nonlobar ICH, with stronger associations for lobar bleeds. Of note, the genetically instrumented HDL-C and triglycerides levels were not associated with ICH risk.

Previous studies provided promising, but inconclusive, evidence on the relevance of LDL-C for risk of ICH. The evidence from randomized control trials of statins is inconsistent. The SPARCL trial, a study focused on the utilization of statins for secondary prevention after a first stroke or transient ischemic attack, found an unexpected increment in ICH risk as a side effect.[2] However, large meta-analyses of statin trials yielded conflicting conclusions for this question, with some finding similar associations[5,12] and others finding nulls results.[4] These inconsistencies may be driven by a lack of statistical power, as ICH is a rare event and statin trials, although large, do not accrue the necessary number of events to appropriately evaluate this relationship. From an observational perspective, the GERFHS study reported a reduction in ICH risk among study participants with a history of hypercholesterolemia.[7] In terms of genetic evidence, a candidate gene study focused on the powerful lipid regulatory gene *CETP* found an association between variants at this locus and ICH risk.[32] Although promising, each of these pieces of evidence has an important limitation: the inconsistency of results observed in clinical trials of statin treatment, the inability to draw causal conclusions in observational studies, and the candidate gene design of the *CETP* study.

The present study provides important additional evidence to support a causal role of LDL-C in risk of ICH. Genetic variants known to be associated with lipid levels can be used as instruments to evaluate a causal relationship between different lipid fractions and ICH risk.[33] We deployed 2 specific strategies to maximize the accuracy and power of this analytical strategy. First, all analytical steps used individual level phenotypic and genotypic data, permitting the utilization of rigorous quality control procedures and the implementation of sensitivity analyses to evaluate whether results were robust to different modeling strategies. Second, we estimated the effect of our instruments, the 4 lipid-related PRSs, on newly released data on lipid fractions from the UK Biobank. The sample size of this study (400,000+ study participants) maximizes the discovery power of the MR analysis by improving the precision of the estimates.

Beyond providing support for a causal role of lipid metabolism in ICH in Europeans, our results also support a specific role of LDL-C as the operative lipid trait mediating the observed inverse associations. Previous studies in Asians, who have lower LDL-C levels than Western populations, reported that this lipid fraction is the likely mediator underlying the inverse association between TC and ICH risk. A nested case-control study within the prospective China Kadoorie Biobank involving ~5,000 ICH cases reported that elevated levels of LDL-C were inversely associated with risk of ICH.[12] MR analyses in this study yielded concordant effect estimates, although these were not statistically significant, possibly due to the lower number of SNPs utilized to build the instrument (59 variants) and the European origin of the populations where these lipid-related SNPs had originally been identified. Our results confirm the role of LDL-C as the mediating lipid fraction and provides evidence supporting its role in persons of European ancestry. We acknowledge that, whereas concordant in the direction of effect, the point estimates for the MR analysis of LDL-C and ICH risk yielded by the present study (OR = 0.59; 95% CI = 0.42–0.82) are significantly more extreme than those reported in the China Kadoorie Biobank (OR = 0.89; 95% CI = 0.62–1.16). The discrepant results could reflect the lower mean age at ICH onset and lower mean LDL-C levels in Chinese compared with Europeans or between-population differences in the distribution of LDL-C, which could have biased such comparisons. Although the overlap in CIs between European and Chinese studies indicate that differences in the estimates between studies are not statistically significant, precise estimates of effects of LDL-C on ICH risk will require additional analyses in further studies involving larger numbers of ICH cases.

The independent replication of our findings constitutes an important next step to consolidate lipid metabolism as an actionable biological target in ICH

prevention. These follow-up studies will be greatly facilitated by increasingly available data from large biobanks and multipurpose repositories like dbGaP[34] and the European Genome-phenome Archive (EGA).[35] Another important future direction involves the clarification of the pathophysiology underlying the observed association. Histopathological evidence in humans suggests that lower cholesterol concentrations may increase the frailty and permeability of brain vessel walls, triggering arterionecrosis, microaneurysm formation and, ultimately, ICH.[36,37] Because our findings point to an effect that is present for both lobar and deep hemorrhages, it is possible that low lipid levels could work as an effect modifier of the risk conveyed by the underlying small vessel disease responsible for the bleed.

The results of this study prompt questions about the risk–benefit ratio of lowering LDL-C for risk of different stroke types. The China Kadoorie Biobank demonstrated equal and opposite proportional differences in risk of ischemic stroke and ICH for equivalent differences in LDL-C. Because the absolute number of ischemic stroke cases exceeded those of ICH by 4-fold, any beneficial effects of lowering LDL-C on ischemic stroke were likely to outweigh risks of ICH. In light of this evidence, it is reasonable to use extreme caution when evaluating possible applications of these results to clinical decision making.

An important limitation of our study was the inability to evaluate the effect of other lipid fractions. Alongside the vast majority of related studies, we evaluated TC, LDL-C, HDL-C, and triglycerides, the 4 lipid traits routinely used in clinical practice, and did not account for several other cardiovascular risk-stratifying lipid fractions, such as apolipoprotein levels.[38–40] A second important limitation is the absence of an independent dataset to replicate the association analysis between the lipid-related PRS and ICH risk. The relatively low incidence of ICH in Western populations limits the amount of appropriately ascertained cases within available genetic and location information. Nevertheless, the estimates observed in 3 different genetic studies of ICH were consistent with each other. In addition, because this study was not intended at risk loci discovery, it could be argued that independent replication for this specific analysis is not strictly needed. Finally, the limited available data on medical history and use of medication in ICH cases precluded any detailed analysis of possible interactions or confounding effects by these variables.

Summary

In conclusion, we report an inverse association between the genetic load of risk alleles for total and LDL-C and risk of ICH in persons of European ancestry. We also found that genetically instrumented higher total and LDL-C were inversely associated with this same risk. Similar associations were observed for both lobar and nonlobar ICH. Our results support a potential causal role of LDL-C in risk of primary, nontraumatic ICH.

Acknowledgments

This research has been conducted using the UK Biobank Resource. Financial support: G.J.F. receives grant support from the National Institutes of Health (K76AG059992 and R03NS112859), the American Heart Association (18IDDG34280056), the Yale Pepper Scholar Award (P30AG021342), and the Neurocritical Care Society Research Fellowship. T.M.G. receives grant support from the National Institutes of Health (P30AG021342 and K07AG043587). J.F.M. receives grant support from the Earl & Nyda Swanson Neurosciences Research Fund and the Harley N. and Rebecca N. Hotchkiss Endowed Fund in Neuroscience Research honoring Ken and Marietta. C.L.P. receives grant support from the American Heart Association (19CDA34620004). M.S. receives grant support from the National Institutes of Health (U01NS074425 and U01NS102289). M.H. receives grant support by the British Heart Foundation Intermediate Clinical Research Fellowship (FS/18/23/33512) and the National Institute for Health Research Oxford Biomedical Research Centre and works in a unit that receives grant support from the UK Medical Research Council. S.M. receives grant support from the American Heart Association/American Stroke Association fellowship (18POST34080063). M.V.H. is supported by the British Heart Foundation (FS/18/23/33512) and the National Institute for Health Research Oxford Biomedical Research Centre. D.W. receives grant support from the National Institutes of Health (U24NS107200 and R01NS100417). J.R. receives grant support from the National Institutes of Health (R24NS092983, R01NS093870, T32NS100663, and R01NS100417). K.N.S. receives grant support from the National Institutes of Health (R03NS112859, U24NS107136, U24NS107215, and R01NR018335) and American Heart Association (17CSA33550004). C.D.A. receives grant support from the National Institutes of Health (K23NS086873 and R01NS103924), the American Heart Association (18SFRN34250007), the MGH Center for Genomic Medicine, and research grants from Bayer AG, and had has consulted for ApoPharma, Inc.

Author Contributions

Conception and design of the study: G.J.F., D.G., M.G., L.S., T.M.G., S.K., C.D.L., D.W., J.R., K.N.S., and C.D.A.

Acquisition and analysis of data: G.J.F., E.K., J.N.A., R.B.N., A.L., S.M., J.C., M.S., J.F.M., D.L.B., B.B.W., D.L.T., J.M.J., H.S., J.J.C., I.F.C., A.L., A.S., C.L.P., N.H.P., C.N.M., and A.B.

Drafting the text and preparing the figures: G.J.F., M.H., D.B., Z.C., L.L.S., R.C., R.G.W., and C.D.A.

Potential Conflicts of Interest

The authors declared no conflict of interest.

Data Sharing

All data utilized in this study is publicly available through dbGaP (https://www.ncbi.nlm.nih.gov/gap/) and the UK Biobank (https://www.ukbiobank.ac.uk/).

References

1. An SJ, Kim TJ, Yoon B-W. Epidemiology, risk factors, and clinical features of intracerebral hemorrhage: an update. J Stroke 2017; 19:3–10.
2. Goldstein LB, Amarenco P, Szarek M, et al. Hemorrhagic stroke in the stroke prevention by aggressive reduction in cholesterol levels study. Neurology 2008;70:2364–2370.
3. Liao JK, Laufs U. Pleiotropic effects of statins-NIH public access. Annu Rev Pharmacol Toxicol 2005;45:89-118.
4. McKinney JS, Kostis WJ. Statin therapy and the risk of intracerebral hemorrhage. Stroke 2012;43:2149–2156.
5. Pandit AK, Kumar P, Kumar A, et al. High-dose statin therapy and risk of intracerebral hemorrhage: a meta-analysis. Acta Neurol Scand 2016;134:22–28.
6. Brown MJ. MRC/BHF heart protection study. Lancet 2002;360:1782. https://doi.org/10.1016/S0140-6736(02)11690-4.
7. Woo D, Sauerbeck LR, Kissela BM, et al. Genetic and environmental risk factors for intracerebral hemorrhage. Stroke 2002;33:1190–1196.
8. Rodriguez-Luna D, Rubiera M, Ribo M, et al. Serum low-density lipoprotein cholesterol level predicts hematoma growth and clinical outcome after acute intracerebral hemorrhage. Stroke 2011;42: 2447–2452.
9. Smith GD, Ebrahim S. Mendelian randomization: genetic variants as instruments for strengthening causal inference in observational studies. *Biosocial Surveys*, Washington, DC: National Academies Press; 2008:78-95.
10. Falcone GJ, Biffi A, Devan WJ, et al. Burden of blood pressure-related alleles is associated with larger hematoma volume and worse outcome in intracerebral hemorrhage. Stroke 2013;44:321–326.
11. Lewis CM, Vassos E. Prospects for using risk scores in polygenic medicine. Genome Med 2017;9:96.
12. Sun L, Clarke R, Bennett D, et al. Causal associations of blood lipids with risk of ischemic stroke and intracerebral hemorrhage in Chinese adults. Nat Med 2019;25:569–574. https://doi.org/10.1038/s41591-019-0366-x.
13. Smith JA, Ware EB, Middha P, et al. Current applications of genetic risk scores to cardiovascular outcomes and subclinical phenotypes. Curr Epidemiol Rep 2015;2:180–190.
14. Martini SR, Flaherty ML, Brown WM, et al. Risk factors for intracerebral hemorrhage differ according to hemorrhage location. Neurology 2012;79:2275–2282.
15. Hoffmann TJ, Theusch E, Haldar T, et al. A large electronic-health-record-based genome-wide study of serum lipids. Nat Genet 2018; 50:401–413.
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;50:1514–1523.
17. Consortium GLG, Willer CJ, Schmidt EM, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45: 1274–1283.
18. Genes for Cerebral Hemorrhage on Anticoagulation Collaborative G. Exploiting common genetic variation to make anticoagulation safer. Stroke 2009;40:S64–S66.
19. Woo D, Falcone GJ, Devan WJ, et al. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. Am J Hum Genet 2014;94:511–521.
20. Cox N. UKbiobank shares the promise of big data. Nature 2018;562: 194–195.
21. Falcone GJ, Woo D. Genetics of spontaneous intracerebral hemorrhage. Stroke 2017;48:3420–3424.
22. Marees AT, de Kluiver H, Stringer S, et al. A tutorial on conducting genome-wide association studies: quality control and statistical analysis. Int J Methods Psychiatr Res 2018;27:76–80. https://doi.org/10.1016/j.taml.2017.02.002.
23. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–909.
24. The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature 2010;467: 1061–1073.
25. Biffi A, Sonni A, Anderson CD, et al. Variants at APOE influence risk of deep and lobar intracerebral hemorrhage. Ann Neurol 2010;68: 934–943.
26. Malik R, Chauhan G, Traylor M, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet 2018;50:524–537.
27. 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;81:559–575.
28. Wu C, Dewan A, Hoh J, Wang Z. A comparison of association methods correcting for population stratification in case-control studies. Ann Hum Genet 2011;75:418–427.
29. Howie B, Fuchsberger C, Stephens M, et al. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet 2012;44:955–959.
30. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009;5:e1000529. https://doi.org/10.1371/journal.pgen.1000529.
31. RStudio Team. *RStudio: Integrated Development for R*. Boston, MA: *RStudio, Inc*, 2015 *URL http//www rstudio com*.
32. Anderson CD, Falcone GJ, Phuah C-L, et al. Genetic variants in CETP increase risk of intracerebral hemorrhage. Ann Neurol 2016; 80:730–740.
33. Allman PH, Aban IB, Tiwari HK, Cutter GR. An introduction to Mendelian randomization with applications in neurology. Mult Scler Relat Disord 2018;24:72–78. https://doi.org/10.1016/j.msard.2018.06.017.

34. Tryka KA, Hao L, Sturcke A, et al. NCBI's database of genotypes and phenotypes: DbGaP. Nucleic Acids Res 2014;42:D975–D979. https://doi.org/10.1093/nar/gkt1211.

35. Lappalainen I, Almeida-King J, Kumanduri V, et al. The European genome-phenome archive of human data consented for biomedical research. Nat Genet 2015;47:692–695.

36. Ooneda G, Yoshida Y, Suzuki K, et al. Smooth muscle cells in the development of plasmatic arterionecrosis, arteriosclerosis, and arterial contraction. J Vasc Res 1978;15:148–156.

37. Bang OY, Saver JL, Liebeskind DS, et al. Cholesterol level and symptomatic hemorrhagic transformation after ischemic stroke thrombolysis. Neurology 2007;68:737–742.

38. Contois JH, McConnell JP, Sethi AA, et al. Apolipoprotein B and cardiovascular disease risk: position statement from the AACC lipoproteins and vascular diseases division working group on best practices. Clin Chem 2009;55:407–419.

39. Ryoo JH, Ha EH, Kim SG, et al. Apolipoprotein B is highly associated with the risk of coronary heart disease as estimated by the Framingham risk score in healthy Korean men. J Korean Med Sci 2011;26:631–636.

40. Upadhyay RK. Emerging risk biomarkers in cardiovascular. J Lipids 2015;2015:1–50.