

Genetic Variants in *CETP* Increase Risk of Intracerebral Hemorrhage

Christopher D. Anderson, MD, MMSc,^{1,2,3,4}
 Guido J. Falcone, MD, ScD, MPH,^{1,2,3,4,5} Chia-Ling Phuah, MD,^{1,2,3,4}
 Farid Radmanesh, MD, MPH,^{1,2,3,4} H. Bart Brouwers, MD,^{1,2,3,4}
 Thomas W. K. Battey, BS,^{1,2,3,4} Alessandro Biffi, MD,^{1,2,4,6,7}
 Gina M. Peloso, PhD,^{1,4} Dajiang J. Liu, PhD,⁸ Alison M. Ayres, BA,^{1,2}
 Joshua N. Goldstein, MD, PhD,⁹ Anand Viswanathan, MD, PhD,²
 Steven M. Greenberg, MD, PhD,² Magdy Selim, MD, PhD,¹⁰
 James F. Meschia, MD,¹¹ Devin L. Brown, MD, MS,¹²
 Bradford B. Worrall, MD, MSc,¹³ Scott L. Silliman, MD,¹⁴

View this article online at wileyonlinelibrary.com. DOI: 10.1002/ana.24780

Received May 17, 2016, and in revised form Sep 13, 2016. Accepted for publication Sep 13, 2016.

Members of the International Stroke Genetics Consortium are available as an online supplementary file. Please refer to Supplementary Table 7 for contributors and affiliations.

The copyright line for this article was changed on 21 November 2016 after original online publication.

Address correspondence to Dr Anderson, Center for Human Genetic Research, Massachusetts General Hospital, 185 Cambridge Street; CPZN-6818, Boston, MA 02114. E-mail: cdanderson@mgh.harvard.edu

From the ¹Center for Human Genetic Research, Massachusetts General Hospital (MGH), Boston, MA; ²J. Philip Kistler Stroke Research Center, Department of Neurology, MGH, Boston, MA; ³Division of Neurocritical Care and Emergency Neurology, Department of Neurology, MGH, Boston, MA; ⁴Program in Medical and Population Genetics, Broad Institute, Cambridge, MA; ⁵Departments of Epidemiology and Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA; ⁶Division of Behavioral Neurology, Department of Neurology, MGH, Boston, MA; ⁷Division of Psychiatry, Department of Psychiatry, MGH, Boston, MA; ⁸Department of Public Health Sciences, Institute of Personalized Medicine, Penn State College of Medicine, Hershey, PA; ⁹Department of Emergency Medicine, MGH, Boston, MA; ¹⁰Department of Neurology, Beth Israel Deaconess Medical Center, Boston, MA; ¹¹Department of Neurology, Mayo Clinic, Jacksonville, FL; ¹²Stroke Program, Department of Neurology, University of Michigan Health System, Ann Arbor, MI; ¹³Departments of Neurology and Public Health Sciences, University of Virginia Health System, Charlottesville, VA; ¹⁴Department of Neurology, University of Florida College of Medicine, Jacksonville, FL; ¹⁵Stroke Center, Harborview Medical Center, University of Washington, Seattle, WA; ¹⁶Department of Neurology, University of Cincinnati College of Medicine, Cincinnati, OH; ¹⁷Department of Neurology, Jagiellonian University Medical College, Krakow, Poland; ¹⁸Institute of Molecular Biology and Biochemistry, Medical University Graz, Graz, Austria; ¹⁹Division of Neurology, Department of Clinical Sciences Lund, Lund University, Lund, Sweden; ²⁰Division of Neurology, Department of Neurology and Rehabilitation Medicine, Skåne University Hospital, Lund, Sweden; ²¹Neurovascular Research Unit, Department of Neurology, Municipal Institute of Medical Investigation–Hospital of the Sea, Autonomous University of Barcelona, Barcelona, Spain; ²²Program in Inflammation and Cardiovascular Disorders, Municipal Institute of Medical Investigation–Hospital of the Sea, Autonomous University of Barcelona, Barcelona, Spain; ²³Department of Neurology and Neurosurgery, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, the Netherlands; ²⁴Department of Neurology, Donders Institute for Brain, Cognition, and Behavior, Radboud University Medical Center, Nijmegen, the Netherlands; ²⁵Division of Clinical Brain Sciences, University of Edinburgh, Edinburgh, United Kingdom; ²⁶Institute for Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom; ²⁷Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, United Kingdom; ²⁸Department of Clinical and Experimental Sciences, Neurology Clinic, University of Brescia, Brescia, Italy; ²⁹Department of Neurology, Medical University of Graz, Graz, Austria; ³⁰Division of Neuroradiology, Department of Radiology, Medical University of Graz, Graz, Austria; ³¹Neurovascular Research Laboratory and Neurovascular Unit, Research Institute, Vall d'Hebron Hospital, Autonomous University of Barcelona, Barcelona, Spain; ³²Stroke Pharmacogenomics and Genetics, Terrassa Mutual Teaching and Research Foundation, Terrassa Mutual Hospital, Terrassa, Spain; ³³Department of Neurology, University of Arizona, Tucson, AZ; ³⁴Department of Neurology, Baltimore Veterans Administration Medical Center and University of Maryland School of Medicine, Baltimore, MD; ³⁵National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD; ³⁶Center for Public Health Genomics and Department of Biostatistical Sciences, Wake Forest University, Winston-Salem, NC; ³⁷Center for Statistical Genetics, Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, MI; ³⁸Division of Cardiology, Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI; ³⁹Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI; and ⁴⁰Cardiovascular Disease Prevention Center, MGH, Boston, MA

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

David L. Tirschwell, MD, MSc,¹⁵ Matthew L. Flaherty, MD,¹⁶ Peter Kraft, PhD,⁵
 Jeremiasz M. Jagiella, MD, PhD,¹⁷ Helena Schmidt, MD,¹⁸
 Björn M. Hansen, MD,^{19,20} Jordi Jimenez-Conde, MD, PhD,^{21,22}
 Eva Giralto-Steinhauer, MD,^{21,22} Roberto Elosua, MD, PhD,^{21,22}
 Elisa Cuadrado-Godia, MD,^{21,22} Carolina Soriano, PhD, BSc,^{21,22}
 Koen M. van Nieuwenhuizen, MD,²³ Catharina J. M. Klijn, MD, PhD,^{23,24}
 Kristiina Rannikmae, MD,²⁵ Neshika Samarasekera, PhD, MRCP,²⁵
 Rustam Al-Shahi Salman, MA, PhD, FRCP,²⁵
 Catherine L. Sudlow, BMBCh, MSc, DPhil, FRCPE,^{25,26}
 Ian J. Deary, FBA, FRSE, FMedSci,²⁷ Andrea Morotti, MD,²⁸
 Alessandro Pezzini, MD,²⁸ Joanna Pera, MD,¹⁷ Andrzej Urbanik, MD, PhD,¹⁷
 Alexander Pichler, MD,²⁹ Christian Enzinger, MD,^{29,30} Bo Norrving, MD,^{19,20}
 Joan Montaner, MD, PhD,³¹ Israel Fernandez-Cadenas, PhD,^{31,32}
 Pilar Delgado, MD, PhD,³¹ Jaume Roquer, MD, PhD,^{21,22} Arne Lindgren, MD,^{19,20}
 Agnieszka Slowik, MD, PhD,¹⁷ Reinhold Schmidt, MD,²⁹ Chelsea S. Kidwell, MD,³³
 Steven J. Kittner, MD, MPH,³⁴ Salina P. Waddy, MD,³⁵ Carl D. Langefeld, PhD,³⁶
 Goncalo Abecasis, PhD,³⁷ Cristen J. Willer, PhD,^{38,39} Sekar Kathiresan, MD,^{1,4,40}
 Daniel Woo, MD,¹⁶ and Jonathan Rosand, MD, MSc,^{1,2,3,4}
 on behalf of the Global Lipids Genetics Consortium and International
 Stroke Genetics Consortium

Objective: In observational epidemiologic studies, higher plasma high-density lipoprotein cholesterol (HDL-C) has been associated with increased risk of intracerebral hemorrhage (ICH). DNA sequence variants that decrease cholesteryl ester transfer protein (*CETP*) gene activity increase plasma HDL-C; as such, medicines that inhibit *CETP* and raise HDL-C are in clinical development. Here, we test the hypothesis that *CETP* DNA sequence variants associated with higher HDL-C also increase risk for ICH.

Methods: We performed 2 candidate-gene analyses of *CETP*. First, we tested individual *CETP* variants in a discovery cohort of 1,149 ICH cases and 1,238 controls from 3 studies, followed by replication in 1,625 cases and 1,845 controls from 5 studies. Second, we constructed a genetic risk score comprised of 7 independent variants at the *CETP* locus and tested this score for association with HDL-C as well as ICH risk.

Results: Twelve variants within *CETP* demonstrated nominal association with ICH, with the strongest association at the rs173539 locus (odds ratio [OR] = 1.25, standard error [SE] = 0.06, $p = 6.0 \times 10^{-4}$) with no heterogeneity across studies ($I^2 = 0\%$). This association was replicated in patients of European ancestry ($p = 0.03$). A genetic score of *CETP* variants found to increase HDL-C by ~ 2.85 mg/dl in the Global Lipids Genetics Consortium was strongly associated with ICH risk (OR = 1.86, SE = 0.13, $p = 1.39 \times 10^{-6}$).

Interpretation: Genetic variants in *CETP* associated with increased HDL-C raise the risk of ICH. Given ongoing therapeutic development in *CETP* inhibition and other HDL-raising strategies, further exploration of potential adverse cerebrovascular outcomes may be warranted.

ANN NEUROL 2016;80:730–740

Serum levels of high-density lipoprotein cholesterol (HDL-C) are strongly and inversely associated with coronary artery disease (CAD) risk.¹ Of the many single nucleotide polymorphisms (SNPs) associated with HDL-C levels, those within cholesteryl ester transfer protein (*CETP*) have the strongest effect.^{2–4} Inhibitory variants within *CETP* associated with increased HDL-C correlate with reduced risk of multiple cardiac risk factors, including metabolic syndrome and myocardial infarction.^{5–8}

Inhibitors of the *CETP* gene product, designed to raise HDL-C by limiting *CETP*-mediated exchange of cholesteryl esters and triglycerides between HDL and low-density lipoprotein (LDL)/very low-density lipoprotein particles, are being investigated in ongoing phase III trials as treatments to reduce CAD risk.^{9,10}

In contrast, substantial data suggest that elevations in HDL-C may increase risk of spontaneous intracerebral hemorrhage (ICH).^{11,12} Furthermore, clinical trial data

TABLE 1. Discovery Populations

Variable	GOCHA		ISGC ICH Study		GERFHS	
	Cases	Controls	Cases	Controls	Cases	Controls
No.	371	389	404	530	374	319
Age, mean (SD)	74 (10)	72 (8)	70 (13)	66 (16)	67 (15)	67 (14)
Female, No. [%]	172 [46]	195 [50]	189 [47]	266 [50]	194 [52]	172 [54]
HTN, No. [%]	274 [75]	227 [58]	278 [69]	247 [47]	241 [64]	166 [52]
T2D, No. [%]	68 [18]	35 [9]	89 [22]	68 [13]	72 [19]	42 [13]
HL, No. [%]	144 [39]	195 [50]	87 [22]	48 [9]	131 [35]	133 [42]
Smoking, No. [%]	56 [15]	15 [4]	58 [14]	74 [14]	79 [21]	46 [14]
Genotyping platform	Illumina 610	Illumina 610	Illumina 610	Illumina 610	Affymetrix 6.0	Affymetrix 6.0
Lobar, ^a No. [%]	205 [55]	–	135 [33]	–	156 [42]	–

Discovery totals: 2,387 individuals (1,149 cases, 1,238 controls), 43% lobar ICH.
^aLobar ICH location.
 GERFHS = Genetic and Environmental Risk Factors for Hemorrhagic Stroke study; GOCHA = Genes and Outcomes of Cerebral Hemorrhage on Anticoagulation study; HL = hyperlipidemia; HTN = hypertension; ICH = intracerebral hemorrhage; ISGC = International Stroke Genetics Consortium; SD = standard deviation; T2D = type 2 diabetes mellitus.

suggest an increased risk of ICH on statins despite a lack of significant differences in lipid levels.^{13,14} Because of small sample sizes and confounding by environmental or medical exposures, prior studies have not been able to resolve this potentially paradoxical role of elevated HDL-C in ICH. Although ICH comprises only 15 to 20% of all strokes, it accounts for 50% of all stroke-related mortality and 30% of total costs.^{15,16} Blood pressure control remains the only available preventive strategy.¹⁷ As HDL-C evolves as a cardiovascular treatment target and clinical trial data on therapeutic modifiers accrue, an improved mechanistic understanding of the pathways involved in hemorrhagic cerebrovascular disease could lead to alternative treatments and prevention strategies for ICH.

It is not known whether CETP inhibitors, which endeavor to produce a biological effect similar to known genetic variants in *CETP*, increase ICH risk. The objective of this study was to use genome-wide genotypes from individuals with and without ICH from the International Stroke Genetics Consortium (ISGC) to test genetic variants within *CETP* for association with ICH risk, under the hypothesis that the HDL-raising effects of inhibitory variants within *CETP* will result in increased ICH. *CETP* genetic variants that impact HDL-C are unconfounded by other exposures, remain constant throughout life, and may be more reflective of long-term levels than periodic lipid measurements.¹⁸ Thus, examination of *CETP* genetic variation constitutes a valuable

causal inference tool to help strengthen or disclaim prior observations of association between elevated HDL-C and ICH, and could provide additional clues about potential adverse effects of pharmacologic CETP inhibition.

Materials and Methods

Study Design

We performed a 2-stage (discovery and replication) case-control candidate-gene association study using both genome-wide data and direct genotyping. The discovery phase utilized data from 3 genome-wide association studies (GWASs) of ICH, sampling patients of European ancestry (Table 1).¹⁹ Replication involved direct genotyping of variants of interest from individuals recruited through 5 case-control studies of ICH, with no overlap between individuals from the discovery phase (Table 2). Detailed description of discovery and replication case and control recruitment architectures can be found in Supplementary Table 1.

All studies had approval from the local institutional review board or ethics committee at each participating institution. Informed consent was obtained from all patients or their legally authorized representatives, or was waived via protocol-specific allowance.

Cases

ICH was defined as a new and acute neurological deficit with compatible brain imaging. Enrolled patients were adult consenting primary acute ICH cases that presented to participating institutions with confirmation of primary ICH through computed tomography or magnetic resonance imaging. Exclusion criteria included trauma, brain tumor, hemorrhagic transformation of a

TABLE 2. Replication Populations

Variable	MGH		ERICH		University of Brescia		UMC Utrecht		University of Edinburgh	
	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
No.	240	458	920	826	198	185	157	160	110	216
Age, No. (SD)	76 (10)	69 (11)	69 (14)	68 (13)	69 (13)	63 (14)	62 (13)	56 (11)	75 (9)	76 (10)
Female, No. [%]	96 [40]	206 [45]	397 [43]	371 [45]	81 [41]	85 [46]	66 [42]	67 [42]	59 [54]	118 [54]
Lobar, No. [%]	120 [48]	–	380 [41]	–	82 [41]	–	60 [38]	–	61 [55]	–
Genotyping platform	iPLEX	iPLEX	Taqman	Taqman	iPLEX	iPLEX	iPLEX	iPLEX	iPLEX	iPLEX

Replication totals: 3,470 individuals (1,625 cases, 1,845 controls), 42% lobar ICH. Discovery + replication totals: 5,625 individuals (2,595 cases, 3,030 controls), 45% lobar ICH.
 ERICH = Ethnic/Racial Variations of Intracerebral Hemorrhage; ICH = intracerebral hemorrhage; iPLEX = Sequenom MassARRAY iPLEX Platform; MGH = Massachusetts General Hospital; SD = standard deviation; TaqMan = Applied Biosystems Taqman Genotyping Assay; UMC = University Medical Center.

cerebral infarction, vascular malformation, or any other cause of secondary ICH in all participating studies.

CASE POPULATIONS. ICH cases were recruited across multiple centers participating in the ISGC from sites in the USA and Europe. For the purposes of reducing confounding by population stratification, only individuals of self-reported European (Caucasian) ancestry were included in the analysis. Likewise, several studies (Genetics of Cerebral Hemorrhage with Anticoagulation, Edinburgh Stroke Study, LINCHPIN) recruited ICH patients with ICH in the presence of anticoagulation (typically warfarin) exposure. These individuals were excluded from analyses due to the etiopathological distinctness of warfarin-related primary ICH from other forms. Discovery case populations were enrolled according to methods previously described.¹⁹ Replication cases were recruited from ISGC participating centers using similar criteria as discovery cases (Supplementary Table 2). Briefly, the University Medical Center (UMC) Utrecht ICH study included additional screening for secondary ICH cases in follow-up. The Edinburgh Stroke Study recruited subjects aged >55 years only, and specifically excluded individuals with antecedent illicit drug use or presentation >1 week from onset of symptoms. The LINCHPIN study identified ICH cases aged >16 years with acute or chronic ICH from a prospective cohort of individuals living in the Lothian region of Scotland, United Kingdom.

NEUROIMAGING. Stroke neurologists and neuroradiologists at each participating site performed the neuroimaging assessment. Following known differences in underlying biology, ICH was classified as lobar or nonlobar according to location.²⁰ ICH originating in the corticosubcortical junction (with or without involvement of subcortical white matter) was defined as lobar, whereas ICH selectively involving the thalamus, internal capsule, basal ganglia, brainstem, or cerebellum was defined as nonlobar.

Controls

Controls were ICH-free individuals >18 years of age and were enrolled from the same populations that gave rise to the cases. Controls were confirmed to have no history of previous ICH by interview and/or medical record review. Control population age restrictions were identical to case population age restrictions for all included studies.

CONTROL POPULATIONS. ICH-free controls were recruited from the same populations that gave rise to the ICH cases, through inpatient recruitment, ambulatory centers in the local communities, blood donation centers serving the same population, and in the case of the Lothian Birth Cohort, a population cohort study (Supplementary Table 3). The Genetic and Environmental Risk Factors for Hemorrhagic Stroke (GERFHS) and Ethnic/Racial Variations of Intracerebral Hemorrhage (ERICH) studies^{19,21} used random digit dialing, the Lothian Birth Cohort individuals were matched to case samples by local investigators,²² and UMC Utrecht identified controls from the local blood donor population. The remainder of the studies used random selection from ambulatory clinics or geographically matched populations where cases were being recruited.

Exposure: Common Genetic Variants within CETP

In the discovery phase, we ascertained variants within *CETP* by means of genome-wide genotyping followed by imputation using methods and quality control procedures previously described.¹⁹ Briefly, DNA was isolated from fresh or frozen peripheral whole blood collected from cases and controls at each participating institution at the time of consent, quantified with a quantification kit (Qiagen, Valencia, CA), and normalized to a concentration of 30ng/μl. Cases and controls were plated together and genotyped on Illumina (San Diego, CA) 610 or Affymetrix (Santa Clara, CA) 6.0 platforms. Standard quality controls for genome-wide data were applied, and the

resulting set of individuals and SNPs were carried forward to imputation, which was completed using IMPUTE2 with 1000 Genomes-based reference panels (March 2012 version).²³ Post-imputation exclusion filters were minor allele frequency (MAF) < 0.01 and information score < 0.5. SNPs were extracted from the *CETP* gene region according to the human genome reference GRCh38.p2 annotated location (<http://www.ncbi.nlm.nih.gov>), \pm 50 kilobases.

Independent Replication

CETP variants exceeding Bonferroni-corrected significance and without significant heterogeneity ($I^2 < 40\%$) for association with ICH in the discovery phase were selected for replication.²⁴ Replication SNPs were chosen based on proxy status with index SNPs. Because replication of *CETP* variants was carried out as part of an ongoing GWAS of ICH, a constraint for the selection of replication SNPs was predicted genotyping success using iPLEX (Sequenom, San Diego, CA) and Taqman (Applied Biosystems, Foster City, CA) methodologies, which were employed at the Massachusetts General Hospital and University of Miami genotyping centers, respectively (see Table 2). Ancestry-informative markers were also genotyped to facilitate adjustment for population admixture.

Data Analysis

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

POPULATION STRUCTURE. Principal component analysis was implemented in both discovery and replication to account for population structure, using genome-wide data in discovery and prespecified ancestry-informative markers in replication.^{25,26} Caucasian population outliers were identified and removed by visual inspection of plots generated with principal components 1 and 2, and these principal components were included as covariates in regression models fitted for association testing. In the GERFHS and ERICH samples, further refinement of population structure was achieved using the ADMIXTURE software tool to remove outliers.²⁷

ASSOCIATION TESTING. Prior to discovery association testing, SNPs within *CETP* were clumped into loci sharing linkage disequilibrium (LD) $r^2 > 0.5$ using PLINK to allow discrimination of semi-independent loci across the gene. Association testing for SNPs within the *CETP* locus and ICH risk was completed separately for all ICH, as well as for lobar and nonlobar hemorrhages. Logistic regression models were fitted assuming independent additive genetic effects for dosage of the minor allele (1 degree of freedom additive trend test), and adjusted for age, gender, and principal components 1 and 2. A similar analytic approach was undertaken for analysis of replication data, using additive allele genotype data rather than dosage.

META-ANALYSIS. Fixed effects, inverse variance weighted meta-analysis was used to pool effect estimates across studies, assessing heterogeneity by computing Cochran's *Q* (with corresponding *p*) and I^2 (percentage of effect size attributable to

heterogeneity). Identical meta-analysis procedures were used for pooling of effects across studies in discovery and replication, and across all studies.²⁸

Genetic Risk Score Analysis

Variants within the *CETP* locus with established association with HDL-C levels in the most recent Global Lipids Genetics Consortium (GLGC) analysis (unpublished data) were extracted from the discovery data set and tested for association with ICH using an additive multi-SNP genetic risk score approach using the GTX package (<http://CRAN.R-project.org/package=gtx>) in R (version 3.0). Ten variants surpassing exome array-wide significance ($p < 2.1 \times 10^{-7}$) and demonstrating independence using a sequential forward selection model in the GLGC data set were identified, of which 7 were available in our ICH discovery data set.²⁹ These 7 variants, on average, were associated with a 0.19-SD increase in HDL-C (~ 2.85 mg/dl) in the GLGC population ($p < 1 \times 10^{-200}$). This corresponds to a proportion of variance explained of 0.032. ICH risk was predicted from summary statistics, weighted according to the established HDL-C effect, and oriented to the HDL-C increasing allele.

Statistical Testing and Software

We used a conservative Bonferroni-corrected threshold for statistical significance of $p < 0.004$, adjusted for the number of semi-independent loci within *CETP* with $r^2 < 0.5$ (12 tests in this analysis). Quality control procedures, genetic association testing for single variants, and score calculations were performed in SNPTest and PLINK v1.07.^{26,30} Imputation was completed using IMPUTE2.²³ All other statistical analyses were performed in SAS 9.2 (SAS Institute, Cary, NC).

Results

Following relevant exclusions during quality control and principal component analysis, 1,149 ICH cases and 1,238 controls from 3 case-control studies of ICH were included in the discovery phase, 43% of which were of the lobar ICH subtype (see Table 1).

CETP Genetic Variants

After imputation using 1000 Genomes reference panels and application of genome-wide quality control filters, a total of 390 common variants of MAF > 0.01 were extracted from the *CETP* gene and 50kb flanking regions (Supplementary Table 4).³¹ These 390 variants were present either via array-based ascertainment or imputation in all 3 of the discovery data sets, and were used for association testing.

Single-SNP Association Testing

After testing all 390 SNPs within *CETP* clumped into regions sharing $r^2 > 0.5$, 12 loci demonstrating nominal association with ICH ($p < 0.05$) were identified (Supplementary Table 5). Three of these loci surpassed Bonferroni correction (Table 3) with residual $r^2 = 0.25$ to 0.39

TABLE 3. Discovery *CETP* Loci Demonstrating Bonferroni-Significant Association with ICH

Lead SNP	CHR	Tested Allele	MAF	Effect Direction	OR	SE	Discovery p	r^2
rs173539	16	T	0.31	+++	1.25	0.06	6.00E-4	0
rs820299	16	G	0.38	---	0.81	0.06	7.50E-4	48
rs158478	16	A	0.48	+++	1.21	0.06	1.48E-3	56

+ = variant increases ICH risk; - = variant decreases ICH risk; CHR = chromosome; ICH = intracerebral hemorrhage; MAF = minor allele frequency; OR = odds ratio; SE = standard error; SNP = single nucleotide polymorphism.

between them. Among these, only rs173539 (odds ratio [OR] = 1.25, standard error [SE] = 0.06, $p = 6.00 \times 10^{-4}$) met prespecified criteria for replication due to its homogeneity across discovery data sets ($r^2 = 0\%$). Of note, rs173539 was in high LD with rs3764261 ($r^2 = 0.98$), the strongest associated SNP with HDL-C in published GWASs of lipid levels (Fig 1).³² Comparison of effects of the rs173539 locus on risk of lobar versus non-lobar hemorrhage revealed no significant differences by ICH subtype (Supplementary Table 6).

Replication and Meta-Analysis of the rs173539 Locus

A total of 1,625 ICH cases and 1,845 controls of Caucasian ancestry were available for replication. Following application of predictive algorithms for SNP genotype ascertainment success using both genotyping methodologies employed, 4 SNPs in LD with rs173539 locus were selected for replication genotyping according to the constraints described (Tables 4 and 5). Both rs173539 and rs3764261 were predicted to fail in one or both replication pools. All 4 selected SNPs were successfully genotyped in all replication data sets. All replication results showed minimal heterogeneity and consistent directions

of effect, and 2 variants replicated at $p < 0.05$. In meta-analysis, all 4 SNPs within the rs173539 locus chosen for replication were strengthened by addition of the replication SNP data, with minimal heterogeneity in the final total sample size of 2,595 ICH cases and 3,030 controls (see Table 5).

Genetic Risk Score Analysis

An additive multi-SNP genetic risk score was constructed using independent HDL-association data.²⁹ Ten variants were selected, of which 7 were present in the ICH discovery data set (Table 6). Three variants were unavailable in the ICH data set due to differences in genotyping platforms (exome array vs GWAS array) between the two studies. The genetic risk score of these 7 variants demonstrated association with ICH (OR = 1.86, SE = 0.13, $p = 1.39 \times 10^{-6}$).

Discussion

Our results demonstrate an association between *CETP* gene variants in the rs173539 locus and risk of ICH, opposite in direction from their effect on risk of CAD and metabolic syndrome.^{5,7,8} Furthermore, an aggregated score of variants within *CETP* that raise HDL-C is

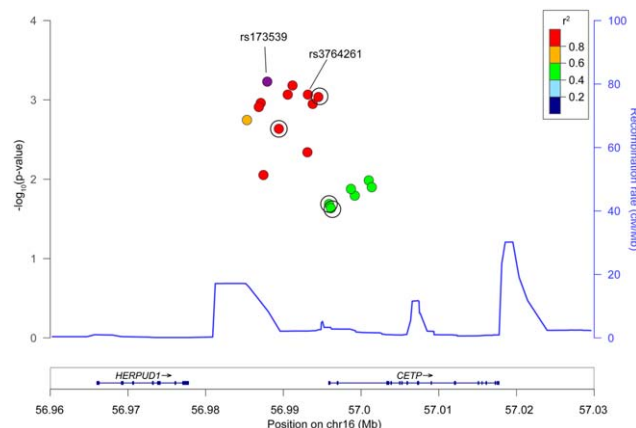


FIGURE 1: Regional association plot of rs173539 and single nucleotide polymorphisms (SNPs) exhibiting $r^2 > 0.5$ in association with intracerebral hemorrhage. SNPs available for replication are circled. Mean recombination rate across the locus is represented by the continuous line. The rs3764261 variant identified was the leading SNP in prior genome-wide association studies of high-density lipoprotein cholesterol. chr = chromosome; cM = centimorgans; Mb = megabase. [Color figure can be viewed in the online issue, which is available at www.annalsofneurology.org.]

TABLE 4. Discovery SNP rs173539 and Local Proxies in Association with ICH Risk

SNP	CHR	Tested Allele	MAF	Effect Direction	OR	SE	Discovery <i>p</i>	<i>r</i> ²
rs173539	16	T	0.31	+++	1.25	0.06	6.00 × 10 ⁻⁴	0
-rs247617, <i>r</i> ² = 0.99	16	A	0.31	+++	1.24	0.06	8.74 × 10 ⁻⁴	0
-rs17231506, <i>r</i> ² = 0.99	16	T	0.31	+++	1.23	0.06	9.13 × 10 ⁻⁴	0
-rs711752, <i>r</i> ² = 0.62	16	A	0.42	++-	1.15	0.06	2.08 × 10 ⁻²	14
-rs708272, <i>r</i> ² = 0.61	16	A	0.42	++-	1.15	0.06	2.23 × 10 ⁻²	18

Association results for rs173539 in association with ICH risk, as well as 4 additional SNPs in LD with rs173539 chosen for replication. - = variant decreases ICH risk; + = variant increases ICH risk; CHR = chromosome; ICH = intracerebral hemorrhage; LD = linkage disequilibrium; MAF = minor allele frequency; OR = odds ratio; *r*² = degree of LD with rs173539; SE = standard error; SNP = single nucleotide polymorphism.

strongly associated with increased ICH risk. These results suggest that there may be substantial differences in the roles of lipids in the progression of cerebrovascular and cardiometabolic diseases. Novel therapies targeting CETP along with other approaches to increase HDL-C are currently under active investigation in an effort to reduce the risk of CAD.³³ Because the cerebral small vessel diseases that lead to ICH are common in the aging population and frequently coincide with risk factors for cardiometabolic disease,^{34,35} our observations supporting opposing effects of HDL-C on ICH and CAD underscore the need for a better understanding of which patients could be at increased risk of ICH on therapies aimed at increasing HDL-C.

Our findings support prior studies linking elevated HDL-C with increased risk of ICH. Unlike prior studies, however, our genetic approach limits confounding by dietary, environmental, or medication exposures. A recent meta-analysis of epidemiological studies examining associations between cholesterol levels and ICH found a

dose-response relationship between HDL-C and ICH risk, with each 1mmol/l increase in HDL-C associated with a 17% increase in ICH risk.¹¹ This result was nullified when studies of subarachnoid hemorrhage patients were included, but strengthened by restriction to studies from the United States, highlighting the potential confounds of case misspecification and unmeasured environmental exposures in testing associations of this nature.

HDL-C appears to have a complex and context-dependent role in cerebrovascular disease. In contrast to ICH, elevated HDL-C is associated with reduced risk of ischemic stroke, particularly strokes caused by large artery atherosclerotic disease, consistent with the observed associations of HDL-C in CAD.³⁶ However, Mendelian randomization (MR) studies of genetic variants predisposing to elevated HDL-C have not demonstrated association with either ischemic stroke or CAD, suggesting the observed relationships may not be causal.^{37,38} Unfortunately, the limited sample size of genetics efforts in ICH coupled with acute changes in lipid values around the

TABLE 5. Replication Results for SNPs in Linkage Disequilibrium with rs173539 and Meta-Analysis of All Samples

SNP	Replication					Discovery/Replication Meta-Analysis				
	Effect	OR	SE	<i>p</i>	<i>r</i> ²	Effect	OR	SE	<i>p</i>	<i>r</i> ²
rs247617	+++++	1.08	0.05	0.18	2	++++/+++++	1.13	0.04	1.0 × 10 ⁻³	0
rs17231506	+++++	1.08	0.05	0.17	1	++++/+++++	1.13	0.04	1.0 × 10 ⁻³	0
rs711752	++++-	1.12	0.05	0.03	7	++-/++++-	1.13	0.04	1.0 × 10 ⁻³	0
rs708272	++++-	1.14	0.05	0.01	4	++-/++++-	1.14	0.04	5.0 × 10 ⁻⁴	0

+ = variant increases ICH risk; - = variant decreases ICH risk; ICH = intracerebral hemorrhage; OR = odds ratio; SE = standard error; SNP = single nucleotide polymorphism.

TABLE 6. ICH Association Results for Variants of Known HDL-C Effect Used to Compute Genetic Risk Score

SNP	Ref Allele	MAF	ICH OR	ICH Beta	ICH SE	ICH <i>p</i>	HDL Effect Allele	HDL Beta	HDL SE	Type
rs173539	T	0.31	1.25	0.222	0.065	0.0006	T	0.230	0.0028	Intergenic
rs3764261	A	0.31	1.23	0.210	0.063	0.0009	A	0.239	0.0028	Intergenic
rs247616	T	0.30	1.22	0.196	0.064	0.0023	T	0.242	0.0028	Intergenic
rs9989419	A	0.40	0.92	-0.079	0.059	0.1808	G	0.131	0.0026	Intergenic
rs5880	C	0.04	1.22	0.202	0.151	0.1812	G	0.258	0.0067	Nonsyn.
rs5882	G	0.32	1.06	0.057	0.065	0.3803	G	0.092	0.0028	Nonsyn.
rs7499892	T	0.19	1.02	0.022	0.076	0.7758	C	0.230	0.0033	Intronic

HDL-C = high-density lipoprotein cholesterol; ICH = intracerebral hemorrhage; MAF = minor allele frequency; Nonsyn. = nonsynonymous; OR = odds ratio; Ref = reference; SE = standard error; SNP = single nucleotide polymorphism.

time of ICH currently preclude the use of this MR approach in our analyses.³⁹

No study, including the present, has yet established a direct causal relationship between HDL-C and ICH risk. Although associations between *CETP* genetic variants and ICH are almost certainly unidirectional due to the immutability of the genetic code, they still could impact an unseen risk factor that lies outside of the known HDL-C level determining effects of the gene. Even if causality can be ultimately established, the mechanism by which a *CETP*-mediated increase in HDL-C may worsen ICH risk remains unclear. Inhibition of *CETP* results in changes to HDL particle size and cholesterol efflux capacity in addition to the observed changes in HDL-C serum levels, and it may be through these accompanying changes in HDL function that ICH risk is conferred.⁴⁰ Furthermore, accumulating evidence suggests that HDL effects on endothelium are dynamic and modifiable, and can even become proinflammatory with the incorporation of serum amyloid A1, complement C3, and ceramides, resulting in altered immune regulation and reduced antioxidant effects.^{41,42} It is therefore possible that elevated HDL-C provides a platform to further the vascular inflammatory processes that play a substantial role in the cerebral small vessel disease underlying ICH.⁴³ Further studies will be needed to dissect the pathways intersecting with HDL-C to clarify the foundational biology of its role in ICH.

Therapeutic development of small molecule and biologic compounds designed to raise HDL-C continue.⁴⁴ Although the first wave of phase III trials of *CETP* inhibitors were plagued by off-target effects and futility,⁴⁵ the REVEAL trial of anacetrapib was recently continued

after unblinded interim review. Other HDL-raising strategies, including apolipoprotein-A1 (ApoA1)-rich reconstituted HDL particle infusions and ApoA1-mimetic peptides, continue to be evaluated in preclinical and early phase trials.⁴⁴ Given this pipeline of HDL-based therapeutic development, it is imperative that potential adverse clinical effects of such strategies be clarified. Early experiences with US Food and Drug Administration-approved PCSK-9 inhibitors have led to predictions of widespread adoption of this new class of drugs, and it is reasonable to expect that HDL-C targeted treatments would be no different, resulting in a potentially large population of aging individuals with pharmacologically induced high HDL-C levels of uncertain long-term cerebrovascular risk.⁴⁶ The proportion of variance in HDL-C levels explained by our genetic risk score was 0.032. This is roughly commensurate with observed effects of statins, which in clinical trials raised HDL by 0.04 to 0.10mg/dl.⁴⁷ With emerging HDL-C modifying strategies likely to exert more profound effects, the impact on ICH risk, if confirmed and verified to be causal, could be more substantial than indicated by our *CETP* genetic risk score.

As noted above, our study cannot determine whether the observed association between *CETP* and ICH risk is through HDL-C alone. Although they exhibit their largest effect on HDL-C levels, *CETP* variants are also associated with LDL, triglycerides, and total cholesterol levels.³ Although we cannot perform formal MR, the association between our HDL-C increasing genetic risk score at *CETP* and risk of ICH provides support for an HDL-specific effect. Even with this suggestion of HDL-C specificity, the composition of HDL particles can vary

with respect to ratios of esterified to unesterified cholesterol as well as apolipoprotein content. Genetic variation that determines circulating HDL-C does not necessarily capture these secondary characteristics, which could have a substantial impact on biological effects.

An additional limitation of our study is the aggregation of case and control data across multiple sites, which could result in biases between cases and controls. We have attempted to control for study demographics and population structure in our regression analyses, and performed independent replication, but unmeasured confounding could still have impacted the observed associations. Related to this point, all analyses presented were in individuals of European ancestry due to small study populations, and therefore low statistical power, in individuals of other racial and ethnic backgrounds. As a result, our findings cannot be extended to minority populations at this time.

Although our study utilized genome-wide data for discovery and genetic risk score analyses, our approach was fundamentally a candidate gene study of *CETP*. Using GWAS data allowed for control of population stratification, which can be a major confounder in traditional candidate gene designs employing only direct genotyping. However, it was still based on an a priori hypothesis about *CETP* association with ICH. Therefore, the false discovery rate for association between variants at *CETP* and ICH risk, although stringently controlled using Bonferroni correction at the *CETP* locus, may still be elevated in comparison with a standard GWAS. Due to the hypothesis-driven nature of our study, we by definition cannot provide novel results about lipid-related genetic loci that lie outside of the tested gene region.

Finally, the *CETP* gene contains several independent loci which have been associated with lipid levels and clinical endpoints.^{3,5,7,32} This resulted in a more complex replication phase than would have been needed if the genetic architecture of the locus were centered about a single region of association. Coupled with the limitations of variant selection in our replication phase, we cannot distinguish a culprit variant to the exclusion of others. Although all variants chosen for replication demonstrated refined effect size estimates and greater statistical significance in meta-analysis with discovery data, replication was strongest for variants in slightly lower LD than the lead variant from discovery, and with slightly higher between-study heterogeneity. Whether this observation represents true heterogeneity of effect at the replicated variants will depend on future validation and extension studies.

We have demonstrated an association between genetic variants in *CETP* and risk of ICH, and have shown that *CETP*'s HDL-C raising effects could play a role in the pathogenesis of ICH. Further work will be

needed to identify how the biological pathways impacted by HDL-C may impart increased risk of hemorrhage. These pathways may yield crucial novel targets for prevention of ICH and the cerebral small vessel diseases that lead to vessel rupture.

Acknowledgment

This work was supported by the NIH National Institute of Neurological Disorders and Stroke (NINDS; K23NS086873, R01NS059727, P50NS061343, R01NS36695, U01NS069763, R01NS30678) and the NIH National Institute on Aging (R01AG26484). Project support for the GLGC through C.J.W. and S.K. was provided by the NIH National Heart, Lung, and Blood Institute (R01HL127564). Lund Stroke Register has been supported by the Swedish Heart and Lung Foundation, Skåne University Hospital, Region Skåne, the Freemasons Lodge of Instruction EOS in Lund, King Gustaf V and Queen Victoria's Foundation, Lund University, the Swedish Stroke Association, and Spain's Ministry of Health (Ministerio de Sanidad y Consumo, Instituto de Salud Carlos III FEDER, RD12/0042/0020).

This report does not represent the official view of the NINDS, the NIH, or any part of the US Federal Government. No official support or endorsement of this article by the NINDS or NIH is intended or should be inferred. No funding entities had involvement in study design; data collection, analysis, and interpretation; writing of the manuscript; or the decision to submit for publication.

We thank M. Hernan for valuable counsel on research methods.

This work will be presented at the International Stroke Conference in Feb 2017 in partial form, if that is necessary for inclusion (accepted for presentation prior to acceptance for publication, disclosed as under review).

Author Contributions

C.D.A. and G.J.F. contributed equally to the presented work. Conception and design of study: C.D.A., G.J.F., C.-L.P., F.R., A.B., G.M.P., S.K., D.W., J. Ros. Acquisition and analysis of data: all authors. Drafting manuscript and figure: C.D.A., G.J.F., C.-L.P., F.R., A.B., G.M.P., A.M.A., S.K., J.Ros.

Potential Conflicts of Interest

Nothing to report.

References

1. Rader DJ, Hovingh GK. HDL and cardiovascular disease. *Lancet* 2014;384:618–625.
2. Asselbergs FW, Guo Y, van Iperen EP, et al. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. *Am J Hum Genet* 2012;91:823–838.

3. Global Lipids Genetics Consortium, Willer CJ, Schmidt EM, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013;45:1274–1283.
4. Tada H, Won HH, Melander O, et al. Multiple associated variants increase the heritability explained for plasma lipids and coronary artery disease. *Circ Cardiovasc Genet* 2014;7:583–587.
5. Kraja AT, Vaidya D, Pankow JS, et al. A bivariate genome-wide approach to metabolic syndrome: STAMPEED consortium. *Diabetes* 2011;60:1329–1339.
6. Chambers JC, Elliott P, Zabaneh D, et al. Common genetic variation near MC4R is associated with waist circumference and insulin resistance. *Nat Genet* 2008;40:716–718.
7. Meiner V, Friedlander Y, Milo H, et al. Cholesteryl ester transfer protein (CETP) genetic variation and early onset of non-fatal myocardial infarction. *Ann Hum Genet* 2008;72(pt 6):732–741.
8. Thompson A, Di Angelantonio E, Sarwar N, et al. Association of cholesteryl ester transfer protein genotypes with CETP mass and activity, lipid levels, and coronary risk. *JAMA* 2008;299:2777–2788.
9. Barter PJ, Rye KA. Cholesteryl ester transfer protein inhibition as a strategy to reduce cardiovascular risk. *J Lipid Res* 2012;53:1755–1766.
10. Kastelein JJ, Besseling J, Shah S, et al. Anacetrapib as lipid-modifying therapy in patients with heterozygous familial hypercholesterolaemia (REALIZE): a randomised, double-blind, placebo-controlled, phase 3 study. *Lancet* 2015;385:2153–2161.
11. Wang X, Dong Y, Qi X, et al. Cholesterol levels and risk of hemorrhagic stroke: a systematic review and meta-analysis. *Stroke* 2013;44:1833–1839.
12. Raffeld MR, Biffi A, Battey TW, et al. APOE epsilon4 and lipid levels affect risk of recurrent nonlobar intracerebral hemorrhage. *Neurology* 2015;85:349–356.
13. Goldstein LB, Amarenco P, Szarek M, et al. Hemorrhagic stroke in the Stroke Prevention by Aggressive Reduction in Cholesterol Levels study. *Neurology* 2008;70(24 pt 2):2364–2370.
14. Collins R, Armitage J, Parish S, et al. Effects of cholesterol-lowering with simvastatin on stroke and other major vascular events in 20536 people with cerebrovascular disease or other high-risk conditions. *Lancet* 2004;363:757–767.
15. Taylor TN, Davis PH, Torner JC, et al. Lifetime cost of stroke in the United States. *Stroke* 1996;27:1459–1466.
16. Ikram MA, Wieberdink RG, Koudstaal PJ. International epidemiology of intracerebral hemorrhage. *Curr Atheroscler Rep* 2012;14:300–306.
17. Passero S, Burgalassi L, D'Andrea P, et al. Recurrence of bleeding in patients with primary intracerebral hemorrhage. *Stroke* 1995;26:1189–1192.
18. Ference BA, Yoo W, Alesh I, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. *J Am Coll Cardiol* 2012;60:2631–2639.
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. Martini SR, Flaherty ML, Brown WM, et al. Risk factors for intracerebral hemorrhage differ according to hemorrhage location. *Neurology* 2012;79:2275–2282.
21. Woo D, Rosand J, Kidwell C, et al. The Ethnic/Racial Variations of Intracerebral Hemorrhage (ERICA) study protocol. *Stroke* 2013;44:e120–e125.
22. Deary IJ, Gow AJ, Pattie A, et al. Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. *Int J Epidemiol* 2012;41:1576–1584.
23. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* 2011;1:457–470.
24. Higgins JPT, Green S, Cochrane Collaboration. *Cochrane handbook for systematic reviews of interventions*. Chichester, UK: Wiley-Blackwell, 2008.
25. 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.
26. 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.
27. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 2009;19:1655–1664.
28. Evangelou E, Ioannidis JP. Meta-analysis methods for genome-wide association studies and beyond. *Nat Rev Genet* 2013;14:379–389.
29. Saey Y, Inza I, Larranaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics* 2007;23:2507–2517.
30. Pei YF, Zhang L, Li J, et al. Analyses and comparison of imputation-based association methods. *PLoS One* 2010;5:e10827.
31. Hall JB, Cooke Bailey JN, Hoffman JD, et al. Estimating cumulative pathway effects on risk for age-related macular degeneration using mixed linear models. *BMC Bioinformatics* 2015;16:329.
32. Kathiresan S, Willer CJ, Peloso GM, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* 2009;41:56–65.
33. Brinton EA, Kher U, Shah S, et al. Effects of anacetrapib on plasma lipids in specific patient subgroups in the DEFINE (Determining the Efficacy and Tolerability of CETP INhibition with AnacEtrapib) trial. *J Clin Lipidol* 2015;9:65–71.
34. Ding J, Sigurdsson S, Garcia M, et al. Risk factors associated with incident cerebral microbleeds according to location in older people: the Age, Gene/Environment Susceptibility (AGES)-Reykjavik Study. *JAMA Neurol* 2015;72:682–688.
35. Palacio S, McClure LA, Benavente OR, et al. Lacunar strokes in patients with diabetes mellitus: risk factors, infarct location, and prognosis: the secondary prevention of small subcortical strokes study. *Stroke* 2014;45:2689–2694.
36. Amarenco P, Goldstein LB, Szarek M, et al. Effects of intense low-density lipoprotein cholesterol reduction in patients with stroke or transient ischemic attack: the Stroke Prevention by Aggressive Reduction in Cholesterol Levels (SPARCL) trial. *Stroke* 2007;38:3198–3204.
37. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a Mendelian randomization study. *Lancet* 2012;380:572–580.
38. Pikula A, Beiser AS, Wang J, et al. Lipid and lipoprotein measurements and the risk of ischemic vascular events: Framingham Study. *Neurology* 2015;84:472–479.
39. Phuah CL, Raffeld MR, Ayres AM, et al. Subacute decline in serum lipids precedes the occurrence of primary intracerebral hemorrhage. *Neurology* 2016;86:2034–2041.
40. Mohammadpour AH, Akhlaghi F. Future of cholesteryl ester transfer protein (CETP) inhibitors: a pharmacological perspective. *Clin Pharmacokinet* 2013;52:615–626.
41. Weichhart T, Kopecky C, Kubicek M, et al. Serum amyloid A in uremic HDL promotes inflammation. *J Am Soc Nephrol* 2012;23:934–947.
42. Papageorgiou N, Zacharia E, Androulakis E, et al. HDL as a prognostic biomarker for coronary atherosclerosis: the role of inflammation. *Expert Opin Ther Targets* 2016;20:907–921.

43. Rouhl RP, Damoiseaux JG, Lodder J, et al. Vascular inflammation in cerebral small vessel disease. *Neurobiol Aging* 2012;33:1800–1806.
44. Degoma EM, Rader DJ. Novel HDL-directed pharmacotherapeutic strategies. *Nat Rev Cardiol* 2011;8:266–277.
45. Johns DG, Duffy J, Fisher T, et al. On- and off-target pharmacology of torcetrapib: current understanding and implications for the structure activity relationships (SAR), discovery and development of cholesteryl ester-transfer protein (CETP) inhibitors. *Drugs* 2012;72:491–507.
46. Rodriguez-Gutierrez R, Shah ND, Montori VM. Predicting the overuse of PCSK-9 inhibitors. *JAMA* 2015;314:1909–1910.
47. McTaggart F, Jones P. Effects of statins on high-density lipoproteins: a potential contribution to cardiovascular benefit. *Cardiovasc Drugs Ther* 2008;22:321–338.