BRAIN
ORIGINAL ARTICLE

Genetically predicted on-statin LDL response is associated with higher intracerebral haemorrhage risk

Ernst Mayerhofer,[1,2,3,4,5] Rainer Malik,[6] Livia Parodi,[1,2,3,4] Stephen Burgess,[7]
Andreas Harloff,[5] Martin Dichgans,[6,8,9] Jonathan Rosand,[1,2,3,4]
Christopher D. Anderson[1,2,3,4,10,†] and Marios K. Georgakis[1,2,3,4,6,†]

†These authors contributed equally to this work.
See Katsanos and Shoamanesh (https://doi.org/10.1093/brain/awac259) for a scientific commentary on this article.

Statins lower low-density lipoprotein cholesterol and are widely used for the prevention of atherosclerotic cardiovascular disease. Whether statin-induced low-density lipoprotein reduction increases risk of intracerebral haemorrhage has been debated for almost two decades. Here, we explored whether genetically predicted on-statin low-density lipoprotein response is associated with intracerebral haemorrhage risk using Mendelian randomization. Using genomic data from randomized trials, we derived a polygenic score from 35 single nucleotide polymorphisms of on-statin low-density lipoprotein response and tested it in the population-based UK Biobank. We extracted statin drug and dose information from primary care data on a subset of 225 195 UK Biobank participants covering a period of 29 years. We validated the effects of the genetic score on longitudinal low-density lipoprotein measurements with generalized mixed models and explored associations with incident intracerebral haemorrhage using Cox regression analysis. Statins were prescribed at least once to 75 973 (31%) of the study participants (mean 57 years, 55% females). Among statin users, mean low-density lipoprotein decreased by 3.45 mg/dl per year [95% confidence interval (CI): (−3.47, −3.42)] over follow-up. A higher genetic score of statin response [1 standard deviation (SD) increment] was associated with significant additional reductions in low-density lipoprotein levels [−0.05 mg/dl per year, (−0.07, −0.02)], showed concordant lipidomic effects on other lipid traits as statin use and was associated with a lower risk for incident myocardial infarction [hazard ratio per SD increment 0.98 95% CI (0.96, 0.99)] and peripheral artery disease [hazard ratio per SD increment 0.93 95% CI (0.87, 0.99)]. Over a 11-year follow-up period, a higher genetically predicted statin response among statin users was associated with higher intracerebral haemorrhage risk in a model adjusting for statin dose [hazard ratio per SD increment 1.16, 95% CI (1.05, 1.28)]. On the contrary, there was no association with intracerebral haemorrhage risk among statin non-users (P = 0.89). These results provide further support for the hypothesis that statin-induced low-density lipoprotein reduction may be causally associated with intracerebral haemorrhage risk. While the net benefit of statins for preventing vascular disease is well-established, these results provide insights about the personalized response to statin intake and the role of pharmacological low-density lipoprotein lowering in the pathogenesis of intracerebral haemorrhage.

1 Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA
2 Program in Medical and Population Genetics, Broad Institute of Harvard and the Massachusetts Institute of Technology, Boston, MA, USA
3 Henry and Allison McCance Center for Brain Health, Massachusetts General Hospital, Boston, MA, USA
4 Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
5 Department of Neurology and Neurophysiology, Medical Center—University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany
6 Institute for Stroke and Dementia Research (ISD), University Hospital, Ludwig-Maximilians-University (LMU) Munich, Munich, Germany

7   University of Cambridge, MRC Biostatistics Unit, Cambridge, UK
8   German Center for Neurodegenerative Diseases (DZNE, Munich), Munich, Germany
9   Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
10 Department of Neurology, Brigham and Women's Hospital, Boston, MA, USA

Correspondence to: Marios K. Georgakis, MD, PhD
Center for Genomic Medicine
Massachusetts General Hospital
185 Cambridge Street, Boston, MA 02114, USA
E-mail: mgeorgakis@mgh.harvard.edu; marios.georgakis@med.uni-muenchen.de

Introduction

Intracerebral haemorrhage (ICH) is a devastating disease associated with a 50% 30-day mortality and major disability among survivors.[1,2] HMG-CoA-reductase inhibitors, commonly known as statins, reduce low-density lipoprotein (LDL) and are widely used for prevention of atherosclerotic cardiovascular disease.[3] The role of LDL in the pathogenesis of ICH, and whether statin intake increases ICH risk, has been a matter of continued debate[4,5] with conflicting data from observational studies and *post hoc* analyses of clinical trials.[5–20]

Human genetic data are a valuable resource for unravelling the role of specific mechanisms in the pathogenesis of diseases. Because genetic liability to polygenic traits is randomly assigned at birth, using genetic variants that are reliably associated with a trait of interest but do not vary with correlated confounders can reduce bias in associations between genetically predicted traits and outcomes.[21–25] Applying this framework, called Mendelian randomization (MR), previous studies have explored whether genetically predicted levels of lipids influence the risk of ICH.[26–29] However, these findings do not address the specific question of whether statin-induced LDL lowering actually increases ICH risk, because the genetic variants capture lifelong small effects on blood lipid levels and not the much stronger short-term effects that result from taking a medication in adulthood.

Expanding the MR concept, genetic variants associated with response to a drug could be used for stratification of individuals in observational studies. Because the innate drug sensitivity is unknown to physicians at the time of prescription, it could be used as instrument for randomizing participants to different levels of drug exposure. This could determine dose-dependent effects of specific drugs on potential side-effects or for exploring repurposing opportunities with the use of observational data, thus overcoming key limitations of conventional MR analyses.[30] In this study, we applied this concept to study the relationship between statin use and ICH risk by leveraging large-scale genetic data for on-statin LDL response from clinical trials and population-based observational data from the UK Biobank (UKB).

Materials and methods

Study population

We used data from the UKB, a population-based prospective cohort study of 502 419 UK residents aged 40–69 years recruited between 2006 and 2010 from 22 assessment centres across the UK.[31] A wide range of phenotyping assessments, biochemical assays, genome-wide genotyping and ongoing longitudinal follow-up data are available for most study participants. For the purposes of the current analyses, we restricted our sample to 46% of the study participants (*n* = 231 336) with detailed linked electronic medical records from their primary care general practitioners. These primary care data include medication prescriptions for a time period ranging from as early as 1978 until 2019, thus allowing a detailed assessment of duration and dose of statin intake both before and after baseline assessments. We excluded 15 137 individuals with missing genetic data and 155 individuals with a history of ICH at baseline (defined as presence of the illness code 'brain haemorrhage', Fig. 1).

The UKB has institutional review board approval from the Northwest Multi-Center Research Ethics Committee (Manchester, UK). All participants provided written informed consent. We accessed the data following approval of an application by the UKB Ethics and Governance Council (Application No. 36993).

Preparation of the UK Biobank primary care data

We extracted data on statin prescriptions and LDL measurements from the UKB primary care data. For obtaining statin exposure metrics, we harmonized the dosages of different statins on the basis of comparison factors from trials evaluating statin efficacy and calculated a mean statin dose per participant across the different prescriptions in the equivalent atorvastatin dose.[3,32–34] The data extraction and quality control process are described in detail in the Supplementary material. According to the 2018 AHA guidelines on cholesterol management, statin intensity was categorized as low (<10 mg), medium (≥10 mg and <40 mg) and high (≥ 40 mg) on the basis of the atorvastatin equivalent dose.[3]

Polygenic score for estimating on-statin LDL response

We used data from the Genomic Investigation of Statin Therapy Consortium, a two-stage genome-wide association study (GWAS) for on-statin LDL cholesterol response among 40 914 statin-treated subjects of European ancestry (30 246 from 10 randomized controlled trials and 10 668 from 11 observational studies),[35] to construct a polygenic score of LDL lowering following statin intake. There was no participant overlap between those studies and the UKB. Following a previously described approach, we used a set of 35 single
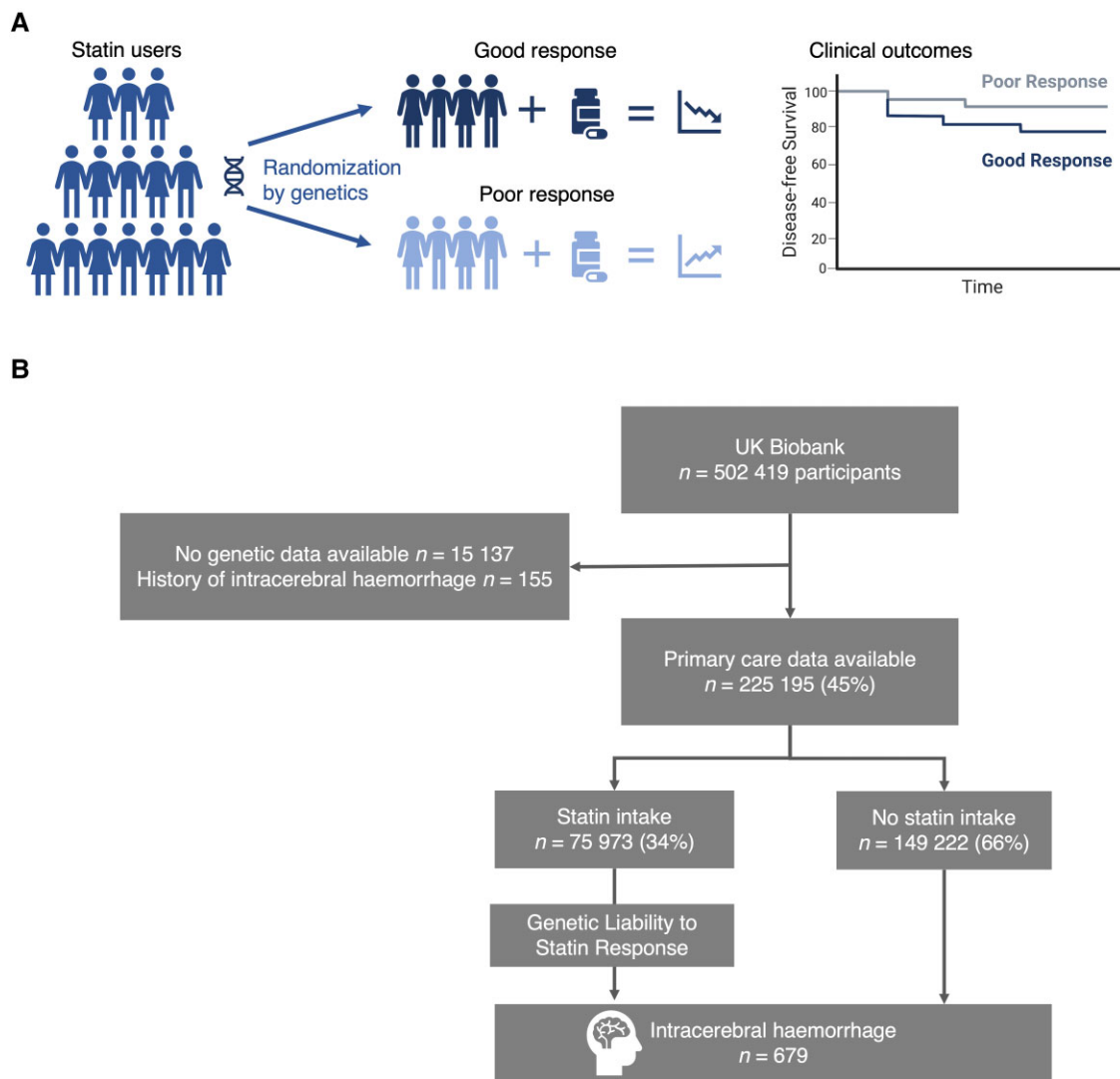
Figure 1 Study concept and study population. (**A**) Individuals respond differently to statins according to their genetic profile.[35] Assuming that there is no selection pressure, drug response is assorted randomly within a population and can be used to explore causal effects of the drug on outcomes using observational data. (**B**) Flow chart of the study participants. Individuals without genetic data or history of ICH at baseline were excluded.

nucleotide polymorphisms (SNPs), selected on the basis of associations with on-statin LDL lowering at $P < 5 \times 10^{-5}$ and clumped at $r^2 < 0.001$ on the basis of the European reference panel of the 1000 Genomes.[36] We then calculated a genetic score with the imputed genotype data of UKB. To confirm that the observed effects were specifically due to genetically predicted on-statin LDL and not off-statin LDL, in sensitivity analyses we tested the association of each of the 35 SNPs included in the score and LDL levels measured at the UKB baseline assessment among those who had never used statins (linear regression models adjustments for age, sex, principal components (PC) 1–10 of population structure, kinship and genotyping assay) and removed the SNPs that associated with off-statin LDL at $P < 0.0014$ (0.05/35 according to Bonferroni). All SNPs for the genetic and the alternative score are provided in Supplementary Table 3.

Validation of statin response genetic scores on LDL trajectories

To test the relevance assumption of MR, we aimed to confirm the effect of the genetic score used on on-statin LDL levels by exploring associations with longitudinal LDL level changes in the primary care data among statin users. Only participants with at least one LDL measurement before and one measurement after their first recorded statin prescription (off- and on-statin LDL) were included in this analysis ($n = 40 633$, 53% of statin users). To account for multiple LDL values per participant over time, we used a mixed model clustered by participant with LDL levels as the outcome and the genetic score, time and their interaction as the exposure. The modes were further adjusted for age, sex, statin equivalency dose, PC1–10, race, kinship and genotyping assay.

Influence of the genetic scores on baseline LDL and lipid particle metabolites

To explore whether a higher genetic score for on-statin LDL lowering mimics an exposure to higher statin intake, we compared associations of a higher score and a higher statin dose with the entire spectrum of 228 lipid particle metabolites among statin users, as measured by nuclear magnetic resonance at baseline. We constructed linear regression models with each metabolite as outcome

and the genetic scores or statin equivalent dose as exposure. The models were further adjusted for age and sex; those with the genetic score as exposure were further adjusted for PC1–10, race, kinship and genotyping assay. We corrected for multiple hypothesis testing with the Bonferroni method setting a significance threshold at 0.05/228. Correlations in the derived estimates for the genetic scores or statin equivalent dose across the lipid traits were tested with Pearson's correlation.

Outcome ascertainment

UKB participants' records have been linked with inpatient hospital codes, primary care data and death registry for longitudinal follow-up. Incident ICH was defined as events occurring after baseline, documented in either hospital admissions or death registry data by the following International Classification of Diseases (ICD) codes: ICD-9 431.X and ICD-10 I61. These criteria were aligned with the diagnostic algorithm for stroke in the UKB (https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/alg_outcome_stroke.pdf) that captured events up to December 2018. We manually applied the same criteria to capture events occurring thereafter up to the end of follow-up (June 2020). Types of intracranial haemorrhage other than ICH were not studied. As positive controls, we also tested associations of the genetic scores with incident myocardial infarction (MI) and peripheral artery disease (PAD), which were defined on the basis of the following ICD-10 codes: I21.X, I22.X, I23.X, I24.1, I25.2 (for MI) and I70.0, I70.00, I70.01, I70.2, I70.20, I70.21, I70.8, I70.80, I70.9, I70.90, I73.8, I73.9 (for PAD).

Effect of the genetic score on ICH and cardiovascular end points

To explore the effects of on-statin genetically predicted LDL response on risk for incident ICH, we used Cox proportional hazard models adjusted for previously published risk factors for ICH[1,2] and genetic covariates: age, sex, BMI, smoking status, history of diabetes, systolic blood pressure, mean statin dose, mean LDL levels, use of anticoagulation and antiplatelet drugs at baseline, PC1–10, race, kinship and genotyping assay. As positive controls, we explored associations between on-statin genetically predicted LDL response and risk for MI and PAD using similar Cox models, additionally adjusting for history of hypertension, hypercholesterolaemia, MI, stroke or PAD without adjusting for antiplatelet and anticoagulation intake. To test the independence and exclusion restriction assumptions of MR and exclude the possibility that any associations are driven by pleiotropic effects of the score independent of on-statin LDL lowering, we tested the same associations among non-statin users. Because statin users are, due to indication bias, at higher risk for MI and PAD, the selection of the population on the basis of statin use could have introduced collider bias. To address this issue, in a sensitivity analysis we used inverse probability weighting to confirm our findings for MI and PAD. Specifically, in the full UKB cohort, we constructed a linear regression model with statin use as outcome and age, sex, BMI, smoking status, hypertension, systolic blood pressure, history of diabetes, intake of diabetes drugs, hypercholesterolaemia, LDL, history of MI, stroke or PAD and the genetic score as covariates. For statin users, we then used the inverse of the fitted values of that model as weights in the respective Cox models to account for the probability of statin prescription in an individual.

Software used

For SNP extraction, genetic score calculation, SNP association tests and relationship inference we used PLINK, bcftools and KING.[37–40] For data extraction, curation, preparation and figure generation, we used RStudio 2021.09.0 with R v.4.1.1 on Mac OS X (aarch64-apple-darwin20) with the packages coxphw, data.table, dplyr, FSA, ggplot2, gmodels, lmerTest, lme4, PheWAS, readr, readxl, stringr, survival, survivalAnalysis, survminer, tidyr and writexl.[41] Figure 1 was partly created with BioRender.com. The analysis plan followed the STROBE-MR statement for the usage of MR in observational studies.[42]

Data availability

The data that support the findings of this study are available from the UKB on submission of a research proposal. The summary statistics of the GWAS for on-statin LDL response used to create the tested polygenic risk score are publicly available.[31,35]

Results
Baseline characteristics

A total of 225 195 of the UKB participants had available genetic, primary care and outcome data, as well as no history of ICH at baseline and were thus eligible for inclusion in the analysis (Fig. 1). Baseline characteristics and outcome data of participants included in the analyses are presented in Table 1. A total of 4 151 471 statin prescriptions for six statin agents were extracted from the primary care data. The majority of the prescriptions referred to simvastatin (60.8%) or atorvastatin (30.4%) and the number of statin prescriptions increased over time. We extracted at least one statin prescription for 75 973 of the participants (33.7%) with available primary care data. Of the statin users, 41% had prescriptions for two or more different drugs at different time points indicating a medication switch. The distributions of statin prescriptions over time, age, dose and presence of vascular risk factors are depicted in Fig. 2 and the detailed distribution of different statins is shown in Supplementary Table 3. Of all prescriptions, 7.2% accounted for a low (<10 mg), 79.9% for a medium (≥10 and <40 mg) and 12.9% for a high (≥40 mg) atorvastatin equivalency dose with proportion of individuals prescribed a statin and statin dose increasing with age. Similarly, statin use and higher statin doses were more common among individuals with more vascular risk factors (active smoking, diabetes, hypertension, hypercholesterolaemia, prevalence of MI, stroke or PAD, or age >65 years) (Fig. 2).

Genetic score for on-statin LDL and LDL trajectories in primary care data

To validate the genetic scores for on-statin LDL lowering in the UKB, we extracted LDL measurements recorded in the primary care data. A total of 46 909 participants (62% of the total statin users) had at least one LDL value before and one after their first statin prescription. There were on average 8.1 ± 5.1 measurements spanning a total of 27.4 years (8.7 ± 4.2 years between the first and last measurements). The mean pre-statin LDL was 147.7 ± 38 mg/dl and the mean post-statin LDL 133.0 ± 42.7 mg/dl. The mean LDL decreased significantly over time [−3.45 mg/dl per year, 95% CI: (−3.47, −3.42)] among statin users. In a mixed linear model adjusting for age and sex, the genetic score and statin dose both had significant effects on absolute LDL levels [−2.3 mg/dl for each SD increase

Table 1 Participant characteristics

Baseline characteristics	All	Statin users	Statin non-users
All, *n*	225 195	75 973	149 222
Female, *n* (%)	122 793 (54.5)	32 491 (42.8)	90 302 (60.5)
European genetic ancestry, *n* (%)	201 856 (89.7)	67 968 (89.5)	133 888 (89.8)
Age, years, mean ± SD	56.5 ± 8.1	60.4 ± 6.6	54.6 ± 8.0
BMI, kg/m^2, mean ± SD	27.5 ± 4.8	28.9 ± 4.9	26.8 ± 4.6
Ever smokers, *n* (%)	101 098 (44.9)	40 382 (53.4)	60 717 (40.8)
Former smoker, *n* (%)	77 469 (34.4)	30 884 (40.9)	46 585 (31.3)
Current smoker, *n* (%)	23 630 (10.5)	9498 (12.6)	14 132 (9.5)
Hypertension, *n* (%)	60 036 (26.7)	36 202 (47.7)	23 834 (16.0)
Systolic blood pressure, mmHg, mean ± SD	141.3 ± 20.6	150.0 ± 20.4	136.9 ± 19.3
Diastolic blood pressure, mmHg, mean ± SD	84.5 ± 11.3	88.1 ± 11.3	82.7 ± 10.9
Hypercholesterolaemia, *n* (%)	27 573 (12.2)	24 708 (32.5)	2865 (1.9)
LDL, mg/dl, mean ± SD	137.8 ± 33.9	134.4 ± 39.9	139.5 ± 30.2
HDL, mg/dl, mean ± SD	56.3 ± 14.7	52.6 ± 13.9	58.0 ± 14.7
Diabetes, *n* (%)	11 370 (5.0)	9798 (12.9)	1527 (1.1)
Antiplatelet use, *n* (%)	33 056 (14.6)	23 971 (31.6)	9085 (6.1)
Anticoagulant use, *n* (%)	2541 (1.1)	1698 (2.2)	843 (0.6)
Incidence of outcomes after enrolment			
ICH, *n* (%)	679 (0.3)	383 (0.5)	296 (0.2)
MI, *n* (%)	12 355(5.5)	8756 (13.0)	3599 (2.4)
PAD, *n* (%)	1711 (0.8)	1378 (1.8)	333 (0.2)
Ischaemic stroke, *n* (%)	2742 (1.2)	1973 (2.6)	769 (0.5)

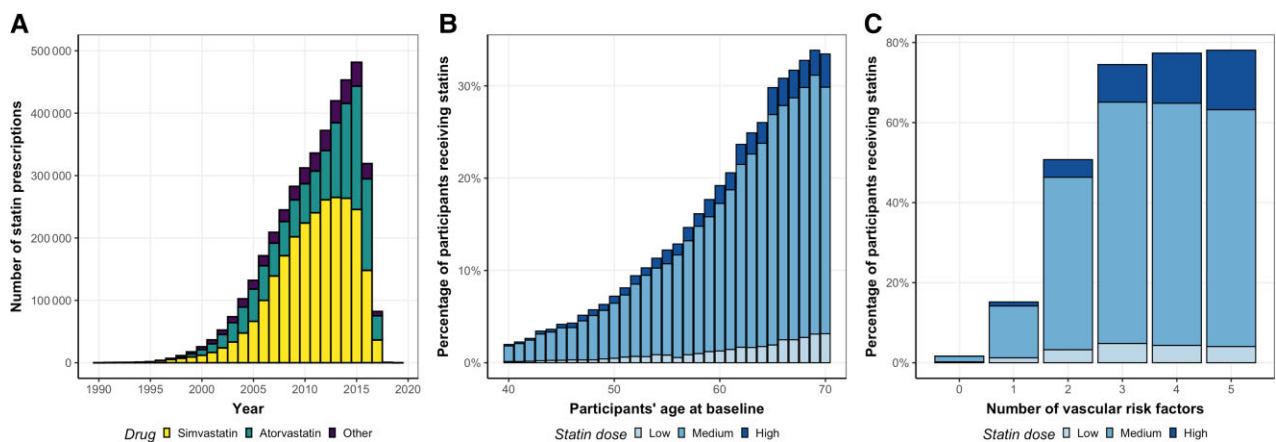HDL = high-density lipoprotein; SD = standard deviation.

Figure 2 Statin prescriptions in the UKB primary care data. (**A**) Number of statin prescriptions per year among 75 973 statin users. (**B**) Percentage of participant's receiving statins per participant's age at baseline. (**C**) Percentage of participants at baseline receiving statins per number of vascular risk factors (diabetes mellitus, hypercholesterolaemia, hypertension, active smoking, age >65 years). Statin intensity in **B** and **C** was classified as low (<10 mg), medium (≥10 mg and <40 mg) and high (≥40 mg) based on the atorvastatin equivalent dose according to the 2018 AHA guideline on cholesterol management.[3] Estimated drug potencies were used to harmonize all drug doses to atorvastatin equivalent doses (see 'Materials and methods' section).

of genetic score, 95% CI: (−2.59, −2.00), Fig. 3A, and −18.8 mg/dl for each SD of statin dose, 95% CI (−18.91, −18.66), Fig. 3B and Supplementary Table 4]. Importantly, there was a significant interaction of the genetic score with time implying a more rapid on-statin LDL decrease among participants with a higher genetic score [−0.05 mg/dl per year for 1 SD of the genetic score, 95% CI (−0.07, −0.02)].

Because we found the genetic score to be also associated with absolute off-statin LDL levels among non-users [−3.0 mg/dl for each SD increase of genetic score, 95% CI: (−3.3, −2.8)], we tested the effect of each SNP and found four of them to be significantly associated with off-statin LDL levels (Supplementary Table 5). Thus,

in a sensitivity analysis, we constructed an alternative genetic score with the remaining 31 SNPs that was no longer associated with off-statin LDL levels among non-users (P > 0.05), but was associated with significant on-statin LDL lowering among statin users [−0.03 mg/dl per year per 1 SD, 95% CI (−0.05, −0.01)].

Influence of the genetic scores on lipid particle metabolites

Next, to explore whether a higher genetic score for on-statin LDL lowering mimics an exposure to higher statin intake at a metabolomic level, we investigated its effects on cholesterol measurements,
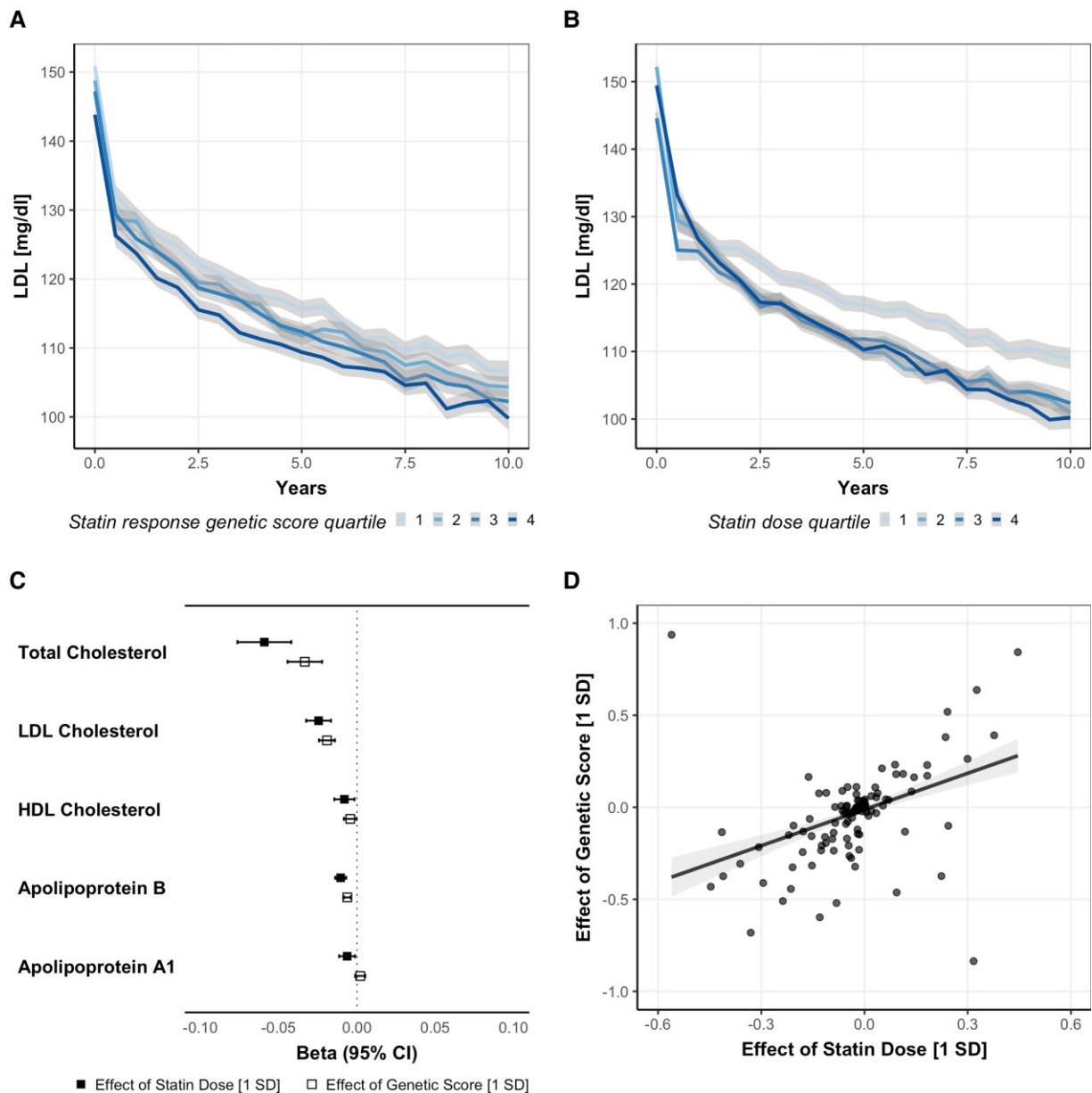
Figure 3 Effect of the genetic score and statin dose on on-statin LDL levels (A and B) and lipid metabolites (C and D) among statin users. LDL trajectories derived from the primary care data of the UKB separated (**A**) by quartiles of the genetic score for LDL response after statin intake and (**B**) by mean statin dose over lifetime in 46 909 individuals with statin intake and at least one LDL measurement before the first statin prescription. (**C**) Estimate and 95% CIs of the effect of statin dose and the genetic score for statin response on total, LDL and high-density lipoprotein cholesterol (mmol/l) and on apolipoprotein B and A1 (g/l) derived from nuclear magnetic resonance among statin users. (**D**) Comparison of the effect size of the genetic score for statin response (1 SD increment) versus the statin dose (1 SD increment) on lipidomic metabolites among statin users. Each point represents the effect sizes for 1 of 228 lipidomic particles. Correlation coefficient $r = 0.65$. The results from **C** and **D** are derived from linear regression models adjusted for age, sex, PC1–10, race, kinship and genetic assay.

as well as lipid particle metabolites, as assessed by standardized methodologies at baseline among statin users. The comparisons of the effect sizes of statin intake on nuclear magnetic resonance-assessed lipid particle metabolites are presented in Fig. 3C and D, respectively. Of the 228 lipid particle metabolites analysed, the genetic score was significantly (Bonferroni-adjusted $P < 0.05$) associated with 161 and a higher statin dose with 97 (Supplementary Table 6). There was a correlation between the effect sizes of the genetic score and statin dose ($r = 0.52$, $P < 0.001$).

Genetically predicted on-statin LDL lowering and risk of incident ICH

Following the validation of the genetic score as a proxy of on-statin LDL lowering, we next tested associations with the risk of incident ICH among statin users (Fig. 4). There were 679 incident ICH over an observation period of 2 514 994 person-years, yielding an incidence of 27 per 100 000 person-years. Over a mean follow-up of 11 years, 383 statin users developed ICH. In Cox proportional hazard models,
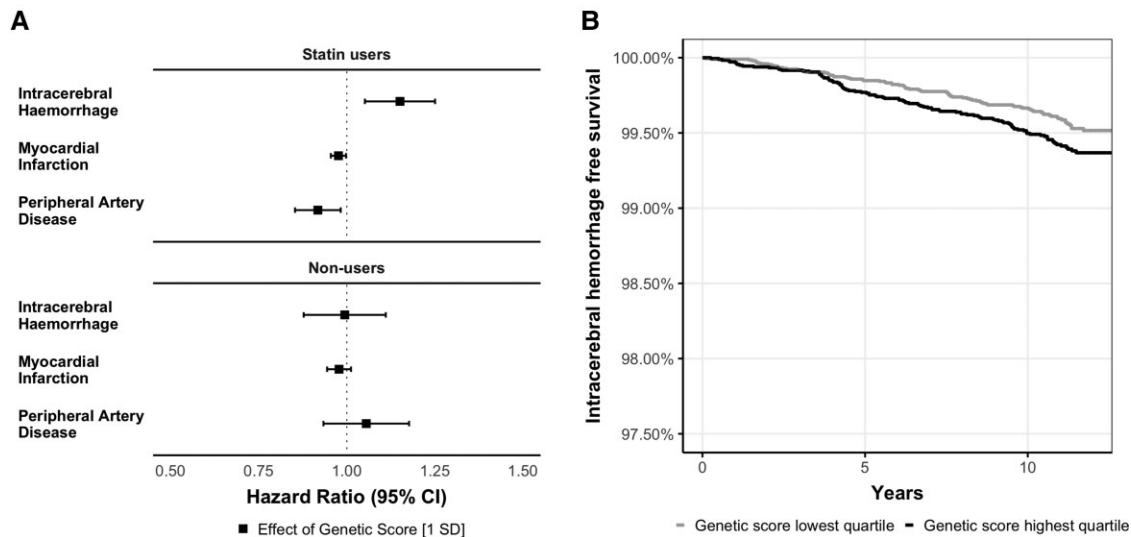
Figure 4 Effect of on-statin genetically predicted LDL response on study outcomes. (**A**) Hazard ratios of statin dose and the genetic scores for statin response among statin users and non-users on ICH, MI and PAD. The results are derived from Cox proportional hazard models adjusted for age, sex, BMI, smoking status, history of diabetes, systolic blood pressure, mean statin dose, LDL levels, PC1–10, race, kinship and genotyping assay; use of anticoagulation and antiplatelet drugs at baseline in models for ICH. (**B**) Kaplan–Meier curves for survival ICH in statin users.

higher genetic scores for on-statin LDL lowering were associated with a higher risk of incident ICH among statin users [hazard ratio (HR) 1.16, 95% CI (1.05, 1.28) for 1 SD difference]. Sensitivity analyses confirmed robustness of the findings among unrelated individuals [kinship coefficient <0.0884, $n = 69\,327$; HR per SD increment 1.18, 95% CI (1.06, 1.31)] as well as with the alternative genetic score not influencing off-statin LDL levels [HR per SD increment 1.12, 95% CI (1.02, 1.24)]. Other variables associated with ICH risk were older age [HR: 1.07 per year, 95% CI: (1.05, 1.09)], lower body mass index [HR: 0.80 per 1 SD, 95% CI: (0.71, 0.90)], higher systolic blood pressure [HR: 1.18 per 1 SD, 95% CI: (1.07, 1.31)] and use of antiplatelet [HR: 1.30 95% CI: (1.06, 1.60)] or anticoagulant [HR: 3.47, 95% CI: (2.30, 5.24)] medications. We found no significant association between mean statin dose and ICH risk [HR: 1.07 per 10 mg atorvastatin equivalent dose, 95% CI: (0.99, 1.15)].

As positive controls, we tested the effects of genetically predicted on-statin LDL lowering on MI and PAD. After adjustment for cardiovascular risk factors, LDL levels and statin dose, we found significant associations of a higher genetic score with a lower risk of both incident MI [HR per SD increment 0.98 95% CI (0.96, 0.99)] and PAD [HR per SD increment 0.93 95% CI (0.87, 0.99)]. To reduce the risk for collider bias, we calculated models weighted for the inverse probability of being prescribed a statin, which yielded effect estimates in the same direction for both MI [HR per SD increment 0.96 95% CI (0.93, 0.99)] and PAD [HR per SD increment 0.84 95% CI (0.70, 0.99)]. Finally, to minimize the possibility that the observed effects are the result of pleiotropy on traits other than on-statin LDL, we also tested the same associations among statin non-users and found no significant effects on ICH, MI or PAD (all P > 0.05).

Discussion

In this study, we used genetic data to stratify statin users by their genetically predicted response to statins and investigated their risk of incident ICH risk. We leveraged data from a GWAS of on-statin LDL lowering from 40 000 statin-treated individuals (75%

clinical trial participants) as well as biochemical, lipidomic and primary care data from 225 000 individuals from a population-based study. We found that a higher genetically predicted LDL response to statins associated with steeper LDL lowering, a similar lipidomic signature as high-dose statin use and a lower risk of atherosclerotic cardiovascular outcomes. In addition, this higher genetically predicted LDL response to statins was associated with a higher risk of ICH among statin users only. There was no such association among individuals who were not taking statins. Our results support a causal effect of more aggressive LDL lowering with statins on risk of ICH and highlight the utility of modelling drug response in addition to dose in examining putative causal associations between biomarkers and outcomes.

Our study extends previous findings from genetic[26,29] and observational analyses,[8–13] providing evidence that beyond lifetime variation in LDL levels, genetic variation in statin-induced LDL lowering also influences ICH risk. This result agrees with *post hoc* analyses of clinical trials supporting a higher risk for haemorrhagic stroke among participants prescribed a high-intensity statin dose.[18] The mechanisms underlying this observation remain poorly understood. It has been speculated that cholesterol is important for vessel integrity, but to date no experimental study has provided evidence for a mechanism connecting low cholesterol levels to vessel damage or loss of vessel structural integrity.[43] As demonstrated in our analyses but also in previous work,[44] statins influence a wide range of lipoprotein particles beyond LDL and thus revealing the main driver of their association with ICH remains a key challenge. Although the follow-up time of existing trials does not exceed 3 years, a meta-analysis did not find aggressive LDL lowering by PCSK9-inhibitors increases ICH risk, even in high-risk patients with previous ischaemic or haemorrhagic stroke, indicating that LDL might not be the sole driver.[17] Because of the widespread lipidomic effect of the genetic score we used, it is not possible from our current analyses to make inferences about which particle class is the causal mediator of this association.

Early clinical trials of statin administration had found a slightly elevated risk for ICH among statin users,[5–7] which was in line with

data from prospective observational studies demonstrating that increased serum total cholesterol and LDL levels are negatively associated with ICH risk in a dose-dependent manner.[8–13] Although subsequent meta-analyses of statin trials found inconsistent results for overall statin use and risk of ICH,[12,14–17] high-dose statin use remained associated with an increased ICH risk.[18] However, *post hoc* analyses from statin trials could not detect statistically significant increases in ICH risk associated with aggressive LDL lowering to <70 mg/dl[19] or <55 mg/dl.[20] These conflicting data about incident ICH among statin users remain a source of concern among medical professionals and are the motivator of the ongoing NINDS-sponsored Statins in Intracerebral Haemorrhage (SATURN) randomized trial (NCT03936361).

By leveraging genetic determinants of response to statin intake, we were able to randomize statin users at the beginning of drug intake, as the prescribing physician is blinded to the genetic variation in statin response. In contrast, when using genetic variants for off-statin LDL or high-density lipoprotein levels in conventional MR approaches,[26–29] randomization is performed at conception and leads to lifelong variations in lipid levels. As such, conventional MR studies have captured lifelong genetically predicted LDL levels and are thus limited in making any inferences about the causal effects of a particular drug prescribed over a shorter timeframe. Our approach overcomes this limitation, facilitating causal inference of the impact of statin intake on ICH using solely observational data. This application could be implemented in other settings as well, and demonstrates the latent utility of additional efforts to develop polygenic predictors of drug response in pharmacogenomic research.[30]

From a methodological perspective, our study also demonstrates the utility of using real-world primary care data for assessing longitudinal trajectories of clinical and biochemical assessments and medication use. Although real-world data are noisier and less standardized than data usually obtained for research purposes, they retain utility to assess drug safety and side-effects, inform clinical trial design and compare drug effectiveness.[45] Leveraging the longitudinal drug prescription and LDL measurement data from primary care data, we were able to track statin prescription and response over a timeframe extending from several years before inclusion of the participants to the study to the end of their follow-up in the UKB. Using data from the rising number of GWAS for drug response,[46,47] future studies could explore in the primary care data from the UKB associations of drug intake with multiple end points. This could allow the detection of previously unreported adverse effects, for which trials are often underpowered,[48] or the investigation of the potential of repurposing opportunities.

Our study has additional specific methodological strengths. Using data from 225 000 participants, including 75 000 statin users and 700 ICH events, we were sufficiently powered to detect meaningful changes in ICH risk by genetically predicted on-statin response. The phenotypic depth of the UKB dataset allowed us to validate the effects of the genetic score statin response on LDL trajectories, lipidomic traits and atherosclerotic end points. Furthermore, we have introduced novel and innovative approaches to leverage GWAS for drug response in large-scale longitudinal population-based datasets. By aggregating data from >4 million drug prescriptions, we were able to precisely phenotype drug intake at an individual level and thus control for statin dose in our outcome models.

Our approach also has limitations. First, the constructed genetic score was associated not only with on-statin LDL lowering but also with off-statin baseline LDL levels. To address this limitation, we introduced an alternative genetic score that was only associated with LDL lowering after statin intake and used that for sensitivity analyses confirming our findings. However, residual confounding due to subthreshold effects of the variants on baseline LDL levels cannot be excluded. Second, we observed a lower incidence of ICH in our study population (27 per 100 000 person-years), as compared to the age-standardized world-wide rate of 42 per 100 000 person-years.[1] This is possibly related to the healthier profile of the UKB population as compared with the general population and necessitates a cautious interpretation of the findings.[49] Third, our study was performed in mainly people of European ancestry and therefore our results cannot be generalized to other populations. Fourth, actual drug intake might also be influenced by poor adherence, which has not been included in our models. Fifth, statins were first introduced in 1988 and prescriptions rose since then, but it was not until 1995 that >90% of the primary care practices in the UK were fully computerized.[50] Sixth, we lacked neuroimaging data from incident ICH events, which would enable stratified analyses by haemorrhage location (lobar versus deep). Seventh, because of the very low number of participants with a prior history of ICH, our study lacked power to explore associations of genetically predicted on-statin LDL response with ICH recurrence. Future studies should focus on exploring the same question among more vulnerable and clinically relevant populations, such as ICH survivors, among whom the balance between the risk of ICH and the prevention of ischaemic cardiovascular events might differ. Finally, because atherosclerotic cardiovascular disease prevention is the main indication for statins, limiting our cohort to statin users might have introduced collider bias for the atherosclerotic endpoints.[51] While we addressed this issue by applying inverse probability weighted models, some relevant bias towards the null might still be present in the measured effect sizes.

In conclusion, we found that higher genetically predicted on-statin LDL response mimics exposure to higher statin doses and increases risk for ICH. These results imply that more aggressive statin-induced LDL lowering might increase risk of ICH and should be balanced against statin benefits in trials of intensive statin treatment. More broadly, our results demonstrate the utility of leveraging genetic data of drug response as a novel method of investigating side-effects and repurposing opportunities of specific drugs with observational data.

Competing interests

J.R. has consulted for Pfizer, Inc. outside of the presented work. C.D.A. has received sponsored research support from Bayer AG and Massachusetts General Hospital and has consulted for ApoPharma, Inc. outside of the presented work. A.H. has received presentation honoraria from AMGEN and Bristol Myers Squibb.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: A systematic analysis for the global burden of disease study 2019. *Lancet Neurol*. 2021;20:795–820.
2. Fernando SM, Qureshi D, Talarico R, *et al*. Intracerebral hemorrhage incidence, mortality, and association with oral anticoagulation use: A population study. *Stroke*. 2021;52:1673–1681.
3. Grundy SM, Stone NJ, Bailey AL, *et al*. 2018 AHA/ACC/AACVPR/ AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation*. 2019;139:e1082–e1143.
4. Chaturvedi S. Lowering cholesterol and intracerebral hemorrhage: There won't be blood! *Stroke*. 2022;53:463–464.
5. Collins R, Armitage J, Parish S, Sleight P, Peto R; Heart Protection Study Collaborative Group. 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.
6. Amarenco P, Bogousslavsky J, Callahan A III, *et al*. High-dose atorvastatin after stroke or transient ischemic attack. *N Engl J Med*. 2006;355:549–559.
7. 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.
8. Ma C, Gurol ME, Huang Z, *et al*. Low-density lipoprotein cholesterol and risk of intracerebral hemorrhage: A prospective study. *Neurology*. 2019;93:e445–e457.
9. Zhang X, Patel A, Horibe H, *et al*. Cholesterol, coronary heart disease, and stroke in the Asia Pacific region. *Int J Epidemiol*. 2003;32:563–572.
10. Iso H, Jacobs DR, Jr., Wentworth D, Neaton JD, Cohen JD. Serum cholesterol levels and six-year mortality from stroke in 350,977 men screened for the multiple risk factor intervention trial. *N Engl J Med*. 1989;320:904–910.
11. Iribarren C, Reed DM, Burchfiel CM, Dwyer JH. Serum total cholesterol and mortality. Confounding factors and risk modification in Japanese-American men. *JAMA*. 1995;273:1926–1932.
12. Law MR, Wald NJ, Rudnicka AR. Quantifying effect of statins on low density lipoprotein cholesterol, ischaemic heart disease, and stroke: systematic review and meta-analysis. *BMJ*. 2003;326:1423.
13. Gu X, Li Y, Chen S, *et al*. Association of lipids with ischemic and hemorrhagic stroke: A prospective cohort study among 267 500 Chinese. *Stroke*. 2019;50:3376–3384.
14. Wang X, Dong Y, Qi X, Huang C, Hou L. Cholesterol levels and risk of hemorrhagic stroke: A systematic review and meta-analysis. *Stroke*. 2013;44:1833–1839.
15. Cholesterol Treatment Trialists' (CTT) Collaboration; Baigent C, Blackwell L, *et al*. Efficacy and safety of more intensive lowering of LDL cholesterol: A meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet*. 2010;376:1670–1681.
16. Amarenco P, Labreuche J. Lipid management in the prevention of stroke: Review and updated meta-analysis of statins for stroke prevention. *Lancet Neurol*. 2009;8:453–463.
17. Sanz-Cuesta BE, Saver JL. Lipid-lowering therapy and hemorrhagic stroke risk: Comparative meta-analysis of statins and PCSK9 inhibitors. *Stroke*. 2021;52:3142–3150.
18. Cheng Y, Qiao L, Jiang Z, *et al*. Significant reduction in the LDL cholesterol increases the risk of intracerebral hemorrhage: A systematic review and meta-analysis of 33 randomized controlled trials. *Am J Transl Res*. 2020;12:463–477.
19. Amarenco P, Kim JS, Labreuche J, *et al*. Intracranial hemorrhage in the TST trial. *Stroke*. 2022;53:457–462.
20. Masson W, Lobo M, Siniawski D, Masson G, Lavalle-Cobo A, Molinero G. LDL-C levels below 55 mg/dl and risk of hemorrhagic stroke: A meta-analysis. *J Stroke Cerebrovasc Dis*. 2021; 30:105655.
21. Sanderson E, Glymour MM, Holmes MV, *et al*. Mendelian randomization. *Nat Rev Methods Primers*. 2022;2:6.
22. Smith GD, Ebrahim S. 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32:1–22.
23. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. *JAMA*. 2017;318:1925–1926.
24. Georgakis MK, Gill D. Mendelian randomization studies in stroke: exploration of risk factors and drug targets with human genetic data. *Stroke*. 2021;52:2992–3003.
25. Dichgans M, Pulit SL, Rosand J. Stroke genetics: discovery, biology, and clinical applications. *Lancet Neurol*. 2019;18:587–599.
26. Falcone GJ, Kirsch E, Acosta JN, *et al*. Genetically elevated LDL associates with lower risk of intracerebral hemorrhage. *Ann Neurol*. 2020;88:56–66.
27. Georgakis MK, Malik R, Anderson CD, Parhofer KG, Hopewell JC, Dichgans M. Genetic determinants of blood lipids and cerebral small vessel disease: Role of high-density lipoprotein cholesterol. *Brain*. 2020;143:597–610.
28. Anderson CD, Falcone GJ, Phuah CL, *et al*. Genetic variants in CETP increase risk of intracerebral hemorrhage. *Ann Neurol*. 2016;80:730–740.
29. 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.
30. Bowden J, Pilling LC, Turkmen D, Kuo CL, Melzer D. The triangulation WIthin a STudy (TWIST) framework for causal inference within pharmacogenetic research. *PLoS Genet*. 2021;17: e1009783.
31. Sudlow C, Gallacher J, Allen N, *et al*. UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12: e1001779.
32. Weng TC, Yang YH, Lin SJ, Tai SH. A systematic review and meta-analysis on the therapeutic equivalence of statins. *J Clin Pharm Ther*. 2010;35:139–151.
33. Adams SP, Tiellet N, Alaeiilkhchi N, Wright JM. Cerivastatin for lowering lipids. *Cochrane Database Syst Rev*. 2020;1: CD012501.
34. Stein EA. Extending therapy options in treating lipid disorders: A clinical review of cerivastatin, a novel HMG-CoA reductase inhibitor. *Drugs*. 1998;56:25–31; discussion 33.
35. Postmus I, Trompet S, Deshmukh HA, *et al*. Pharmacogenetic meta-analysis of genome-wide association studies of LDL cholesterol response to statins. *Nat Commun*. 2014;5:5068.
36. Smit RAJ, Trompet S, Leong A, *et al*. Statin-induced LDL cholesterol response and type 2 diabetes: A bidirectional two-sample

Mendelian randomization study. *Pharmacogenomics J.* 2020;20: 462–470.

37. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics.* 2011;27:2987–2993.

38. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4:7.

39. v2.00a2.3LM 64-bit Intel (24 Jan 2020). www.cog-genomics.org/plink/2.0/

40. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics.* 2010;26:2867–2873.

41. RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA. Version 1.4.1106 2021. http://www.rstudio.com/

42. Skrivankova VW, Richmond RC, Woolf BAR, *et al.* Strengthening the reporting of observational studies in epidemiology using Mendelian randomization: The STROBE-MR statement. *JAMA.* 2021;326:1614–1621.

43. Lauer A, Greenberg SM, Gurol ME. Statins in intracerebral hemorrhage. *Curr Atheroscler Rep.* 2015;17:46.

44. Richardson TG, Leyden GM, Wang Q, *et al.* Characterising metabolomic signatures of lipid-modifying therapies through drug target Mendelian randomisation. *PLoS Biol.* 2022;20: e3001547.

45. Rudrapatna VA, Butte AJ. Opportunities and challenges in using real-world data for health care. *J Clin Invest.* 2020;130:565–574.

46. Kamide K, Asayama K, Katsuya T, *et al.* Genome-wide response to antihypertensive medication using home blood pressure measurements: A pilot study nested within the HOMED-BP study. *Pharmacogenomics.* 2013;14:1709–1721.

47. Singh S, Warren HR, Hiltunen TP, *et al.* Genome-wide meta-analysis of blood pressure response to beta1-blockers: Results from ICAPS (International Consortium of Antihypertensive Pharmacogenomics Studies). *J Am Heart Assoc.* 2019;8:e013115.

48. Tsang R, Colley L, Lynd LD. Inadequate statistical power to detect clinically significant differences in adverse event rates in randomized controlled trials. *J Clin Epidemiol.* 2009; 62:609–616.

49. Fry A, Littlejohns TJ, Sudlow C, *et al.* Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol.* 2017;186:1026–1034.

50. Millman A, Lee N, Brooke A. ABC Of medical computing. Computers in general practice–I. *BMJ.* 1995;311:800–802.

51. Tonnies T, Kahl S, Kuss O. Collider bias in observational studies. *Dtsch Arztebl Int.* 2022;119:107–122.