medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Genetic variation supports a causal role for valproate in prevention of ischemic stroke

Authors

Ernst Mayerhofer, MD[1-3], Livia Parodi, PhD[1-4], Kaavya Narasimhalu, MD, PhD[1-3], Stefan Wolking, MD[5], Andreas Harloff, MD[6], Marios K Georgakis MD, PhD[1-3,7], Jonathan Rosand, MD, MSc[1-3], Christopher D Anderson, MD, MMSc[1-4]

Affiliations

[1] Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA
[2] Program in Medical and Population Genetics, Broad Institute of Harvard and the Massachusetts Institute of Technology, Boston, MA, USA
[3] McCance Center for Brain Health, Massachusetts General Hospital, Boston, MA, USA
[4] Department of Neurology, Brigham and Women's Hospital, Boston, MA, USA
[5] Department of Neurology and Epileptology, University Hospital Aachen, Germany
[6] Department of Neurology and Neurophysiology, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Germany
[7] Institute for Stroke and Dementia Research (ISD), University Hospital, Ludwig-Maximilians-University (LMU) Munich, Munich, Germany

Corresponding author

Christopher D Anderson, Brigham and Women's Hospital, Boston, cdanderson@partners.org

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Valproate is a candidate for ischemic stroke prevention due to its anti-atherosclerotic effects in vivo. Although valproate use is associated with decreased ischemic stroke risk in observational studies, confounding by indication precludes causal conclusions. To overcome this limitation, we applied Mendelian randomization to determine whether genetic variants that influence seizure response among valproate users associate with ischemic stroke. We derived a genetic score for valproate response using genome-wide association data of seizure response after valproate intake from the Epilepsy Pharmacogenomics Consortium. We then tested this score among valproate users of the UK Biobank for association with incident and recurrent ischemic stroke using Cox proportional hazard models. Among 2,150 valproate users (mean 56 years, 54% females), 82 ischemic strokes occurred over a mean 12-year follow-up. Higher valproate response genetic score was associated with higher serum valproate levels (+5.78 µg/ml per one SD, 95% CI [3.45, 8.11]). After adjusting for age and sex, higher valproate response genetic score was associated with lower ischemic stroke risk (HR per one SD 0.73, [0.58, 0.91]) with a halving of absolute risk in the highest compared to the lowest score tertile (4.8% vs 2.5%, p-trend=0.027). Among 194 valproate users with prevalent stroke at baseline, a higher valproate response genetic score was associated with lower recurrent ischemic stroke risk (HR per one SD 0.53, [0.32, 0.86]) with reduced absolute risk in the highest compared to the lowest score tertile (3/51, 5.9% vs. 13/71, 18.3%, p-trend=0.026). The valproate response genetic score was not associated with ischemic stroke among the 427,997 valproate non-users (p=0.61), suggesting minimal pleiotropy. In an independent cohort of 1,241 valproate users of the Mass General Brigham Biobank with 99 ischemic stroke events over 6.5 years follow-up, we replicated our observed associations between the valproate response genetic score and ischemic stroke (HR per one SD 0.77, 95% CI: [0.61, 0.97]). These results demonstrate that a genetically predicted favorable seizure response to valproate is associated with higher serum valproate levels and reduced ischemic stroke risk among valproate users, providing causal support for valproate effectiveness in ischemic stroke prevention. The strongest effect was

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

found for recurrent ischemic stroke, suggesting potential dual-use benefits of valproate for post-stroke epilepsy. Clinical trials will be required in order to identify populations that may benefit most from valproate for stroke prevention.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Valproate is a widely used antiepileptic drug that has been associated with decreased risk for ischemic stroke in observational studies.[1-4] Valproate is assumed to exert this preventive effect by inhibiting histone deacetylase 9 (HDAC9),[5,6] which in animal studies has been found to lead to a stabilizing and anti-inflammatory effect on atherosclerotic plaques.[7-9] Although genetic variants in the *HDAC9* gene have been repeatedly and robustly associated with large-artery stroke in population-based and case-control genome-wide association studies (GWAS),[10,11] causal evidence supporting a role for valproate in stroke prevention through this mechanism in humans is still missing. One of the reasons for this gap is that observational studies are prone to biases and thus cannot deliver evidence for a causal drug effect.[12] While randomized clinical trials provide the needed evidence, they are often tailored to a specific indication and can be underpowered for secondary endpoints or uncommon side effects,[13] making efficient evaluation of drug repurposing challenging.

Germline genetic variation, constant throughout the lifespan and thus not prone to confounding, can be leveraged to assess whether a drug causally contributes to a specific outcome. If there are genetic variants that are known to influence drug response, the drug users in an observational study can be divided according to the predicted genetic response. Because the prescribing health care providers and the patients are not aware of these genetic variants at the time of prescription, the cohort of drug users can be considered randomized and blinded by genetics. If the genetic variants that predispose to better drug response are associated with the outcome of interest in users of the drug, the results support the hypothesis that the drug causally contributes to the outcome. Pleiotropic effects of the genetic variants that have an independent effect on the outcome can be ruled out if no associations are found among non-drug users. We have recently applied this form of *in silico* simulation of a randomized controlled trial, a special case of Mendelian randomization, to show that statins causally contribute to intracerebral hemorrhage.[14]

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Utilizing this framework, we investigated the causal contribution of valproate use on risk of incident and recurrent ischemic stroke. To gain insight into clinically relevant effects on atherosclerosis, we additionally investigated the effect on myocardial infarction. We leveraged data from a previously published GWAS of clinical response to three antiepileptic drugs (valproate, lamotrigine, and levetiracetam)[15] to construct and validate a genetic score for valproate response among valproate users in the UK Biobank (UKB) and study its association with the selected outcomes **(Figure 1A)**. To rule out the possibility that detected associations are driven by an antiepileptic drug class effect, we also used the same approach to study the effects of lamotrigine and levetiracetam on ischemic stroke. To assess the robustness of our findings, we replicated observed associations in an independent cohort of valproate users of the Mass General Brigham Biobank (MGBB).

Methods

Study population

The UKB is an ongoing, population-based prospective cohort study of over 500,000 individuals aged 37-73 years at baseline, recruited from 2006-2010 in 22 assessment centers across the UK.[16] A wide range of baseline data including phenotyping assessments, biochemical assays, genome-wide genotyping, and primary care data (in a subset of the total cohort) together with ongoing follow-up outcome data is available. Only individuals with available genetic data were included in the present study. Because we used genetic variants[15] and a linkage disequilibrium reference panel[17] that were both discovered only in individuals of European ancestry (see below), we further restricted our cohort to European ancestry participants.

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).

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Mendelian randomization approach

Because the genetic variants that are used for the exposure (in our study, valproate, lamotrigine, and levetiracetam clinical response) are derived from a GWAS including only individuals exposed to those medications, but stroke and myocardial infarction outcome GWAS have been performed among drug users and non-users, we used an individual-level approach in the UK Biobank to test for drug-specific effects and assess for pleiotropic effects of our genetic instruments. We constructed a genetic score for response to each drug and tested each score for association with the outcomes of interest among individuals exposed or not exposed to each medication. Because of the random assortment of common alleles in a population, genetically predicted drug response is randomly allocated, and thus an association of the genetic response score with the outcome of interest in those exposed to the drug provides evidence of a causal drug effect. Further, pleiotropic effects of the genetic variants that inadvertently modify the risk for chosen outcomes independent of the drug can be ruled out if there is no association among individuals not exposed to the drug. This approach has been described by us and others in previously published work.[14,18]

Identification of antiepileptic drug users

We identified valproate, lamotrigine, and levetiracetam drug users via verbal baseline interview and primary care prescription data, allowing us to capture drug prescriptions within a time period between 1978 to 2018. We have previously reported the details of our pipeline to extract medication data from UKB primary care data.[14] First, all available ever-approved formulations for valproate, lamotrigine, and levetiracetam were gathered by using international nonproprietary (INN) names, former and current trade names in the UK (via the National Health Service Dictionary of Medicines and Devices [DM+D] browser, https://services.nhsbsa.nhs.uk/dmd-browser/search), and their associated DM+D and British National Formulary (BNF) codes (**Supplemental Table S1**). Then, all primary care prescription data and the verbal interview data were searched for these formulations. Individuals were considered as users of a drug if they reported intake of one of the formulation names containing the drug at any verbal

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

interview (baseline or follow-up, UKB field 20003) or if they had two or more prescriptions of a formulation containing the drug in the primary care data (gp_scripts table). The first drug prescription date for each individual was defined either as the first prescription date from primary care data or as the verbal interview date, whichever was earlier and available. To assess whether patients were on monotherapy or had other antiepileptic drugs prescribed, we extracted prescriptions for the most common antiepileptic drugs (**Supplemental Table S1**) between first prescription of valproate and end of follow-up.

Construction of the genetic scores for antiepileptic drug response

We used genome-wide association data from the Epilepsy Pharmacogenomics Consortium (EpiPGX) on seizure freedom after antiepileptic drug intake in European ancestry patients with generalized epilepsy.[15] In that study, participants were defined as treatment responder if they were seizure free under continuous treatment for at least one year, and as treatment non-responders if they had 50% of pretreatment seizure frequency or higher under adequate dosing of the drug according to a specialist.[15] The cohort included patients on valproate (n=565), lamotrigine (n=387), and levetiracetam (n=209).[15] Association tests were performed based on responder vs. non-responder status for each of the drugs.[15] There was no participant overlap between the EpiPGX GWAS and the UKB. Association results were available for single nucleotide polymorphisms (SNPs) associated with drug response at $p<0.05$ (n=162,242 for valproate, n=162,666 for lamotrigine, and n=162,430 for levetiracetam).

To construct the genetic scores to be used as instruments for valproate, lamotrigine, and levetiracetam response, we leveraged PRS-CS (polygenic prediction via bayesian regression and continuous shrinkage priors), a novel unsupervised polygenic prediction method for that uses a high-dimensional Bayesian regression framework to derive a genetic score from GWAS summary statistics without requiring an external validation cohort.[17] PRS-CS takes linkage disequilibrium of genetic variants into account by using an external linkage disequilibrium reference panel and outperforms traditional clumping and thresholding approaches such as PRSice.[17,19] We used PRS-CS with default parameters

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

to generate SNP weights for response to each drug, which yielded weights for 33,089, 33,300, and 32,736 SNPs for valproate, lamotrigine, and levetiracetam, respectively.

To test robustness of the discovered associations, we performed sensitivity analyses with an alternative genetic instrument for valproate response that was derived using a clumping and thresholding approach. Following previously described approaches for drug response Mendelian randomization studies,[14,20] we selected the SNPs from the EpiPGX GWAS results that were associated with valproate response at $p<5x10^{-5}$ and clumped at $r^2<0.001$ based on the 1000 Genomes European reference panel.[20] The alternative genetic score for valproate response consisted of 20 SNPs that were retained after clumping of 139 SNPs.

Finally, using the weights derived from PRS-CS and clumping and thresholding, we calculated the individual genetic scores for corresponding drug users in the UKB using imputed genotype data. For assessment of appropriate randomization, Kaplan-Meier curves, and calculation of absolute risk differences, individuals were divided in genetic score tertiles. The SNPs and weights for the genetic scores are provided in **Supplemental Tables S2-S5**.

Validation of the genetic score on serum valproate response to normalized valproate dosing

We aimed to test the effect of the derived genetic score for valproate response on valproate serum levels to confirm its validity to test the hypothesis whether valproate has an effect on ischemic stroke through serum level-dependent effects. Genetic variants that are associated with seizure response to valproate could be unrelated to the effect of valproate on ischemic stroke, if they influence a pathway that is not related to drug metabolism, but rather further downstream in its effect on seizure prevention. However, if the genetic variants predict valproate response through an effect on valproate serum levels below the threshold of impacting prescriber behavior, they are a proxy for genetically predicted drug exposure and thus can be used as an instrument for randomization in the test for an effect on ischemic

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

stroke. In this special case of Mendelian randomization, this is assertion of the relevance assumption of the genetic variants.

Valproate serum levels were gathered from the primary care clinical data by using the Read codes '44W4.' and 'XE25d'. Values that were 0 (indicating off-valproate situations) and those higher than 200 µg/ml (potentially erroneous or not in µg/ml) were discarded. For each serum level value, the taken valproate dose at the time of measurement was approximated. First, the duration in days between the prescriptions before and after the date of the serum level measurement was calculated. Then, the quantity of tablets was multiplied by the dose of each tablet, divided by the duration of the prescription interval, yielding the average daily dose in mg per day. The association of valproate dose with valproate serum levels was tested in a linear regression model with valproate serum level as dependent variable and average daily valproate dose and the genetic score as independent variables. The model was additionally adjusted for age at the time of serum level measurement and sex. Levels for lamotrigine or levetiracetam were not available in the UK Biobank.

Outcome ascertainment

UKB participants' records have been linked with inpatient hospital codes, primary care data, and death registry for longitudinal follow-up. Outcome events were gathered from hospital admissions and death registry data using International Classification of Diseases (ICD) 10 codes that were aligned with the diagnostic algorithm in the UKB (https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/alg_outcome_main.pdf). Incident events were defined as events occurring after baseline and after the first drug prescription date. Because only very limited information on ischemic vs. hemorrhagic stroke subtypes before baseline exists in the UKB, recurrent ischemic strokes were defined as ischemic strokes occurring in individuals with a history of any stroke at baseline, as defined in the UKB field 42006. The ICD-10 codes used for ascertainment of the outcomes are supplied in **Supplemental Table S6.** Stroke outcomes in the UKB have been routinely used in genetic association studies, including the most recent GWAS of stroke risk.[11]

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Association of the genetic scores with outcomes

To explore the effects of valproate on the selected outcomes, cause-specific Cox proportional hazard models censored for death were used with the valproate-specific genetic scores as independent variable, adjusted for age, sex, principal components (PC) 1-3, and genotyping assay in the cohort of valproate users. Although ischemic stroke subtypes are not available in the UKB, we tried to investigate the pathophysiological mechanism of valproate's action on stroke prevention. To approximate cardioembolic stroke, ischemic stroke in the setting of atrial fibrillation was analyzed by adding an interaction term of the genetic score with prevalent atrial fibrillation in the model for ischemic stroke and performing subgroup analyses in individuals with and without a diagnosis of atrial fibrillation before or within six months after ischemic stroke. Because of the small cohort size, the main analyses were performed in all valproate users regardless of potential cryptic relatedness, and sensitivity analyses were performed in a cohort restricted to unrelated individuals (KING kinship coefficient < 0.0884). Chi-Square test for trend in proportions were used to assess significant differences in absolute stroke risk between genetic score tertiles. To rule out that observed associations between the genetic scores and the outcomes are caused by pleiotropic effects of the genetic score not related to valproate use, we tested the same associations among non-valproate users, thus assessing the independence and exclusion restriction assumptions of Mendelian randomization. To further rule out an antiepileptic drug class effect, all analyses were repeated with the genetic scores for lamotrigine and levetiracetam response among users of the respective drugs. To exclude drug interactions, we also performed a sensitivity analysis among the cohorts of patients on valproate monotherapy.

Replication analyses

We aimed to confirm our discoveries in the MGBB, an ongoing prospective clinical research cohort of patients of Mass General Brigham (MGB), the parent organization of Massachusetts General Hospital (MGH) and Brigham and Women's Hospital (BWH) in Boston, Massachusetts, USA. All patients aged 18 years or older presenting to any of the MGB clinics consenting to broad research are included.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Patients are recruited in-person at MGH and BWH and online through an electronic patient gateway. The MGBB provides ongoing electronic health record data with ICD codes for outcomes, imaging, and, on a subset of individuals, genetic data. Recruitment has been ongoing since 1998 and to date, more than 133,000 patients have been included in the Biobank and over 65,000 have been genotyped. Genotyping has been performed in batches on different arrays, with batches 1-9 performed by MGH and batches 10-13 by the Broad Institute. To minimize batch effects and variation across different genotyping arrays, many of the individuals that have been genotyped in batches 1-9 have been re-genotyped in batch 13. For the current analyses, we considered only individuals that were genotyped in batches 10-13, which we combined and imputated together on the Michigan Imputation Server using the Haplotype Reference Consortium v.1.1 reference panel. We used exactly the same PRS-CS approach to construct the genetic score in MGBB participants.

We queried the MGB Biobank database using the same criteria that we used in the UKB. We identified patients with available genetic data and two or more oral prescriptions of valproate. Most of the prescriptions contained instruction texts from which we could extract the daily dose by multiplying the medication dose by the prescribed number of times taken daily. Outcomes were identified using inpatient ICD-9 and ICD-10 codes (**Supplemental Table S6**) and were classified as prevalent or incident events regarding their occurrence before or after the date of the first valproate prescription. We used the same statistical analysis approaches for the replication analyses in the MGBB as in the UKB. Linear regression models were used to assess the associations of the average valproate dose and the genetic score with valproate serum levels. For the association of the genetic score with outcomes, Cox proportional hazard models were constructed with the time to event as the number of days between the date of the first valproate prescription and incident event date for patients for which an outcome occured, and the number of days between the first valproate prescription and the last encounter for patients for which no outcome occured. We detected significant associations between the principal components and age and sex, most likely due to chance imbalance across genotyping batches; people

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

in later batches were better powered and consisted of older individuals and more women, and thus models adjusted for age and sex and the principal components introduced collinearity to our model. Because we found no associations between the genetic score for valproate response with age and sex, indicating near-perfect randomization, we removed age and sex from our Cox models and the reported effect estimates are from models with the genetic score as main predictor, adjusted for principal components 1-3 which still contain age and sex information.

Software and statistical methods used

ANOVA and Chi-Square test were used for comparison of continuous and discrete baseline variables, respectively. SNP extraction and genetic score calculation was performed with PLINK and bcftools, relationship inference was performed with KING.[21-24] Data extraction, curation, preparation, statistical analysis, and figure generation was done with RStudio on Mac OS X.[25] This study followed the guideline for the usage of Mendelian Randomization in observational studies (STROBE-MR).[26]

Data availability

UK Biobank participant data are available through the UK Biobank after approval of a research proposal. The SNPs used for construction of the genetic scores were obtained from the authors of the original publication.[15]

Results

Baseline characteristics

We identified 2,150 valproate users after exclusion of non-European ancestry individuals and those with unavailable genetic data (**Figure 1B, Table 1**). The date range for the prescriptions was July 1987 to March 2018, and the mean time between first and last prescription was 6.5 ± 6.9 years. Valproate users had a significant higher rate of vascular risk factors compared to valproate non-users, but no difference in the genetic score (**Table 1**). The genetic score for valproate response was normally distributed among valproate users (**Figure 2A**). When comparing valproate users stratified by genetic

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

score, no differences in baseline characteristics, cardiovascular risk factors, antiplatelet and statin use, or approximated valproate doses (among the 1,387 individuals with available prescription data) were found across genetic score tertiles, indicating an appropriate randomization (**Table 2**). Most patients were on valproate monotherapy, and all tertiles had average valproate serum levels in the therapeutic range (**Table 2**).

Validation of the genetic score on serum valproate response to normalized valproate dosing

A total of 549 valproate serum level values were available for 202 valproate users after exclusion of 137 values that were zero and 41 values that were higher than 200 (**Figure 2B**). The approximated daily valproate dose was significantly associated with valproate serum levels (1.25 µg/ml per 100mg/day, 95% CI [0.85, 1.66], **Figure 2C**). We found a significant association of the genetic score with valproate serum levels, indicating higher average serum levels in participants with higher genetic scores (+5.78 µg/ml per one SD, 95%CI[3.45, 8.11], **Figure 2C**). These associations were replicated with the alternative genetic score (**Supplemental Table S7**). Serum levels for lamotrigine or levetiracetam were not available for validation of genetic scores.

Association of the valproate genetic instrument with incident ischemic stroke

Among the 2,150 valproate users, 82 ischemic strokes occurred over a mean follow-up of 11.6 years. In Cox proportional hazard models, a higher genetic score for valproate response was associated with a lower risk of incident ischemic stroke (HR 0.73, 95% CI [0.58, 0.91] per one SD increase, **Figure 3A**). Individuals in the lowest tertile of the genetic score had an almost two-fold increased absolute risk for ischemic stroke compared to those in the highest score tertile (4.8% vs 2.5%, p-trend = 0.027, **Figure 3B**). Sensitivity analyses confirmed robustness of the findings among the 1,967 unrelated individuals (HR 0.75, 95% CI [0.59, 0.95] for 76 events, **Supplemental Table S8**) as well as with the alternative genetic score for valproate response using a clumping and thresholding approach (HR 0.75, 95% CI [0.59, 0.95], **Supplemental Table S9**). Restricting our cohort to the patients on valproate monotherapy (n=1,675), the associations were replicated but with wider confidence intervals due to lower power

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

(**Supplemental Table S10**). We found no associations between the genetic scores for lamotrigine and levetiracetam response and ischemic stroke in the 1,108 lamotrigine and 789 levetiracetam users, respectively (p=0.76 and p=0.24, **Supplemental Tables S11-S12**).

In the cohort of valproate users, 67 individuals had a diagnosis of atrial fibrillation before or within six months after ischemic stroke, and 24 ischemic strokes occurred among them. No interaction between prevalent atrial fibrillation and the genetic score was found (p=0.40). No association between the genetic score and ischemic stroke was found in this subgroup (p=0.51), however when removing the 67 individuals with atrial fibrillation from the cohort of valproate users, the association of the genetic score with incident ischemic stroke was more pronounced (HR 0.68, 95% CI [0.52, 0.89], **Figure 3A**), suggesting a better utility of valproate for prevention of ischemic stroke not related to atrial fibrillation.

Association of the genetic score with recurrent ischemic stroke

Among the 194 individuals with prevalent stroke at baseline, 22 recurrent ischemic strokes occurred over a mean follow-up of 11.2 years. In Cox proportional hazard models, a higher genetic score was associated with decreased risk for recurrent ischemic stroke (HR 0.53, 95% CI [0.32, 0.86] per one SD increase). Although only few cases in total, individuals in the lowest tertile of the genetic score had a three-fold higher absolute risk for ischemic stroke compared to those in the highest tertile (13/71, 18.3% vs 3/51, 5.9%, p-trend=0.026, **Figure 3C**).

Associations of the genetic score with myocardial infarction

No associations of the genetic score for valproate response were found for myocardial infarction (133 events, p=0.93) among valproate users. Also, no associations were found among valproate non-users, reassuring that our genetic instruments were not affected by horizontal pleiotropy (p=0.18, **Figure 3A)**.

Replication in the MGB Biobank

We identified 1,241 valproate users with available genetic data in the MGBB, among which 99 ischemic stroke events and 126 myocardial infarction events occurred over a median follow-up of 6.7 years. In

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

this independent cohort we found 839 patients with 6,353 valproate serum level measurements. The associations between prescribed valproate dose and the genetic score with valproate serum levels replicated with almost identical effect estimates (**Supplemental Table S13**). We also found significant association between the genetic score for valproate response and ischemic stroke (HR 0.77, 95% CI: [0.61, 0.97]), and no association between the genetic score and myocardial infarction (p=0.68).

Discussion

In this study, we used common genetic variants to stratify 2,150 valproate users in the UKB by their genetically predicted response to valproate and investigated their risk of incident ischemic stroke and myocardial infarction. We leveraged data from a GWAS of seizure control after valproate intake from a cohort of patients with epilepsy and applied it to medication prescription and intake data of 502,000 individuals from a population-based observational cohort study. We found that a higher genetically predicted seizure response to valproate was associated with higher valproate serum levels, indicating that the included genetic variants predispose to greater valproate exposure by affecting valproate metabolism. In addition, this higher genetically predicted response to valproate was associated with a lower risk of ischemic stroke, among valproate users only. There was no such association among individuals who were not taking valproate, and no association among lamotrigine and levetiracetam users, ruling out independent pleiotropic effects of the genetic variants or an antiepileptic drug effect. The robustness of our results is supported by replication in the MGBB biobank, an independent electronic health record database from a different continent that is representative of a clinical population. Our results support a causal effect of valproate on ischemic stroke risk and demonstrate the utility of leveraging genetic data in observational cohorts to model drug response *in silico* for drug repurposing.

Valproate's anticonvulsant effects were discovered in 1963[27] and it is today a commonly used drug for seizure prevention and mood stabilization.[28-30] Valproate is a nonspecific HDAC inhibitor,[6] and with the

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

detection of *HDAC9* as genetic risk locus for large-artery stroke,[7,10,11] the role of valproate for stroke prevention has been postulated. Various studies have investigated the association of valproate with vascular outcomes in observational studies with conflicting findings. Some have found a decreased risk of ischemic stroke[1,2] and myocardial infarction,[2,3,31,32] while others have found an increased risk of stroke.[33,34] A recent meta-analysis found an increased overall stroke risk in patients with epilepsy, but a decreased stroke risk in patients taking valproate compared to other antiepileptic drugs.[4] Our study confirmed the presumed effect of valproate on ischemic stroke, but failed to confirm previously hypothesized associations with myocardial infarction despite better statistical power, suggesting confounding by indication or attrition bias in previous findings. To test its clinical value for ischemic stroke prevention, valproate is currently under investigation in the Sodium Valproate to Prevent Stroke (SOLVE) trial for the prevention of atherosclerosis progression in patients with large-artery stroke in the UK (ISRCTN12685153).

Our results show a decreased risk of stroke in valproate users. We were unable to investigate valproate's effect on specific stroke subtypes because this information is not available in the UKB, limiting our insight into a specific clinical mechanism. However, we found a stronger association in individuals without a diagnosis of atrial fibrillation even though this subgroup contributed many events and thus statistical power, suggesting that valproate's overall ischemic stroke prevention effect is not by preventing cardioembolism. The lack of association with myocardial infarction could suggest that valproate's stroke prevention mechanism acts through other known pathways beyond slowing of atherosclerosis through *HDAC* inhibition, such as protection of the blood brain barrier,[35] increase in ischemic tolerability through neuroprotective effects as demonstrated *in vivo* for brain ischemia, spinal cord injury and traumatic brain injury,[35-37] inhibition of platelet aggregation,[38] or increase in tissue plasminogen activator.[39]

In our analyses, we found the strongest effect of valproate on ischemic stroke among individuals with a history of prior stroke, despite a small cohort size. Existing data shows that post-stroke epilepsy is

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

common, with an overall incidence of up to 7%[40] and a high recurrence rate of up to 12% in five years[41]. If future studies confirm our observation that valproate contributes most to the prevention of recurrent ischemic stroke among stroke survivors, future trials could evaluate its utility as a treatment for post-stroke epilepsy as a dual-use secondary prevention agent.

Our study provides a compelling example of how genetic data can aid in prioritizing drug repurposing targets when observational studies are limited. The most likely reason for the conflicting evidence for the effect of valproate from observational studies is confounding by indication. As shown in our results, valproate users have a higher number of vascular risk factors compared to the general population (**Table 1**). Thus, association of valproate in solely epidemiological models use would yield an increased risk for vascular outcomes in our cohort, leading to a false conclusion from confounding bias. Our study shows that in these cases, Mendelian randomization is a powerful tool for overcoming this challenge, randomizing individuals to similar baseline characteristics **(Table 2)**, thus providing an unbiased approach that is similar to an actual clinical trial. This intriguing concept is enabled through the increasing number of studies providing genetic markers for drug response.[42] Although valproate is a widely used and known drug, its adverse and teratogenic effects[43] reduce its attractiveness for drug repurposing in the large collective of stroke survivors. However, our findings potentially apply to other HDAC inhibitors that have been tested *in vitro* and *in vivo*,[44] providing further justification for stroke prevention trials employing HDAC inhibitors.

Our research illustrates the potential utility and reliability of the under-utilized primary care prescription data within the UKB. It demonstrates the capacity to generate robust and replicable findings. Our investigation yielded nearly identical estimates for the associations between the prescribed dosage of valproate and serum valproate levels within both the UKB and the MGBB, despite these data being sourced from distinct healthcare systems across two different continents, utilizing two disparate methodologies for daily intake assessment. Furthermore, we observed congruent effect estimates between the genetic score and valproate serum levels, irrespective of the absence of standardized

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

measurements. These findings suggest that with sufficiently large cohorts, it is possible to discern true biological effects despite the diversity in electronic health care records and healthcare systems, along with their inherent weaknesses and limitations.

We acknowledge the limitations of our study. First, the SNPs used to construct the genetic score for valproate response were discovered only in individuals of European ancestry[15] and applied to a predominantly European UKB population, raising the question whether our findings can be applied to Non-European populations. Although the leveraged genetic variants mark valproate response, they act as instrument for randomization in our study and thus it is likely that the findings also apply to individuals of non-European racial/ethnic background. Second, although we identified 2,150 valproate users, the absolute number of outcomes was low, restricting the power of our study. Third, we were not able to investigate valproate's effect on stroke subtypes because detailed stroke phenotyping is unavailable in the UKB. Furthermore, it is possible that some of the ischemic strokes among valproate users might have been misclassified seizures with Todd's paralysis. However, for several reasons we believe that it is unlikely that such a misclassification has substantially biased our findings: i) to drive observed associations, the majority of strokes must have been misclassified, which would raise serious concerns about the UKB phenotyping and all its associated research findings, ii) we did not find an association of the genetic response scores for lamotrigine and levetiracetam with ischemic stroke, and iii) we replicated our findings in the MGBB cohort on inpatient stroke diagnoses. Fourth, since we gathered data on valproate prescriptions, and not on valproate use, we cannot be certain that our whole cohort was in fact using valproate, but we postulate that this might only have diluted our effect estimates towards the null. Finally, our study cannot answer the question which patient collective would benefit most from valproate use, however with our limited power we found the largest effect for prevention of recurrent ischemic stroke.

In conclusion, by using an innovative Mendelian randomization approach leveraging genetic data, our study supports a causal role for valproate in the prevention of ischemic stroke, with the largest

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

decrease in risk for recurrent ischemic stroke. Our results provide actionable evidence for the performance of clinical trials with valproate or other *HDAC* inhibitors that might have a more favorable dosing and side effect profile for the prevention of ischemic stroke.

Disclosures

CDA has received sponsored research support from Bayer AG and has consulted for ApoPharma unrelated to this work. JR reports compensation from National Football League and Takeda Development Center Americas for consultant services, all unrelated to this work.

Sources of Funding

CDA is supported by NIH R01NS103924, U01NS069673, American Heart Association 18SFRN34250007 and 21SFRN812095, and the Massachusetts General Hospital McCance Center for Brain Health for this work. MKG is supported by the FöFoLe program of LMU Munich (Reg.-Nr. 1120), the DFG Germany's Excellence Strategy within the framework of the Munich Cluster for Systems Neurology (EXC 2145 SyNergy - ID 390857198), and the Fritz-Thyssen Foundation (Ref. 10.22.2.024MN) JR receives research grants from NIH and the American Heart Association-Bugher Foundation. AH received funding from the the Berta-Ottenstein programm for Advanced Clinician Scientists, Medical Faculty, University of Freiburg, Germany. SW receives funding from the German Research Foundation (WO 2385/2-1).

Acknowledgments

This research has been conducted using the UK Biobank Resource under Application Number 36993. Figure 1A was created with Biorender.com.

References

1. Brookes RL, Crichton S, Wolfe CDA, Yi Q, Li L, Hankey GJ, Rothwell PM, Markus HS. Sodium Valproate, a Histone Deacetylase Inhibitor, Is Associated With Reduced Stroke Risk After

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Previous Ischemic Stroke or Transient Ischemic Attack. *Stroke*. 2018;49:54-61. doi: 10.1161/STROKEAHA.117.016674

2. Olesen JB, Abildstrom SZ, Erdal J, Gislason GH, Weeke P, Andersson C, Torp-Pedersen C, Hansen PR. Effects of epilepsy and selected antiepileptic drugs on risk of myocardial infarction, stroke, and death in patients with or without previous stroke: a nationwide cohort study. *Pharmacoepidemiol Drug Saf*. 2011;20:964-971. doi: 10.1002/pds.2186
3. Dregan A, Charlton J, Wolfe CD, Gulliford MC, Markus HS. Is sodium valproate, an HDAC inhibitor, associated with reduced risk of stroke and myocardial infarction? A nested case-control study. *Pharmacoepidemiol Drug Saf*. 2014;23:759-767. doi: 10.1002/pds.3651
4. Brigo F, Lochner P, Nardone R, Manganotti P, Lattanzi S. Increased risk of stroke and myocardial infarction in patients with epilepsy: A systematic review of population-based cohort studies. *Epilepsy Behav*. 2020;104:106307. doi: 10.1016/j.yebeh.2019.05.005
5. Phiel CJ, Zhang F, Huang EY, Guenther MG, Lazar MA, Klein PS. Histone deacetylase is a direct target of valproic acid, a potent anticonvulsant, mood stabilizer, and teratogen. *J Biol Chem*. 2001;276:36734-36741. doi: 10.1074/jbc.M101287200
6. Gottlicher M, Minucci S, Zhu P, Kramer OH, Schimpf A, Giavara S, Sleeman JP, Lo Coco F, Nervi C, Pelicci PG, et al. Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. *EMBO J*. 2001;20:6969-6978. doi: 10.1093/emboj/20.24.6969
7. Markus HS, Makela KM, Bevan S, Raitoharju E, Oksala N, Bis JC, O'Donnell C, Hainsworth A, Lehtimaki T. Evidence HDAC9 genetic variant associated with ischemic stroke increases risk via promoting carotid atherosclerosis. *Stroke*. 2013;44:1220-1225. doi: 10.1161/STROKEAHA.111.000217
8. Asare Y, Campbell-James TA, Bokov Y, Yu LL, Prestel M, El Bounkari O, Roth S, Megens RTA, Straub T, Thomas K, et al. Histone Deacetylase 9 Activates IKK to Regulate Atherosclerotic Plaque Vulnerability. *Circ Res*. 2020;127:811-823. doi: 10.1161/CIRCRESAHA.120.316743

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

9. Bowes AJ, Khan MI, Shi Y, Robertson L, Werstuck GH. Valproate attenuates accelerated atherosclerosis in hyperglycemic apoE-deficient mice: evidence in support of a role for endoplasmic reticulum stress and glycogen synthase kinase-3 in lesion development and hepatic steatosis. *Am J Pathol*. 2009;174:330-342. doi: 10.2353/ajpath.2009.080385

10. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L, Giese AK, van der Laan SW, Gretarsdottir S, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet*. 2018;50:524-537. doi: 10.1038/s41588-018-0058-3

11. Mishra A, Malik R, Hachiya T, Jurgenson T, Namba S, Posner DC, Kamanu FK, Koido M, Le Grand Q, Shi M, et al. Stroke genetics informs drug discovery and risk prediction across ancestries. *Nature*. 2022;611:115-123. doi: 10.1038/s41586-022-05165-3

12. Gilmartin-Thomas JF, Liew D, Hopper I. Observational studies and their utility for practice. *Aust Prescr*. 2018;41:82-85. doi: 10.18773/austprescr.2018.017

13. 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. doi: 10.1016/j.jclinepi.2008.08.005

14. Mayerhofer E, Malik R, Parodi L, Burgess S, Harloff A, Dichgans M, Rosand J, Anderson CD, Georgakis MK. Genetically predicted on-statin LDL response is associated with higher intracerebral haemorrhage risk. *Brain*. 2022;145:2677-2686. doi: 10.1093/brain/awac186

15. Wolking S, Schulz H, Nies AT, McCormack M, Schaeffeler E, Auce P, Avbersek A, Becker F, Klein KM, Krenn M, et al. Pharmacoresponse in genetic generalized epilepsy: a genome-wide association study. *Pharmacogenomics*. 2020;21:325-335. doi: 10.2217/pgs-2019-0179

16. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, 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. doi: 10.1371/journal.pmed.1001779

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

17. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun*. 2019;10:1776. doi: 10.1038/s41467-019-09718-5

18. 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. doi: 10.1371/journal.pgen.1009783

19. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience*. 2019;8. doi: 10.1093/gigascience/giz082

20. Smit RAJ, Trompet S, Leong A, Goodarzi MO, Postmus I, Warren H, Theusch E, Barnes MR, Arsenault BJ, Li X, et al. Statin-induced LDL cholesterol response and type 2 diabetes: a bidirectional two-sample Mendelian randomization study. *Pharmacogenomics J*. 2020;20:462-470. doi: 10.1038/s41397-019-0125-x

21. 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. doi: 10.1093/bioinformatics/btr509

22. 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. doi: 10.1186/s13742-015-0047-8

23. Shaun Purcell CC. *v2.00a2.3LM 64-bit Intel (24 Jan 2020)*.

24. 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. doi: 10.1093/bioinformatics/btq559

25. Team R. *RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA; 1.4.1106* 2021.

26. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, VanderWeele TJ, Higgins JPT, Timpson NJ, Dimou N, et al. Strengthening the Reporting of

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. *JAMA*. 2021;326:1614-1621. doi: 10.1001/jama.2021.18236

27. Lopez-Munoz F, Baumeister AA, Hawkins MF, Alamo C. The role of serendipity in the discovery of the clinical effects of psychotropic drugs: beyond of the myth. *Actas Esp Psiquiatr*. 2012;40:34-42.

28. Ailani J, Burch RC, Robbins MS, Board of Directors of the American Headache S. The American Headache Society Consensus Statement: Update on integrating new migraine treatments into clinical practice. *Headache*. 2021;61:1021-1039. doi: 10.1111/head.14153

29. Kanner AM, Ashman E, Gloss D, Harden C, Bourgeois B, Bautista JF, Abou-Khalil B, Burakgazi-Dalkilic E, Llanas Park E, Stern J, et al. Practice guideline update summary: Efficacy and tolerability of the new antiepileptic drugs I: Treatment of new-onset epilepsy: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology and the American Epilepsy Society. *Neurology*. 2018;91:74-81. doi: 10.1212/WNL.0000000000005755

30. Yatham LN, Kennedy SH, Parikh SV, Schaffer A, Bond DJ, Frey BN, Sharma V, Goldstein BI, Rej S, Beaulieu S, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. *Bipolar Disord*. 2018;20:97-170. doi: 10.1111/bdi.12609

31. Renoux C, Dell'Aniello S, Saarela O, Filion KB, Boivin JF. Antiepileptic drugs and the risk of ischaemic stroke and myocardial infarction: a population-based cohort study. *BMJ Open*. 2015;5:e008365. doi: 10.1136/bmjopen-2015-008365

32. Olesen JB, Hansen PR, Abildstrom SZ, Andersson C, Weeke P, Schmiegelow M, Erdal J, Torp-Pedersen C, Gislason GH. Valproate attenuates the risk of myocardial infarction in patients with epilepsy: a nationwide cohort study. *Pharmacoepidemiol Drug Saf*. 2011;20:146-153. doi: 10.1002/pds.2073

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

33. Hsieh CY, Lai EC, Yang YH, Lin SJ. Comparative stroke risk of antiepileptic drugs in patients with epilepsy. *Epilepsia*. 2013;54:172-180. doi: 10.1111/j.1528-1167.2012.03693.x

34. Giner-Soriano M, Marsal JR, Gomez-Lumbreras A, Morros R. Risk of ischaemic stroke associated with antiepileptic drugs: a population-based case-control study in Catalonia. *BMC Neurol*. 2021;21:208. doi: 10.1186/s12883-021-02237-1

35. Wang Z, Leng Y, Tsai LK, Leeds P, Chuang DM. Valproic acid attenuates blood-brain barrier disruption in a rat model of transient focal cerebral ischemia: the roles of HDAC and MMP-9 inhibition. *J Cereb Blood Flow Metab*. 2011;31:52-57. doi: 10.1038/jcbfm.2010.195

36. Yu SH, Cho DC, Kim KT, Nam KH, Cho HJ, Sung JK. The neuroprotective effect of treatment of valproic Acid in acute spinal cord injury. *J Korean Neurosurg Soc*. 2012;51:191-198. doi: 10.3340/jkns.2012.51.4.191

37. Tai YT, Lee WY, Lee FP, Lin TJ, Shih CL, Wang JY, Chiu WT, Hung KS. Low dose of valproate improves motor function after traumatic brain injury. *Biomed Res Int*. 2014;2014:980657. doi: 10.1155/2014/980657

38. Kis B, Szupera Z, Mezei Z, Gecse A, Telegdy G, Vecsei L. Valproate treatment and platelet function: the role of arachidonate metabolites. *Epilepsia*. 1999;40:307-310. doi: 10.1111/j.1528-1157.1999.tb00709.x

39. Larsson P, Alwis I, Niego B, Sashindranath M, Fogelstrand P, Wu MC, Glise L, Magnusson M, Daglas M, Bergh N, et al. Valproic acid selectively increases vascular endothelial tissue-type plasminogen activator production and reduces thrombus formation in the mouse. *J Thromb Haemost*. 2016;14:2496-2508. doi: 10.1111/jth.13527

40. Zou S, Wu X, Zhu B, Yu J, Yang B, Shi J. The pooled incidence of post-stroke seizure in 102 008 patients. *Top Stroke Rehabil*. 2015;22:460-467. doi: 10.1179/1074935715Z.00000000062

41. Flach C, Muruet W, Wolfe CDA, Bhalla A, Douiri A. Risk and Secondary Prevention of Stroke Recurrence: A Population-Base Cohort Study. *Stroke*. 2020;51:2435-2444. doi: 10.1161/STROKEAHA.120.028992

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

42. McInnes G, Yee SW, Pershad Y, Altman RB. Genomewide Association Studies in Pharmacogenomics. *Clin Pharmacol Ther*. 2021;110:637-648. doi: 10.1002/cpt.2349

43. Rahman M, Nguyen H. Valproic Acid. In: *StatPearls*. Treasure Island (FL); 2022.

44. Chen X, He Y, Fu W, Sahebkar A, Tan Y, Xu S, Li H. Histone Deacetylases (HDACs) and Atherosclerosis: A Mechanistic and Pharmacological Review. *Front Cell Dev Biol*. 2020;8:581015. doi: 10.3389/fcell.2020.581015

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Tables and Figures

	Valproate users (N=2,150)	Valproate non-users (N=427,997)	p value
Age at baseline, years, mean (SD)	56.4 (8.18)	56.9 (8.00)	0.0106
Female sex, n (%)	1,153 (53.6)	231,608 (54.1)	0.667
Hypertension, n (%)	641 (29.8)	113,259 (26.5)	<0.001
Diabetes, n (%)	156 (7.3)	20,138 (4.7)	<0.001
Hypercholesterolemia, n (%)	335 (15.6)	52,795 (12.3)	<0.001
Smoking status			<0.001
Never Smoker, n (%)	1,111 (51.7)	232,202 (54.3)	
Ex-Smoker, n (%)	702 (32.7)	150,730 (35.2)	
Current Smoker, n (%)	328 (15.3)	43,551 (10.2)	
History of stroke, n (%)	198 (9.2)	6423 (1.5)	<0.001
History of myocardial infarction, peripheral artery disease, or stroke, n (%)	314 (14.6)	27157 (6.3)	<0.001
Statin use at baseline, n (%)	509 (23.7)	70,515 (16.5)	<0.001
Antiplatelet use at baseline, n (%)	428 (19.9)	60,863 (14.2)	<0.001
Genetic score for valproate response, mean (SD)	25.9 (4.07)	25.9 (4.17)	0.85

Table 1. Baseline characteristics of valproate users compared to valproate non-users.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

	Genetic score lowest tertile (N=717)	Genetic score middle tertile (N=717)	Genetic score highest tertile (N=716)	p
Age at baseline, years, mean (SD)	56.6 (8.33)	56.0 (8.15)	56.6 (8.06)	0.298
Female sex, n (%)	386 (53.8)	389 (54.3)	378 (52.8)	0.85
Hypertension, n (%)	213 (29.7)	224 (31.2)	204 (28.5)	0.522
Diabetes, n (%)	48 (6.7)	62 (8.6)	46 (6.4)	0.209
Hypercholesterolemia, n (%)	104 (14.5)	121 (16.9)	110 (15.4)	0.456
Smoking status				0.377
Never Smoker	372 (51.9)	377 (52.6)	362 (50.6)	
Ex-Smoker	233 (32.5)	217 (30.3)	252 (35.2)	
Current Smoker	110 (15.3)	118 (16.5)	100 (14.0)	
History of stroke, n (%)	73 (10.2)	72 (10.0)	53 (7.4)	0.966
History of myocardial infarction, peripheral artery disease, or stroke, n (%)	108 (15.1)	105 (14.6)	101 (14.1)	0.876
Statin use at baseline, n (%)	161 (22.5%)	182 (25.4%)	166 (23.2%)	0.398
Antiplatelet use at baseline, n (%)	132 (18.4%)	153 (21.3%)	143 (20.0%)	0.381
Valproate monotherapy, n (%)	579 (80.8%)	562 (78.4%)	569 (79.5%)	0.538
Genetic score for valproate response, mean (SD)	21.4 (2.18)	25.9 (1.01)	30.4 (2.03)	<0.001
Valproate dose per day, mg, mean (SD)[†]	824 (501)	769 (474)	827 (551)	0.158
Valproate serum levels, µg/ml, mean (SD)*	53.6 (27.0)	58.1 (28.7)	65.8 (32.3)	0.0526

Table 2. Baseline characteristics of valproate users stratified by the genetic score for valproate response. Higher scores indicate higher genetic predisposition for seizure freedom after valproate intake.

[†] Approximated valproate doses were available for 1,387 individuals

* 549 valproate serum levels were available for 202 individuals.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

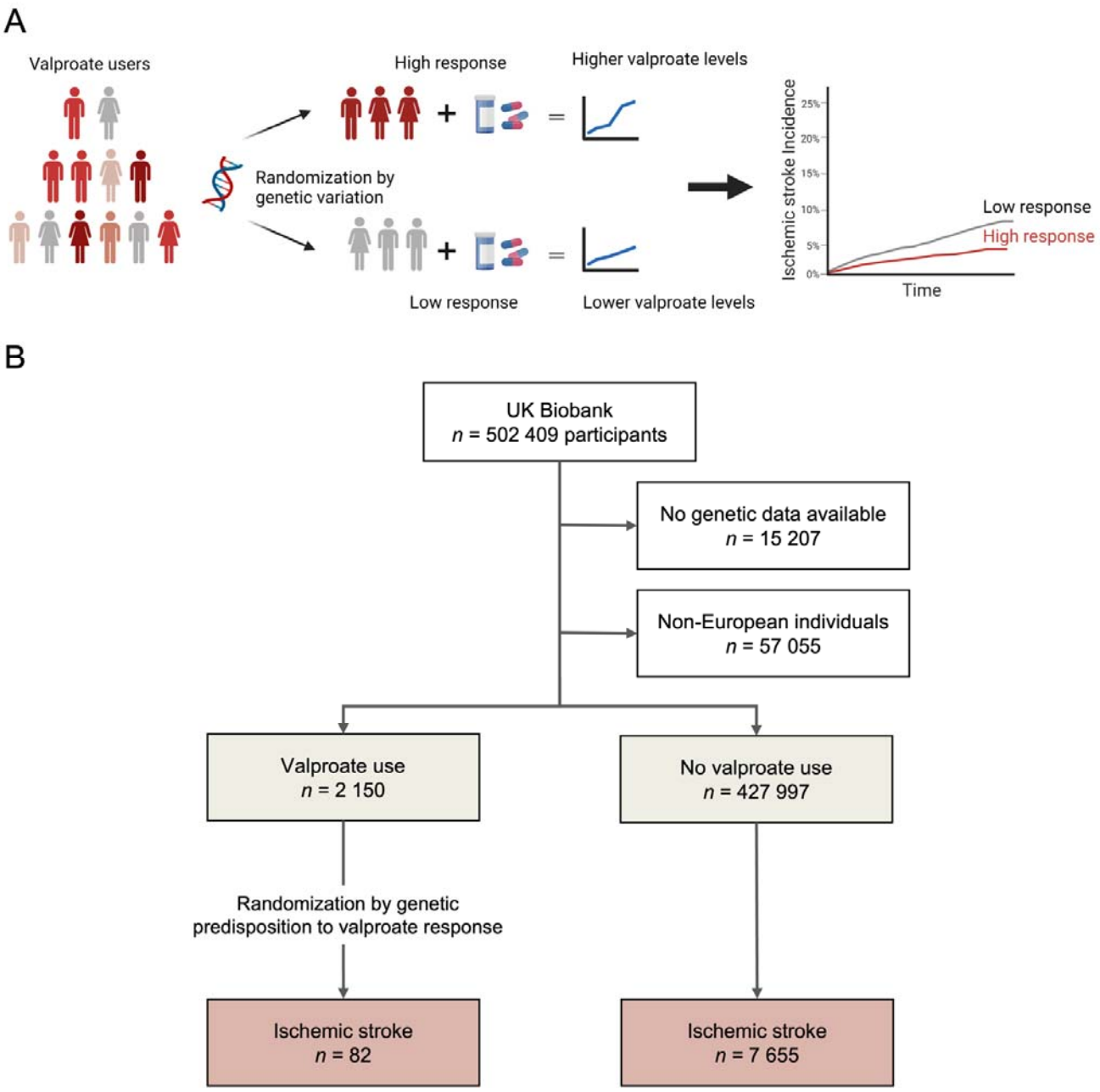

Figure 1. Study overview.

A: Concept of the *in silico* trial randomized by genetic variation. A genetic score consisting of genetic variants known to predispose to higher likelihood of seizure freedom after valproate intake was associated with serum valproate levels and incidence of ischemic stroke.

B: Study flow. After exclusion of Non-European individuals and those without genetic data, 2,150 valproate users and 427,997 valproate non-users were identified.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

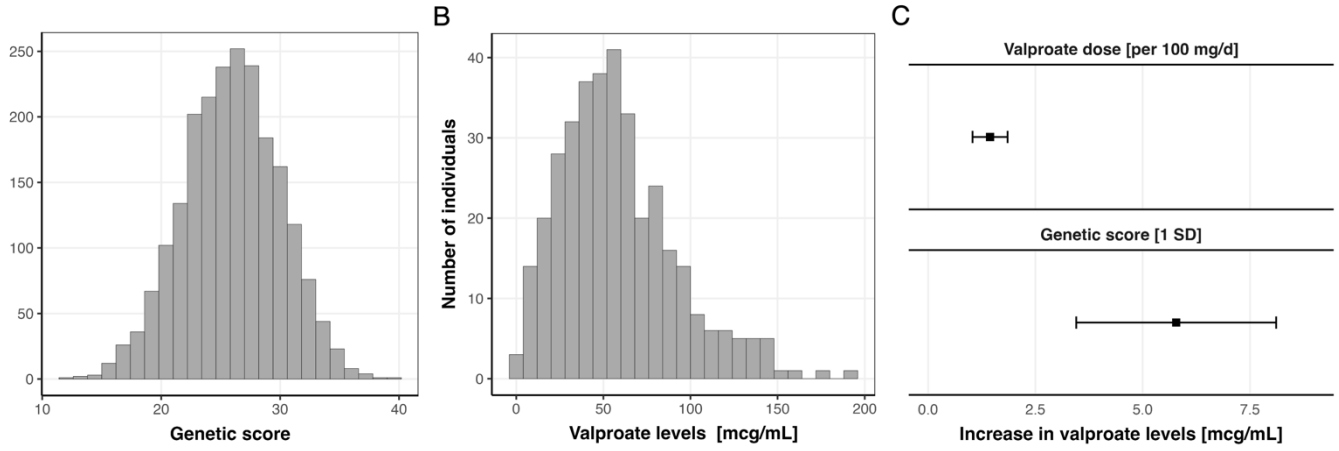

Figure 2. Valproate serum levels and their association with the genetic score for valproate response in the UK Biobank.

A: Distribution of the genetic score among the 2,150 valproate users.

B: Distribution of the 549 valproate serum level values among 202 valproate users.

C: Association of approximated daily valproate dose and the genetic score for valproate response with valproate serum levels in a linear regression model adjusted for age and sex. Depicted in the top is the coefficient for valproate dose, on the bottom it is the genetic score for valproate response, suggesting that the genetically predicted valproate response acts through valproate metabolism.

medRxiv preprint doi: https://doi.org/10.1101/2023.02.14.23285856; this version posted May 30, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

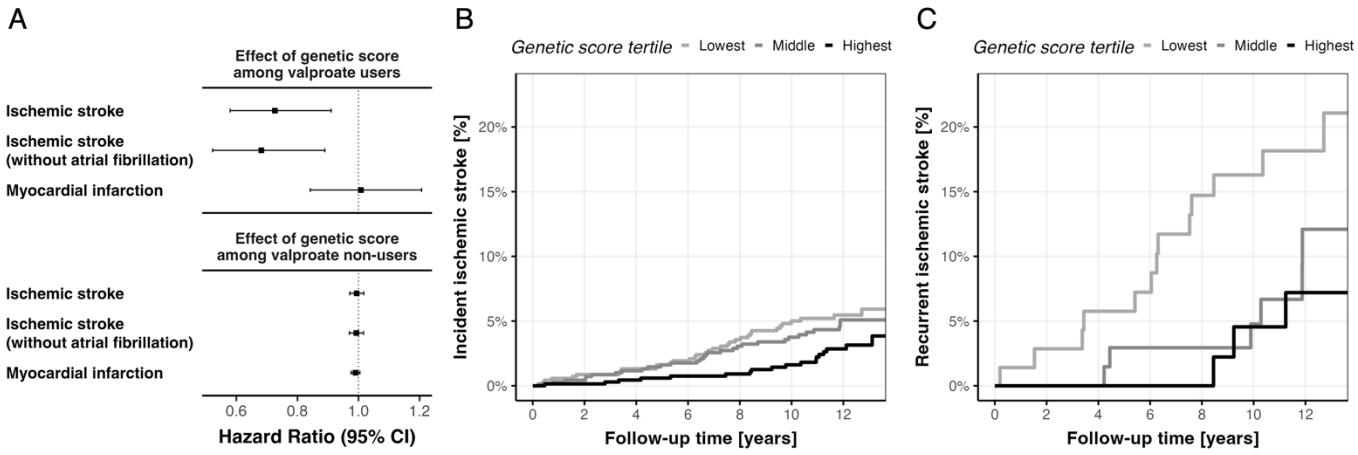

Figure 3. Associations between the genetic score for valproate response with outcomes in the UK Biobank.

A: Hazard ratios of valproate on ischemic stroke, ischemic stroke without concurrent diagnosis of atrial fibrillation, and myocardial infarction.

B and C: Kaplan-Meier plots for incidence of **(A)** 82 ischemic strokes among the 2,150 valproate users and **(B)** 22 recurrent ischemic stroke among the 194 valproate users with prevalent stroke at baseline, stratified by genetic score tertiles.