medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Identification of putative causal relationship between stroke and 1504 complex traits using large-scale phenome-wide screening

Tania Islam[1], Luis M García-Marín[2,3], Miguel E. Rentería[2,3], Gabriel Cuellar-Partida[4], Asaduzzaman Khan[1], Mohammad Ali Moni[1,*]

[1] School of Health and Rehabilitation Sciences, Faculty of Health, The University of Queensland, Brisbane, QLD 4072, Australia

[2]Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

[3]School of Biomedical Sciences, Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia

[4]Diamantina Institute, The University of Queensland & Translational Research Institute, Woolloongabba QLD, Australia

[*]Corresponding author:

Dr. Mohammad Ali Moni

School of Health and Rehabilitation Sciences,

Faculty of Health, The University of Queensland,

Brisbane, QLD 4072, Australia

E-mail: m.moni@uq.edu.au

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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/10.1101/2023.03.08.23287006; this version posted March 12, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Abstract

Stroke is the second leading cause of death and the third leading cause of long-term disability in the world. This study aimed to explore the novel putative causal genetic relationship of stroke with hundreds of complex traits by leveraging genetic data. We used genome-wide association studies (GWAS) data and the latent causal variable method to identify potential causal relationships between stroke and 1,504 complex traits of the UK biobank. We found that 262 traits were genetically correlated with stroke risk at a false discovery rate (FDR <0.05). Of those correlated traits, 28 showed robust evidence of partial genetic causality (GCP) with stroke (|GCP|> 0.60; FDR < 0.05). Our results showed that some conditions, including atrial fibrillation, pulmonary embolism, blood clots in the lung, platelet crit, self-reported deep venous thrombosis, weight gain after depression, and the use of some medications such as insulin, pioglitazone, and gliclazide were inferred to increase stroke risk. On the other hand, greater levels of testosterone, apolipoprotein A, SHBG, and HDL cholesterol decrease the risk of stroke. Also, our results suggest that genetic susceptibility to stroke raises the risk of neck and chest pain and loose teeth. Finally, our findings suggest that cardiac vascular disease, blood clot in lung, deep venous thrombosis, and certain anti-diabetic medications could have a causal role in increasing the risk of stroke, which could be used as novel testable hypotheses for future epidemiological studies.

Keywords: Stroke, latent causal variable method, summary statistic for complex traits, and genome-wide association studies (GWAS).

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

Introduction

Stroke is the second largest cause of mortality and disability worldwide and a leading cause of cognitive impairment and dementia in aged people (1). Stroke affects around 13.7 million individuals, and about 5.5 million die every year (2,3). Stroke is a neurological condition characterised by blockage of a blood vessel. For instance, blood clots can form in the brain, obstructing blood flow and causing blood vessels to rupture, resulting in haemorrhage in the brain. When the brain's arteries rupture during a stroke event, the lack of oxygen causes the brain cells to die (4).

Clinical and epidemiological studies have repeatedly discovered a hereditary component to stroke vulnerability along with more typical risk factors, including smoking, hypertension, and diabetes mellitus (5,6). Studies on twins and family history also provide evidence that genetic susceptibility is significant in stroke aetiology (7). The estimated heritability was 37.9% for all ischemic strokes (8). The heritability of stroke subtypes varied, with small-vessel disease having a lower heritability (16.1%), large-vessel disease having a higher heritability (40.3%), and cardioembolic stroke having a medium heritability (32.6%) (8). In genome-wide association studies (GWAS), several significant risk loci have been identified as associated with stroke (9). For instance, cardioembolic stroke was linked to two atrial fibrillation-related loci (*PITX2* and *ZFHX3*), whereas large-vessel stroke was linked to a locus on chromosome 9p21 that was previously associated with coronary artery disease (9–11). In addition, a multi-ancestry GWAS analysis revealed 35 risk loci that are associated with the aetiology of stroke and its subtypes (12,13). Recently, cross-ancestry meta-analyses of GWAS summary data on stroke identified a total of 89 independent loci significantly associated with stroke and stroke subtypes (14). Mendelian randomisation (MR) was used genetic data to understand the causal role of traits on stroke risk (15). For example, blood pressure, atrial fibrillation, venous thrombosis, circulating lipoproteins, LDL (low-density lipoprotein) cholesterol, obesity, type 2 diabetes, insulin resistance, hyperglycaemia, education, BMI, physical activity, alcohol consumption and smoking have been associated with increased risk of stroke using MR (15,16). Mendelian randomisation (MR) and latent causal variable (LCV) approaches examine the influence of genetic liability for a specific trait on the outcome (17,18). Interestingly, the LCV method represents some advantages over MR methods. Firstly, in contrast to conventional MR methods, LCV can accurately differentiate between genetic correlation and full or partial genetic causality. In addition, positive LCV outcomes are more likely to represent true causal effects. Secondly, the genetic causality proportion (*gcp-value*) can be easily estimated by quantifying the degree of causation, while MR methods, especially MR Egger can be easily confounded by genetic correlations (19).

A genetic correlation might be explained by pleiotropic effects, which can be vertical or horizontal (20). When genetic variants directly influence trait A and trait B, this type of effect is called horizontal pleiotropy (**Fig. 1(A)**). On the other hand, vertical pleiotropy may be seen as a causal cascade where the influence of a genetic variant on a single trait is explained by its impact on another trait (**Fig. 1(B)**) (17). Horizontal pleiotropy could lead to high-false positive results in genetic epidemiology to assess the causality (19,21).

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

Due to the massive public health concerns with serious socio-economic consequences, there is a growing interest in exploring the potential causal genetic risk factors of stroke. Therefore, in the present study, we perform a hypothesis-free exploratory large-scale genetic screening for risk factors of stroke using the LCV method, which is less susceptible to confounding by horizontal pleiotropic effects (19). Our findings support some of the causal associations hypothesized by observational studies and shed light on the relationship between stroke, lifestyle factors, diseases, and health conditions.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

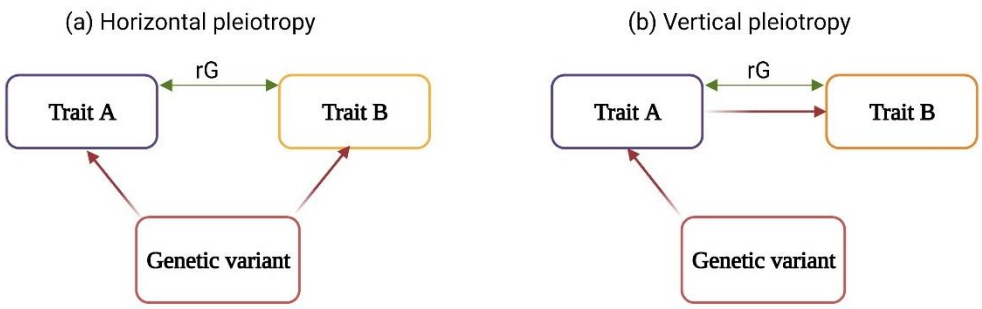

Fig.1. The concept of horizontal and vertical pleiotropic effects on phenotypes (traits). (A) when genetic variants simultaneously affect trait A and trait B then it's called horizontal pleiotropy. (B) when a genetic variant affect trait A and trait A influence the risk of trait B then it's called vertical pleiotropy. rG represents genetic correlation. This figure was generated from Biorender.com.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

Materials and Methods

Stroke data

We obtained GWAS summary statistics for stroke from a meta-analysis of GWAS data by Malik et al. (12) released by the MEGASTROKE consortium (http://megastroke.org/). Briefly, these summary statistics were obtained from a fixed-effects meta-analysis of GWAS in 40,585 stroke cases and 406,111 controls of European ancestry. The details of the study have been described in the original publication (12).

UK biobank cohort datasets

We used 1504 phenotypes available at the Complex Traits Genetics Virtual Lab (CTG-VL) web-based service (https://genoma.io) (22). Most of the GWAS summary statistics accessible in CTG-VL were obtained from Neale's Lab, which collected the UK Biobank datasets (http://www.nealelab.is/uk-biobank). As a result, the majority of these GWAS summary results represent people of European ancestry, which eliminates potential biases due to genetic ancestry variations (22). GWAS data from UK Biobank were corrected for age, age-squared, inferred sex and genetic ancestry (22).

Latent causal variable

The latent causal variable (LCV) is a model used to assess the causal association of two genetically correlated traits. LCV considers a latent variable (L) that has a causal effect on each trait (19). The GCP value indicates if the genetic correlation is due mainly to horizontal or vertical pleiotropy and is calculated by determining the correlations between latent variable L with trait A and trait B respectively (19). A GCP value of 0 indicates that the genetic correlation is mediated by horizontal pleiotropic effects, suggesting the absence of causal genetic effects. On the other hand, $|GCP| = 1$ denotes the presence of vertical pleiotropic effects as well as complete genetic causality. A $|GCP| > 0.6$ is regarded as a robust indicator of potential vertical pleiotropic effects (19). To evaluate LCV outcomes, three things must be considered: the amount of genetic correlation, the magnitude of the GCP estimation, and the directionality of the GCP estimate. The GCP value does not represent the magnitude of prospective causal effects; rather, it represents the proportion of a genetic correlation that could be described by potential causal effects. For example, associations with a low genetic correlation ($|rG| < 0.30$) and a high $|GCP|$ value near 1 indicate that all the genetic elements overlap between two traits, despite its small genetic correlation, which is likely due to vertical pleiotropic effects. Furthermore, a negative GCP value between stroke and trait A indicates that trait A may have causal genetic effects on stroke, whereas a positive GCP value between stroke and trait A indicates that stroke may have causal genetic effects on trait A.

This study estimated the GCP value between stroke and 1504 complex traits. As outlined in previous studies (23,24), we employed the publicly available phenome-wide analysis methodology in the CTG-VL web platform. The phenome-wide analytic framework was built by using an R script for the LCV approach that the original authors of the method made accessible on GitHub (19). We uploaded GWAS summary statistics for stroke to perform the analysis on the CTG-VL web platform. The phenome-wide analytic pipeline was then utilised to

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

perform LD score regression (25) and LCV analyses (19) to assess genetic correlations and potential putative causal links, respectively. Full instructions for utilising and understanding the publicly accessible phenome-wide analysis methodology can be found elsewhere (23,24). The LCV was performed on all traits that showed a genetic correlation with stroke (false discovery rate, FDR<0.05). Similarly, we considered all GCP estimate values that were significant after multiple testing correction (FDR<0.05). The overall analytical approach that we employed in this study is shown in **Fig. 2.**

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

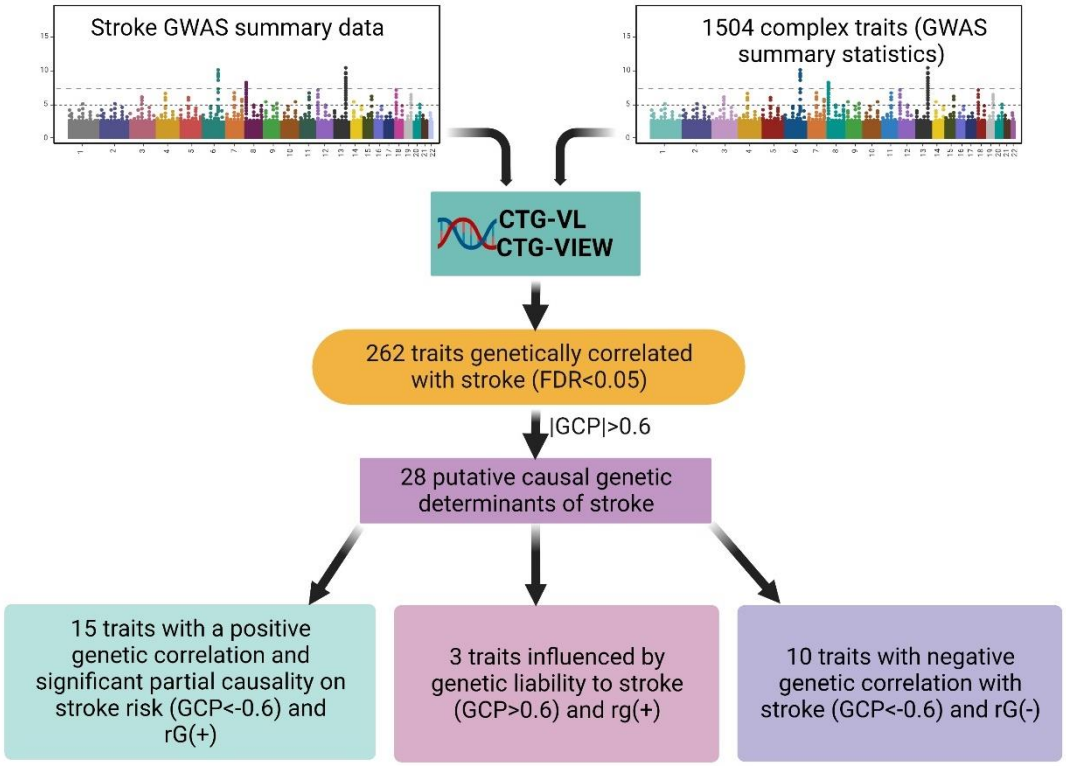

Fig.2. The schematic diagram illustrates the genetic relationship between stroke and 1504 complex traits. The genetic correlation was consistent across the phenotypes, demonstrating that genetic variants specific to each of the 1504 complex traits had consistent positive or negative correlations with the genetic variants associated with stroke. This figure was generated from Biorender.com.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

Results

Of the 1,504 complex traits, we identified 262 traits genetically correlated with stroke risk (FDR<0.05; Supplementary Table (S1). Among those genetically correlated traits, 28 could be explained via robust vertical pleiotropic effects with stroke (|GCP|> 0.60; FDR <0.05), while seven showed limited partial genetic causality (|GCP|< 0.60; FDR < 5%; (Supplementary Table S2 & S3 respectively).

We also found that genetic susceptibility to a variety of diseases, and the use of prescription medication might increase the stroke risk (**Fig.3 and Table 1**). Specifically, we observed that cardiovascular traits, including, self-reported atrial fibrillation, self-reported pulmonary embolism, blood clot in the lung, atrial fibrillation and flutter (ICD10), blood clot in the leg reported by the doctor, deep venous thrombosis of lower extremities and pulmonary embolism, may increase the risk of stroke. Our study also highlighted that genetic susceptibility to stroke may raise the risk of pain in the throat and chest (ICD10), chest pain or discomfort, and loose teeth (**Fig. 3 and Table 2**).

Our study also found four complex traits, which decrease the risk of stroke, including HDL cholesterol, apoliprotein, testosterone and SHBD levels for males and females. These four traits showed negative genetic correlations and significant genetic causal proportions (**Fig. 4 and Table 3**).

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

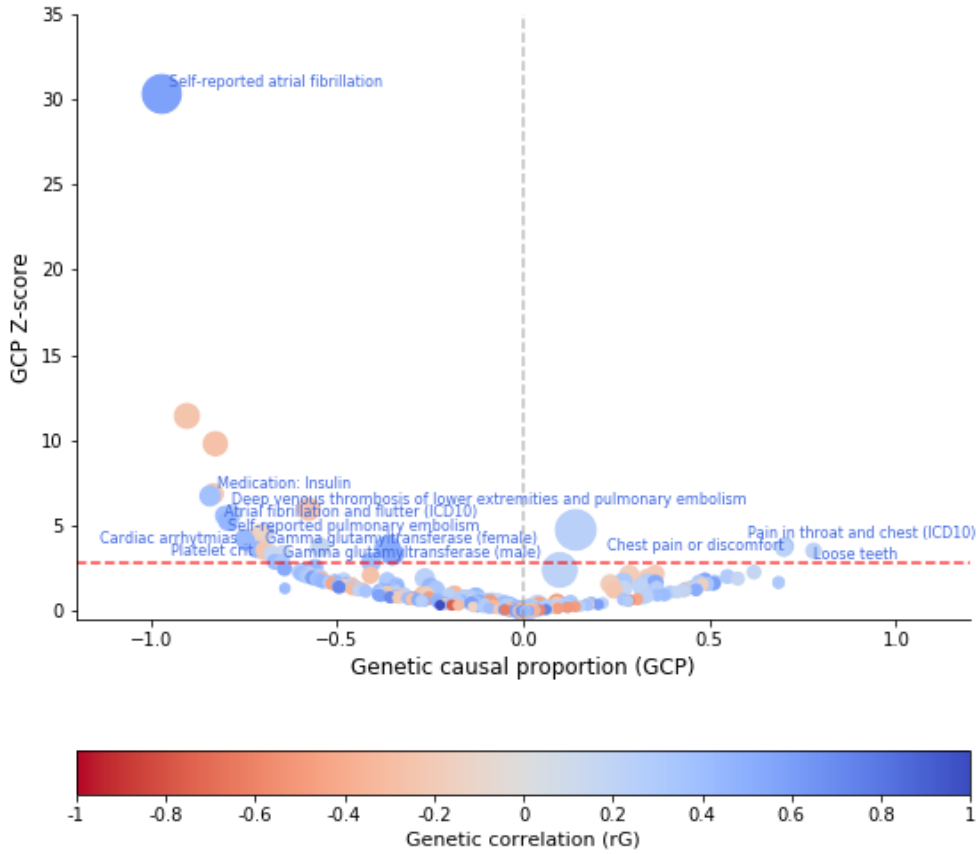

Fig.3. The diagram represents the traits that show positive genetic correlations with stroke. The dot point represents each trait with a significant genetic correlation with stroke. The x-axis shows the genetic causal proportion (GCP) values, and the y-axis shows GCP related Z-scores that are statistically significant. The horizontal red dashed line indicates the significant statistical threshold with a false discovery rate (FDR<.05). The vertical grey dash line indicates the putative causal traits that increase the risk of stroke (left side) and on the right side it shows the traits that are putative consequences of stroke.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

Table 1. Traits with potential causal association with stroke. This table represents the phenotypes with FDR<.05 and robust genetic causal proportions (|GCP| > 0.60) with stroke.

Trait	GCP	GCPse	GCPpval	rG	rGse	rGpval
Self-reported atrial fibrillation	-0.97	0.03	8.97×10^{-202}	0.58	0.18	9.51×10^{-4}
Self-reported pulmonary embolism	-0.79	0.15	1.37×10^{-7}	0.45	0.13	3.20×10^{-4}
Blood clot in the lung diagnosed by a doctor	-0.75	0.17	2.46×10^{-5}	0.44	0.13	3.41×10^{-4}
Cardiac arrhytmias	-0.72	0.19	2.99×10^{-4}	0.44	0.08	1.76×10^{-7}
Blood clot in the leg diagnosed by a doctor	-0.74	0.18	5.25×10^{-5}	0.43	0.11	1.96×10^{-4}
Medication: Gliclazide	-0.66	0.23	3.75×10^{-3}	0.42	0.09	2.33×10^{-5}
Self-reported deep venous thrombosis	-0.74	0.18	4.23×10^{-5}	0.40	0.11	2.60×10^{-4}
Atrial fibrillation and flutter (ICD10)	-0.80	0.14	2.10×10^{-8}	0.40	0.08	1.91×10^{-6}
Medication: Insulin	-0.84	0.12	1.43×10^{-11}	0.38	0.14	6.43×10^{-3}
Medication: Pioglitazone	-0.74	0.17	1.72×10^{-5}	0.33	0.12	8.45×10^{-3}
Weight change during worst episode of depression: Gained weight	-0.64	0.23	4.03×10^{-3}	0.32	0.12	1.01×10^{-2}
Deep venous thrombosis of lower extremities and pulmonary embolism	-0.78	0.14	1.28×10^{-8}	0.24	0.08	6.00×10^{-3}
Gamma glutamyltransferase (female)	-0.69	0.15	4.13×10^{-6}	0.16	0.04	3.23×10^{-4}
Platelet crit	-0.67	0.20	7.61×10^{-4}	0.14	0.04	1.32×10^{-3}
Gamma glutamyltransferase (male)	-0.64	0.19	7.61×10^{-4}	0.14	0.05	6.60×10^{-3}

This table represents the traits with significant false discovery rate (FDR < .05); GCP: genetic causal proportion; GCP se: genetic causal proportion standard deviation; GCP pval: genetic causal proportion unadjusted p-value; rG: genetic correlation; rG se: genetic correlation standard deviation; rG pval: genetic correlation unadjusted p-value.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

Table 2. The genetic variants of stroke influence the risk of certain traits.

Trait	GCP	GCPse	GCPpval	rG	rGse	rGpval
Pain in throat and chest (ICD10)	0.70	0.18	1.35×10^{-4}	0.35	0.07	4.66×10^{-7}
Chest pain or discomfort	0.69	0.18	2.19×10^{-4}	0.23	0.05	7.28×10^{-5}
Loose teeth	0.77	0.21	3.84×10^{-4}	0.21	0.08	1.08×10^{-2}

This table represents the traits with significant false discovery rate (FDR < .05); GCP: genetic causal proportion; GCP se: genetic causal proportion standard deviation; GCP pval: genetic causal proportion unadjusted p-value; rG: genetic correlation; rG se: genetic correlation standard deviation; rG pval: genetic correlation unadjusted p-value.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

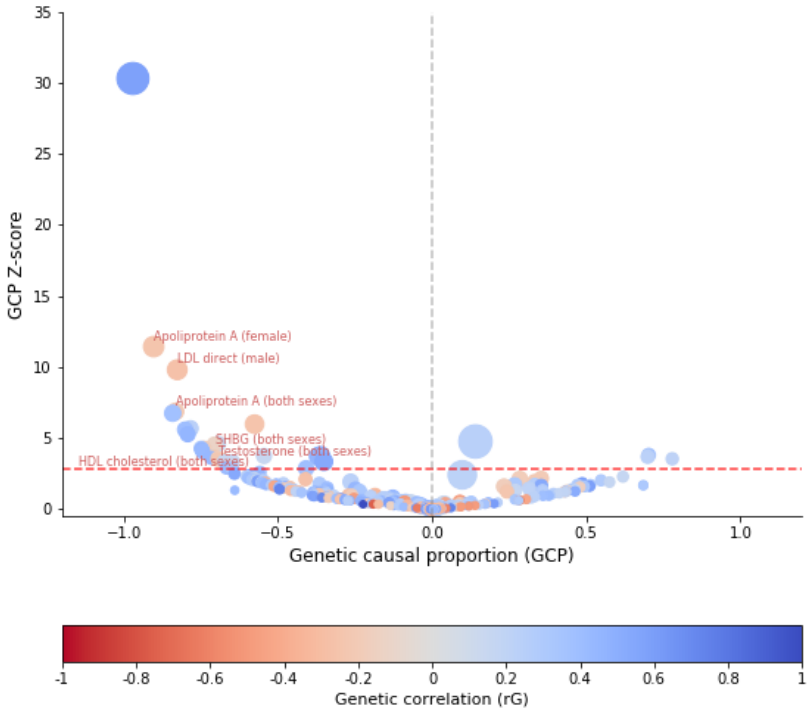

Fig.4. The diagram represents the traits that show negative genetic correlations with stroke. The dot point represents each trait with a significant genetic correlation with stroke. The x-axis shows the genetic causal proportion (GCP) values, and the y-axis shows GCP associated Z-scores that are statistically significant. The horizontal red dashed line indicates the significant statistical threshold with a false discovery rate (FDR<.05). The vertical grey dashed line indicates the traits that decrease the risk of stroke (left side).

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

Table 3. Traits with a negative genetic correlation and significant partial causality on stroke risk.

Name	GCP	GCPse	GCPpval	rG	rGse	rGpval
Testosterone (male)	-0.70	0.15	5.12×10^{-6}	-0.12	0.04	1.06×10^{-2}
SHBG (both sexes)	-0.70	0.19	2.11×10^{-4}	-0.14	0.04	1.44×10^{-3}
Testosterone (female)	-0.68	0.19	4.07×10^{-4}	-0.15	0.05	7.52×10^{-3}
SHBG (female)	-0.74	0.17	1.68×10^{-5}	-0.17	0.05	8.17×10^{-4}
Testosterone (both sexes)	-0.69	0.19	3.14×10^{-4}	-0.18	0.05	4.85×10^{-4}
Apoliprotein A (both sexes)	-0.83	0.12	5.77×10^{-12}	-0.20	0.06	2.0×10^{-3}
HDL cholesterol (both sexes)	-0.69	0.19	5×10^{-4}	-0.23	0.08	7.84×10^{-3}
Apoliprotein A (female)	-0.90	0.07	2.56×10^{-30}	-0.24	0.06	3.1×10^{-4}
LDL direct (male)	-0.82	0.08	9.74×10^{-23}	-0.27	0.07	3.76×10^{-4}
HDL cholesterol (female)	-0.70	0.19	2.9×10^{-4}	-0.27	0.07	3.93×10^{-4}

This table represents the traits with significant false discovery rate (FDR < .05); GCP: genetic causal proportion; GCP se: genetic causal proportion standard deviation; GCP pval: genetic causal proportion unadjusted p-value; rG: genetic correlation; rG se: genetic correlation standard deviation; rG pval: genetic correlation unadjusted p-value

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

Discussion

This study advances our understanding of stroke aetiology by shedding light on the causal architecture of stroke. We leveraged GWAS summary data to examine the potential causal association between stroke and 1,504 different complex traits. Based on genetic evidence, we discovered 28 potential causal relationships. Overall, our findings suggest that certain pathophysiological conditions and vascular disorders may increase the risk of stroke.

Our study revealed that the causal genetic effects of certain traits may influences the risk of stroke. Consistent with previous studies, our study showed that atrial fibrillation (AF) increases the risk of stroke. Recently, two sample MR studies found bi-directional relationships between AF and cardiovascular disease, cognitive impairment, type 2 diabetes mellitus and late-onset Alzheimer's disease, which suggests AF may have a bi-directional causal relationship with stroke and stroke subtypes (26). Previous studies reported that AF contributes to the pathogenesis of stroke by causing stasis in the left atrium and subsequent brain embolism (27). AF is the most prevalent kind of cardiac arrhythmia, and it is widely known as a significant risk factor for ischemic stroke, transient ischaemic attacks, heart attack, and dementia (28–30). A study showed that about 15–25% of people who suffered an ischemic stroke had AF (31–33), and it also enhanced stroke risk five-fold compared with healthy people (29). However, apart from triggering stroke, AF may be associated with other risk factors that may influence stroke. For instance, age, sex, coronary artery disease, high blood pressure, diabetes, inflammatory disorders, and cardiac embolism are all potential risk factors for both stroke and AF (27). Our findings suggest that a genetic predisposition to AF contributes to stroke risk. Venous thromboembolism (VET), also known as deep vein thrombosis (DVT) and pulmonary embolism (PE), may have a role in circulatory collapse and short- and long-term serious effects on quality of life (34). In particular, VTE is a complex disorder, with increasing age and obesity recognised as prevalent atherosclerotic risk factors for ischemic stroke (35). The incidence of DVT is higher in aged people with age $\geq$ 65 years (36,37). DVT happens when a thrombus (blood clot) forms in the deep veins, particularly in the lower extremities (legs), which leads to pain, cramping, and swelling in the lower extremities, and DVT is common among patients with stroke (36). However, DVT happens in 10% to 75% of post-stroke patients depending on the diagnosis method and time of evaluation (38). Our findings suggest that genetic variants of VTE and PE influence the risk of stroke. Weight change (weight gain) after depression is a serious public health concern, and depressed individuals have a 58% chance to develop obesity (39). Research conducted in 2021 revealed that patients who had a history of adolescent overweight and obesity as measured by BMI had a higher risk of developing an early-onset ischemic stroke (40). A study on obesity and cryptogenic ischemic stroke found that a high BMI was associated with a higher risk of ischemic stroke in young individuals when smoking, high blood pressure, and diabetes were taken into consideration (41). Our study found the genetic evidence of weight gain after depression may increase the risk of stroke.

Previous studies reported that vascular factors such as type 2 diabetes and hypertension may increase the risk of ischemic and intracerebral haemorrhagic stroke by 80–90% (42). MR studies revealed that genetically predicted type 2 diabetes was associated with ischemic heart disease, stroke, and cardiovascular diseases (43). In our study, cardiac arrhythmias and a blood clot in the leg diagnosed by a doctor increased stroke risk. Similarly, the use of medication for type 2 diabetes increases the risk of stroke. For instance, pioglitazone is used as a cardioprotective drug for type 2 diabetes patients. This drug is used as a proxy for type 2 diabetes. Pioglitazone is used to reduce myocardial infarction, stroke, and cardiovascular mortality by retarding the atherosclerotic process (44). Insulin,

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

for example, is used for anti-hyperglycaemic therapy (45), so it could be used as a proxy for type 2 diabetes. Insulin use was linked to type 2 diabetes, indicating a shared genetic architecture of type 2 diabetes and insulin. Therefore, in this case, type 2 diabetes would be the risk factor for stroke. In addition, gliclazide is a member of the sulphonylurea drug group commonly used in type 2 diabetes treatment to control blood glucose (46). A population-wide study in Korea observed that gliclazide monotherapy increases all-cause mortality and the risk of acute myocardial infarction and stroke (46). Consistently, in the present study, several cardiovascular phenotypes such as atrial fibrillation and flutter, self-reported pulmonary embolism, a blood clot in the leg diagnosed by a doctor, deep venous thrombosis of lower extremities, pulmonary embolism, and weight change during the worst episode of depression (gaining weight) are significantly genetically causally associated with stroke risk. These observations might be used to develop testable hypotheses for future studies.

Neck and chest pain are common consequences among the patients with stroke. Chest pain and discomfort occurred due to a lack of oxygen-rich blood in the heart muscle. It usually happens when one or more arteries in the heart become blocked or narrowed. Previous studies reported that angina pectoris (chest pain) enhances the risk of ischemic stroke (47), and patients with acute stroke may also develop different medical complications, among which chest pain is common (48). Our results found the inferred causal genetic relationship between stroke and chest pain. In addition, stroke survivors have a lower oral health-related quality of life (49). A recent study found that tooth loss was significantly associated with vascular cognitive impairment (VCI) that causes ischemic stroke (50,51). A study conducted on the dental health and cognitive function of 161 people who had acute ischemic strokes and had lower cognitive function test scores and higher levels of VCI tended to lose more teeth (52). We presume that the relationship between stroke and tooth loss could be mediated by potential vertical pleiotropic effects and may potentially be influenced by genetic predisposition and oral health susceptibility.

Our results demonstrated that certain hormones and enzymes play a significant role in stroke prevention. Apolipoprotein A1 (ApoA1) is a protein component of high-density lipoprotein (HDL) and is involved in the transfer of excessive cholesterol from peripheral cells to the liver. Aside from its atheroprotective properties, apoA1 also contains anti-inflammatory and antioxidant properties (53). A previous study found that low amounts of apolipoprotein A1 and high amounts of apolipoprotein B were associated with an increased risk of cerebrovascular stroke and cardiovascular diseases (53–56). High levels of high-density lipoprotein (HDL) cholesterol significantly reduce the risk of stroke (57). On the other hand, high levels of LDL (low-density lipoprotein) enhance the risk of ischemic stroke, and low levels of LDL are linked to the risk of intracerebral haemorrhagic stroke (58). Genome-wide association studies revealed that A1-SD genetically elevated LDL cholesterol enhanced the risk of ischemic stroke and large artery atherosclerosis stroke (59), while A1-SD genetically elevated HDL cholesterol decreased the risk of small vessel stroke (60). In addition, sex hormone-binding globulin (SHBG) is a protein produced by the liver and usually binds to sex hormones such as estrogen, dihydrotestosterone, and testosterone. Previous studies demonstrated that elevated levels of SHBG protein decrease the risk of diabetes, stroke, and vascular diseases, while low levels of SHBG increase the risk of those diseases (61,62). Thus, SHBG might be used as a risk prediction tool for predicting stroke risk.

We acknowledged several limitations in this present study. Firstly, the GWAS summary statistics used in this study were obtained from the MEGASTROKE and UK Biobank cohort, where the participants are all from European ancestry. Thus, our findings should not be consistent with other ancestries unless they are verified using

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

data from diverse populations. Secondly, our studies included 1504 traits, but other causal relationships may still exist with other traits. Thirdly, by definition, the LCV method evaluates the causal relationship between two traits and assumes no bidirectional causality (19). Consequently, it is still debatable whether the LCV method would be able to determine the bidirectional causality between the traits. This assumption must be systematically evaluated in the future. Null findings in this study might be explained in a way that there is no causal association between the phenotypes or there may exist bidirectional causality. In the current study, caution must be taken when interpreting null findings.

Conclusions

In summary, our study provides evidence for potential causal genetic relationships between stroke and other complex traits. Our findings suggest that the role of several cardiovascular traits, such as atrial fibrillation, cardiac arrhythmias increase the risk of stroke. We also observed that the influence of different physiological conditions, such as blood clots in the lung and legs and sudden weight gain after depression, may increase the risk of stroke. Overall, our findings support from the evidence of previous epidemiological studies, and provide novel insights into the causal genetic architecture of stroke, which in turn could be used as testable hypotheses to improve the development of future studies, treatments, and preventative strategies.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

References

1. Organization, W.H. (2010) World health statistics 2010. *World health statistics 2010*; World Health Organization, (2010) .

2. Roger, V.L., Go, A.S., Lloyd-Jones, D.M., Adams, R.J., Berry, J.D., Brown, T.M., Carnethon, M.R., Dai, S., De Simone, G. and Ford, E.S. (2011) Heart disease and stroke statistics—2011 update: a report from the American Heart Association. *Circulation*, **123**, e18–e209.

3. Global and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016, regional (2017) A systematic analysis for the global burden of disease Study 2016. *Lancet*, **390**, 1211–1259.

4. Kuriakose, D. and Xiao, Z. (2020) Pathophysiology and treatment of stroke: present status and future perspectives. *Int. J. Mol. Sci.*, **21**, 7609.

5. Bak, S., Gaist, D., Sindrup, S.H., Skytthe, A. and Christensen, K. (2002) Genetic liability in stroke: a long-term follow-up study of Danish twins. *Stroke*, **33**, 769–774.

6. Liao, D., Myers, R., Hunt, S., Shahar, E., Paton, C., Burke, G., Province, M. and Heiss, G. (1997) Familial history of stroke and stroke risk: the Family Heart Study. *Stroke*, **28**, 1908–1912.

7. Dichgans, M. (2007) Genetics of ischaemic stroke. *Lancet Neurol.*, **6**, 149–161.

8. Bevan, S., Traylor, M., Adib-Samii, P., Malik, R., Paul, N.L.M., Jackson, C., Farrall, M., Rothwell, P.M., Sudlow, C. and Dichgans, M. (2012) Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. *Stroke*, **43**, 3161–3167.

9. Gretarsdottir, S., Thorleifsson, G., Manolescu, A., Styrkarsdottir, U., Helgadottir, A., Gschwendtner, A., Kostulas, K., Kuhlenbäumer, G., Bevan, S. and Jonsdottir, T. (2008) Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann. Neurol.*, **64**, 402–409.

10. Gudbjartsson, D.F., Holm, H., Gretarsdottir, S., Thorleifsson, G., Walters, G.B., Thorgeirsson, G., Gulcher, J., Mathiesen, E.B., Njølstad, I. and Nyrnes, A. (2009) A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat. Genet.*, **41**, 876–878.

11. Gschwendtner, A., Bevan, S., Cole, J.W., Plourde, A., Matarin, M., Ross-Adams, H., Meitinger, T., Wichmann, E., Mitchell, B.D. and Furie, K. (2009) Sequence variants on chromosome 9p21. 3 confer risk for atherosclerotic stroke. *Ann. Neurol. Off. J. Am. Neurol. Assoc. Child Neurol. Soc.*, **65**, 531–539.

12. Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., Rutten-Jacobs, L., Giese, A.-K., Van Der Laan, S.W. and Gretarsdottir, S. (2018) Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet.*, **50**, 524–537.

13. Malik, R., Rannikmäe, K., Traylor, M., Georgakis, M.K., Sargurupremraj, M., Markus, H.S., Hopewell, J.C., Debette, S., Sudlow, C.L.M. and Dichgans, M. (2018) Genome-wide meta-analysis identifies 3 novel loci associated with stroke. *Ann. Neurol.*, **84**, 934–939.

14. Mishra, A., Malik, R., Hachiya, T., Jürgenson, T., Namba, S., Posner, D.C., Kamanu, F.K., Koido, M., Le Grand, Q. and Shi, M. (2022) Stroke genetics informs drug discovery and risk prediction across ancestries. *Nature*, **611**, 115–123.

15. Georgakis, M.K. and Gill, D. (2021) Mendelian randomization studies in stroke: exploration of risk factors and drug targets with human genetic data. *Stroke*, **52**, 2992–3003.

16. Liu, J., Rutten-Jacobs, L., Liu, M., Markus, H.S. and Traylor, M. (2018) Causal impact of type 2 diabetes mellitus on cerebral small vessel disease: a Mendelian randomization analysis. *Stroke*, **49**, 1325–1331.

17. García-Marín, L.M., Campos, A.I., Martin, N.G., Cuéllar-Partida, G. and Rentería, M.E. (2021) Phenome-wide analysis highlights putative causal relationships between self-reported migraine and other complex traits. *J. Headache Pain*, **22**, 1–8.

18. Siewert, K.M., Klarin, D., Damrauer, S.M., Chang, K.-M., Tsao, P.S., Assimes, T.L., Davey Smith, G. and Voight, B.F. (2020) Cross-trait analyses with migraine reveal widespread pleiotropy and suggest a

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

vascular component to migraine headache. *Int. J. Epidemiol.*, **49**, 1022–1031.

19. O'Connor, L.J. and Price, A.L. (2018) Distinguishing genetic correlation from causation across 52 diseases and complex traits. *Nat. Genet.*, **50**, 1728–1734.

20. Van Rheenen, W., Peyrot, W.J., Schork, A.J., Lee, S.H. and Wray, N.R. (2019) Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.*, **20**, 567–581.

21. Koellinger, P.D. and De Vlaming, R. (2019) Mendelian randomization: the challenge of unobserved environmental confounds. Mendelian randomization: the challenge of unobserved environmental confounds. *Int. J. Epidemiol.* **(2019)** , *48*, 665–671.

22. Cuellar-Partida, G., Lundberg, M., Kho, P.F., D'Urso, S., Gutiérrez-Mondragón, L.F., Ngo, T.T. and Hwang, L.-D. (2019) Complex-Traits Genetics Virtual Lab: A community-driven web platform for post-GWAS analyses. *BioRxiv*, 518027.

23. García-Marín, L.M., Campos, A.I., Martin, N.G., Cuéllar-Partida, G. and Rentería, M.E. (2021) Inference of causal relationships between sleep-related traits and 1,527 phenotypes using genetic data. *Sleep*, **44**, zsaa154.

24. Haworth, S., Kho, P.F., Holgerson, P.L., Hwang, L.-D., Timpson, N.J., Rentería, M.E., Johansson, I. and Cuellar-Partida, G. (2021) Assessment and visualization of phenome-wide causal relationships using genetic data: an application to dental caries and periodontitis. *Eur. J. Hum. Genet.*, **29**, 300–308.

25. Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L. and Neale, B.M. (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.*, **47**, 291–295.

26. Kwok, M.K. and Schooling, C.M. (2021) Mendelian randomization study on atrial fibrillation and cardiovascular disease subtypes. *Sci. Rep.*, **11**, 1–9.

27. Kamel, H., Okin, P.M., Elkind, M.S. V and Iadecola, C. (2016) Atrial fibrillation and mechanisms of stroke: time for a new model. *Stroke*, **47**, 895–900.

28. Chugh, S.S., Havmoeller, R., Narayanan, K., Singh, D., Rienstra, M., Benjamin, E.J., Gillum, R.F., Kim, Y.-H., McAnulty Jr, J.H. and Zheng, Z.-J. (2014) Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation*, **129**, 837–847.

29. Son, M.K., Lim, N.-K., Kim, H.W. and Park, H.-Y. (2017) Risk of ischemic stroke after atrial fibrillation diagnosis: A national sample cohort. *PLoS One*, **12**, e0179687.

30. Hou, L., Xu, M., Yu, Y., Sun, X., Liu, X., Liu, L., Li, Y., Yuan, T., Li, W. and Li, H. (2020) Exploring the causal pathway from ischemic stroke to atrial fibrillation: a network Mendelian randomization study. *Mol. Med.*, **26**, 1–9.

31. Arboix, A., García-Eroles, L., Massons, J.B., Oliveres, M., Pujades, R. and Targa, C. (2000) Atrial fibrillation and stroke: clinical presentation of cardioembolic versus atherothrombotic infarction. *Int. J. Cardiol.*, **73**, 33–42.

32. Paciaroni, M., Agnelli, G., Caso, V., Venti, M., Milia, P., Silvestrelli, G., Parnetti, L. and Biagini, S. (2005) Atrial fibrillation in patients with first-ever stroke: frequency, antithrombotic treatment before the event and effect on clinical outcome. *J. Thromb. Haemost.*, **3**, 1218–1223.

33. Lamassa, M., Di Carlo, A., Pracucci, G., Basile, A.M., Trefoloni, G., Vanni, P., Spolveri, S., Baruffi, M.C., Landini, G. and Ghetti, A. (2001) Characteristics, outcome, and care of stroke associated with atrial fibrillation in Europe: data from a multicenter multinational hospital–based registry (The European Community Stroke Project). *Stroke*, **32**, 392–398.

34. Heit, J.A. (2005) Venous thromboembolism: disease burden, outcomes and risk factors. *J. Thromb. Haemost.*, **3**, 1611–1617.

35. Glynn, R.J. and Rosner, B. (2005) Comparison of risk factors for the competing risks of coronary heart disease, stroke, and venous thromboembolism. *Am. J. Epidemiol.*, **162**, 975–982.

36. Liu, Z., Liu, D., Guo, Z.-N., Jin, H., Sun, T., Ni, C. and Yan, X. (2021) Incidence and Risk Factors of Lower-Extremity Deep Vein Thrombosis After Thrombolysis Among Patients with Acute Ischemic

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

Stroke. *Pharmgenomics. Pers. Med.*, **14**, 1107.

37. Liu, L., Zheng, H., Wang, D.Z., Wang, Y., Hussain, M., Sun, H., Wang, A., Zhao, X., Dong, K. and Wang, C. (2014) Risk assessment of deep-vein thrombosis after acute stroke: a prospective study using clinical factors. *CNS Neurosci. Ther.*, **20**, 403–410.

38. Khan, M.T., Ikram, A., Saeed, O., Afridi, T., Sila, C.A., Smith, M.S., Irshad, K. and Shuaib, A. (2017) Deep vein thrombosis in acute stroke-a systemic review of the literature. *Cureus*, **9**.

39. Blasco, B.V., García-Jiménez, J., Bodoano, I. and Gutiérrez-Rojas, L. (2020) Obesity and depression: Its prevalence and influence as a prognostic factor: A systematic review. *Psychiatry Investig.*, **17**, 715.

40. Bardugo, A., Fishman, B., Libruder, C., Tanne, D., Ram, A., Hershkovitz, Y., Zucker, I., Furer, A., Gilon, R. and Chodick, G. (2021) Body mass index in 1.9 million adolescents and stroke in young adulthood. *Stroke*, **52**, 2043–2052.

41. Jaakonmäki, N., Zedde, M., Sarkanen, T., Martinez-Majander, N., Tuohinen, S., Sinisalo, J., Ryödi, E., Autere, J., Hedman, M. and Junttola, U. (2022) Obesity and the risk of cryptogenic ischemic stroke in young adults. *J. Stroke Cerebrovasc. Dis.*, **31**, 106380.

42. Hauer, A.J., Ruigrok, Y.M., Algra, A., van Dijk, E.J., Koudstaal, P.J., Luijckx, G., Nederkoorn, P.J., van Oostenbrugge, R.J., Visser, M.C. and Wermer, M.J. (2017) Age-Specific vascular risk factor profiles according to stroke subtype. *J. Am. Heart Assoc.*, **6**, e005090.

43. Huang, M., Laina-Nicaise, L.-D., Zha, L., Tang, T. and Cheng, X. (2022) Causal Association of Type 2 diabetes mellitus and glycemic traits with cardiovascular diseases and lipid traits: a Mendelian randomization study. *Front. Endocrinol. (Lausanne).*, **13**.

44. DeFronzo, R.A., Inzucchi, S., Abdul-Ghani, M. and Nissen, S.E. (2019) Pioglitazone: the forgotten, cost-effective cardioprotective drug for type 2 diabetes. *Diabetes Vasc. Dis. Res.*, **16**, 133–143.

45. Lau, A.N.C., Tang, T., Halapy, H., Thorpe, K. and Catherine, H.Y. (2012) Initiating insulin in patients with type 2 diabetes. *Cmaj*, **184**, 767–776.

46. Cho, E.H., Han, K., Kim, B. and Lee, D.H. (2020) Gliclazide monotherapy increases risks of all-cause mortality and has similar risk of acute myocardial infarction and stroke with glimepiride monotherapy in Korean type 2 diabetes mellitus. *Medicine (Baltimore).*, **99**, e21236.

47. Tanne, D., Shotan, A., Goldbourt, U., Haim, M., Boyko, V., Adler, Y., Mandelzweig, L. and Behar, S. (2002) Severity of angina pectoris and risk of ischemic stroke. *Stroke*, **33**, 245–250.

48. Kumar, S., Selim, M.H. and Caplan, L.R. (2010) Medical complications after stroke. *Lancet Neurol.*, **9**, 105–118.

49. Schimmel, M., Leemann, B., Christou, P., Kiliaridis, S., Schnider, A., Herrmann, F.R. and Müller, F. (2011) Oral health-related quality of life in hospitalised stroke patients. *Gerodontology*, **28**, 3–11.

50. Grabe, H.J., Schwahn, C., Völzke, H., Spitzer, C., Freyberger, H.J., John, U., Mundt, T., Biffar, R. and Kocher, T. (2009) Tooth loss and cognitive impairment. *J. Clin. Periodontol.*, **36**, 550–557.

51. Matthews, J.C., You, Z., Wadley, V.G., Cushman, M. and Howard, G. (2011) The association between self-reported tooth loss and cognitive function in the REasons for Geographic And Racial Differences in Stroke study: an assessment of potential pathways. *J. Am. Dent. Assoc.*, **142**, 379–390.

52. Zhu, J., Li, X., Zhu, F., Chen, L., Zhang, C., McGrath, C., He, F., Xiao, Y. and Jin, L. (2015) Multiple tooth loss is associated with vascular cognitive impairment in subjects with acute ischemic stroke. *J. Periodontal Res.*, **50**, 683–688.

53. Eldeeb, M.A., Zaki, A.S., Ashour, S., Abdel Nasser, A., El Bassiouny, A. and Abdulghani, K.O. (2020) Serum apolipoprotein A1: a predictor and prognostic biomarker in acute ischemic stroke. *Egypt. J. Neurol. Psychiatry Neurosurg.*, **56**, 1–5.

54. Lamarche, B., Moorjani, S., Lupien, P.J., Cantin, B., Bernard, P.-M., Dagenais, G.R. and Despre´ s, J.-P. (1996) Apolipoprotein AI and B levels and the risk of ischemic heart disease during a five-year follow-up of men in the Que´ bec cardiovascular study. *Circulation*, **94**, 273–278.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

55. Walldius, G., Jungner, I., Holme, I., Aastveit, A.H., Kolar, W. and Steiner, E. (2001) High apolipoprotein B, low apolipoprotein AI, and improvement in the prediction of fatal myocardial infarction (AMORIS study): a prospective study. *Lancet*, **358**, 2026–2033.

56. Bhatia, M., Howard, S.C., Clark, T.G., Neale, R., Qizilbash, N., Murphy, M.F.G. and Rothwell, P.M. (2006) Apolipoproteins as predictors of ischaemic stroke in patients with a previous transient ischaemic attack. *Cerebrovasc. Dis.*, **21**, 323–328.

57. Wannamethee, S.G., Shaper, A.G. and Ebrahim, S. (2000) HDL-cholesterol, total cholesterol, and the risk of stroke in middle-aged British men. *Stroke*, **31**, 1882–1888.

58. Collaboration, P.S. (2007) Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55 000 vascular deaths. *Lancet*, **370**, 1829–1839.

59. Hindy, G., Engström, G., Larsson, S.C., Traylor, M., Markus, H.S., Melander, O. and Orho-Melander, M. (2018) Role of blood lipids in the development of ischemic stroke and its subtypes: a Mendelian randomization study. *Stroke*, **49**, 820–827.

60. Baigent, C., Blackwell, L., Emberson, J., Holland, L.E., Reith, C., Bhala, N., Peto, R., Barnes, E.H., Keech, A. and Simes, J. (2010) Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet (London, England)*, **376**, 1670–1681.

61. Madsen, T.E., Luo, X., Huang, M., Park, K.E., Stefanick, M.L., Manson, J.E. and Liu, S. (2020) Circulating SHBG (sex hormone-binding globulin) and risk of ischemic stroke: findings from the WHI. *Stroke*, **51**, 1257–1264.

62. Ding, E.L., Song, Y., Manson, J.E., Hunter, D.J., Lee, C.C., Rifai, N., Buring, J.E., Gaziano, J.M. and Liu, S. (2009) Sex hormone–binding globulin and risk of type 2 diabetes in women and men. *N. Engl. J. Med.*, **361**, 1152–1163.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.08.23287006; this version posted March 12, 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.
All rights reserved. No reuse allowed without permission.

Data Availability Statement

All data here analysed are publicly available on MEGASTROKE consortium and UK biobank summary data available on CTG-VL.

Acknowledgments
"The MEGASTROKE project received funding from sources specified at http://www.megastroke.org/acknowledgments.html". The author list of MEGASTROKE is available in Supplementary Table 4. T.I. is supported by a Research Training scholarship from The University of Queensland, Australia.

Conflict of Interest: "Authors declare no conflict of interest".

Author contributions: "Tania Islam and Mohammad Ali Moni contributed to the study conception and design. Data collection and analysis were performed by Tania Islam, Luis M García-Marín. The first draft of the manuscript was written by Tania Islam and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript."

Table Legends:

Table 1. Traits with potential causal association with stroke. This table represents the phenotypes with FDR<.05 and robust genetic causal proportions (|GCP| > 0.60) with stroke.

Table 2: The genetic variants of stroke influence the risk certain traits.

Table 3. Traits with a negative genetic correlation and significant partial causality on stroke risk.

Figure Legend:

Fig.1. The concept of horizontal and vertical pleiotropic effects on phenotypes. (A) when genetic variants simultaneously affect trait A and trait B then it's called horizontal pleiotropy. (B) when a genetic variant affect trait A and trait A influence the risk of trait B then it's called vertical pleiotropy. rG represents genetic correlation.

Fig.2. The schematic diagram illustrates the genetic relationship between stroke and 1504 complex traits. The genetic correlation was consistent across the phenotypes, demonstrating that genetic variants specific to each of the 1504 complex traits had consistent positive or negative correlations with the genetic variants associated with stroke.

Fig.3. The diagram represents the traits that shows positive genetic correlation with stroke. The dot point represents each trait with a significant genetic correlation with stroke. The x-axis shows the genetic causal proportion (GCP) values and y-axis shows GCP related Z-scores that is statistically significant. The horizontal red dash line indicates the significant statistical threshold with false discovery rate (FDR<.05). The vertical grey dash line indicates the putative causal traits that increase the risk of stroke (left side) and on the right side shows the traits that are putative consequences of stroke.

Fig.4. The diagram represents the traits that shows negative genetic correlation with stroke. The dot point represents each trait with a significant genetic correlation with stroke. The x-axis shows the genetic causal proportion (GCP) values and y-axis shows GCP related Z-scores that is statistically significant. The horizontal red dash line indicates the significant statistical threshold with false discovery rate (FDR<.05). The vertical grey dash line indicates the traits that decreases the risk of stroke (left side).