Diabetes mellitus, glycemic traits, and cerebrovascular disease: a Mendelian randomization study

Author(s):
Marios K Georgakis, MD, PhD; Eric L Harshfield, PhD; Rainer Malik, PhD; Nora Franceschini, MD, MPH; Claudia Langenberg, MD, PhD; Nicholas Wareham, MD, PhD; Hugh S Markus, DM, F Med Sci; Martin Dichgans, MD

Equal Author Contributions:
Marios K. Georgakis and Eric L. Harshfield contributed equally as first authors to this work. Hugh S. Markus and Martin Dichgans contributed equally as last authors to this work.

Corresponding Author:
Martin Dichgans
martin.dichgans@med.uni-muenchen.de

Affiliation Information for All Authors: Marios K. Georgakis, MD, PhD, Institute for Stroke and Dementia Research, University Hospital, Ludwig-Maximilians-University, Munich, Germany, Graduate School for Systemic Neurosciences, Ludwig-Maximilians-University, Munich, Germany; Eric L. Harshfield, PhD, Stroke Research Group, Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK; Rainer Malik, PhD, Institute for Stroke and Dementia Research, University Hospital, Ludwig-Maximilians-University, Munich, Germany; Nora Franceschini, MD, MPH, Department of Epidemiology, UNC Gillings Global School of Public Health, Chapel Hill, NC, USA; Claudia Langenberg, MD, PhD5, MRC Epidemiology Unit, University of Cambridge, Cambridge, UK; Nicholas J. Wareham, MD, PhD, MRC Epidemiology Unit, University of Cambridge, Cambridge, UK; Hugh S. Markus, DM, F Med Sci, Stroke Research Group, Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK; Martin Dichgans, MD, Institute for Stroke and Dementia Research, University Hospital, Ludwig-Maximilians-University, Munich, Germany; Munich Cluster for Systems Neurology (SyNergy), Munich, Germany; German Centre for Neurodegenerative Diseases (DZNE), Munich, Germany

Contributions:
Marios K Georgakis: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data
Eric L Harshfield: Drafting/revision of the manuscript for content, including medical writing for content; Study concept or design; Analysis or interpretation of data
Rainer Malik: Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data
Nora Franceschini: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data
Claudia Langenberg: Drafting/revision of the manuscript for content, including medical writing for content; Study concept or design; Analysis or interpretation of data
Nicholas Wareham: Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data
Hugh S Markus: Drafting/revision of the manuscript for content, including medical writing for content; Study concept or design; Analysis or interpretation of data
Martin Dichgans: Drafting/revision of the manuscript for content, including medical writing for content; Study concept or design; Analysis or interpretation of data

Publication History: This manuscript was pre-published in Preprint version on medRχiv platform (doi: https://doi.org/10.1101/2019.12.27.19015834, https://www.medrxiv.org/content/10.1101/2019.12.27.19015834v1)

Number of characters in title: 96

Abstract Word count: 242

Word count of main text: 4649

References: 47

Figures: 2

Tables: 2

Supplemental: STROBE-MR checklist

Statistical Analysis performed by: Marios K. Georgakis, MD, PhD Rainer Malik, PhD Aff.: Institute for Stroke and Dementia Research, University Hospital, Ludwig-Maximilians-University (LMU)

Search Terms: [2] All Cerebrovascular disease/Stroke, [12] Stroke prevention, [59] Risk factors in epidemiology, [95] Association studies in genetics

Study Funding: M. Georgakis has received funding from the Onassis Foundation and the German Academic Exchange Service (DAAD). This project has received funding from the European Unions Horizon 2020 research and innovation programme (No 666881), SVDs@target (to M. Dichgans) and No 667375, CoSTREAM (to M. Dichgans and H. S. Markus); the DFG as part of the Munich Cluster for Systems Neurology (EXC 1010 SyNergy ID 390857198) and the CRC 1123 (B3) (to M. Dichgans); the Corona Foundation (to M. Dichgans); the Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis of Small Vessel Disease of the Brain) (to M. Dichgans); a grant for strategic collaboration between LMU Munich and Cambridge University; British Heart Foundation Programme Grant RG/16/4/32218 (To H. Markus); infrastrsuctural support from the Cambridge Univeriisty Hospitals NIHR Comprehensive Biomedical Research Centre.

Disclosures: Dr. Georgakis reports no disclosures relevant to the manuscript; Dr. Harshfield reports no disclosures relevant to the manuscript; Dr. Malik reports no disclosures relevant to the manuscript; Dr. Franceschini reports no disclosures relevant to the manuscript; Dr. Langenberg reports no disclosures relevant to the manuscript; Dr. Wareham reports no disclosures relevant to the manuscript; Dr. Markus reports no disclosures relevant to the manuscript; Dr. Dichgans reports no disclosuresrelevant to the manuscript;

Diabetes mellitus, glycemic traits, and cerebrovascular disease: a Mendelian randomization study

Marios K. Georgakis, MD, PhD[1,2] *, Eric L. Harshfield, PhD[3] *, Rainer Malik, PhD[1], Nora Franceschini, MD, MPH[4], Claudia Langenberg, MD, PhD[5], Nicholas J. Wareham, MD, PhD[5], Hugh S. Markus, DM, F Med Sci[3] †, Martin Dichgans, MD[1,6,7] †

1 Institute for Stroke and Dementia Research, University Hospital, Ludwig-Maximilians-University, Munich, Germany
2 Graduate School for Systemic Neurosciences, Ludwig-Maximilians-University, Munich, Germany
3 Stroke Research Group, Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK
4 Department of Epidemiology, UNC Gillings Global School of Public Health, Chapel Hill, NC, USA
5 MRC Epidemiology Unit, University of Cambridge, Cambridge, UK
6 Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
7 German Centre for Neurodegenerative Diseases (DZNE), Munich, Germany

* these authors contributed equally to the manuscript
† these authors contributed equally to the manuscript

Corresponding Author:
Martin Dichgans, MD
Institute for Stroke and Dementia Research, University Hospital
Ludwig-Maximilians-University (LMU)
Feodor-Lynen-Str. 17, 81377 Munich, Germany
Phone: +49-89-4400-46018; Fax: +49-89-4400-46040
E-mail: martin.dichgans@med.uni-muenchen.de

Statistical analyses were conducted by Marios K. Georgakis, MD, PhD, and Rainer Malik, PhD, Institute for Stroke and Dementia Research, University Hospital, Ludwig-Maximilians-University (LMU).

Supplemental Data: Online Supplement includes 11 Tables (e-1 to e-11) and 2 Figures (e-1 to e-2).

Word count:
Title character count: 96
Word count abstract: 242
Word count article: 4,649
Number of references: 47
Number of tables: 2
Number of figures: 2

Search terms: [2] All Cerebrovascular disease/Stroke, [7] Intracerebral hemorrhage, [91] All Genetics, [59] Risk factors in epidemiology

Study funding: M. Georgakis has received funding from the Onassis Foundation and the German Academic Exchange Service (DAAD). This project has received funding from the European Union's Horizon 2020 research and innovation programme (No 666881), SVDs@target (to M. Dichgans) and No 667375, CoSTREAM (to M. Dichgans and H. S. Markus); the DFG as part of the Munich Cluster for Systems Neurology (EXC 1010 SyNergy – ID 390857198) and the CRC 1123 (B3) (to M. Dichgans); the Corona Foundation (to M. Dichgans); the Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis of Small Vessel Disease of the Brain) (to M. Dichgans); a grant for strategic collaboration between LMU Munich and Cambridge University; British Heart Foundation Programme Grant RG/16/4/32218 (To H. Markus); infrastructural support from the Cambridge Univeristy Hospitals NIHR Comprehensive Biomedical Research Centre.

Role of the funder/sponsor: The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Acknowledgements: This research has been conducted using the UK Biobank Resource (UK Biobank application 2532). We acknowledge the contributions by the DIAGRAM Consortium, the MAGIC Consortium, the MEGASTROKE Consortium, the ISGC Consortium, and the CHARGE Consortium for making their data publicly available. MEGASTROKE has received funding from the sources detailed at http://www.megastroke.org/acknowledgments.html.

Author Disclosures:

Dr. Georgakis reports no disclosures.
Dr. Harshfield reports no disclosures.
Dr. Malik reports no disclosures.
Dr. Franceschini reports no disclosures.
Dr. Langenberg reports no disclosures.
Dr. Wareham reports no disclosures.
Dr. Markus reports no disclosures.
Dr. Dichgans reports no disclosures.

Publication history: Preprint version on medRχiv platform (doi: https://doi.org/10.1101/2019.12.27.19015834, https://www.medrxiv.org/content/10.1101/2019.12.27.19015834v1)

ABSTRACT

Objective: We employed Mendelian randomization (MR) to explore the effects of genetic predisposition to type 2 diabetes (T2D), hyperglycemia, insulin resistance, and β-cell dysfunction on risk of stroke subtypes and related cerebrovascular phenotypes.

Methods: We selected instruments for genetic predisposition to T2D (74,124 cases, 824,006 controls), HbA1c levels (n=421,923), fasting glucose levels (n=133,010), insulin resistance (n=108,557), and β-cell dysfunction (n=16,378) based on published genome-wide association studies. Applying two-sample MR, we examined associations with ischemic stroke (60,341 cases, 454,450 controls), intracerebral hemorrhage (1,545 cases, 1,481 controls), and ischemic stroke subtypes (large artery, cardioembolic, small vessel stroke), as well as with related phenotypes (carotid atherosclerosis, imaging markers of cerebral white matter integrity, and brain atrophy).

Results: Genetic predisposition to T2D and higher HbA1c levels were associated with higher risk of any ischemic stroke, large artery stroke, and small vessel stroke. Similar associations were also noted for carotid atherosclerotic plaque, fractional anisotropy, a white matter disease marker, and markers of brain atrophy. We further found associations of genetic predisposition to insulin resistance with large artery and small vessel stroke, whereas predisposition to β-cell dysfunction was associated with small vessel stroke, intracerebral hemorrhage, lower grey matter volume, and total brain volume.

Conclusions: This study supports causal effects of T2D and hyperglycemia on large artery and small vessel stroke. We show associations of genetically predicted insulin resistance and β-cell dysfunction with large artery and small vessel stroke that might have implications for anti-diabetic treatments targeting these mechanisms.

Classification of Evidence: This study provides Class II evidence that genetic predisposition to T2D and higher HbA1c levels are associated with a higher risk of large artery and small vessel ischemic stroke.

INTRODUCTION

Cerebrovascular disease is a major public health issue ranking as the second leading cause of mortality and adult disability worldwide [1,2]. Type 2 diabetes (T2D) is an established risk factor for cerebrovascular disease [3,4]. In cohort studies, T2D shows associations with higher risk for both ischemic and hemorrhagic stroke independently of other risk factors [5]. Also, several studies found associations of measures of hyperglycemia (glycated hemoglobin (HbA1c) and fasting glucose levels) with risk of stroke, both in patients with and without diabetes [5]. However, large-scale randomized controlled trials (RCTs) testing intensive glucose-lowering in patients with T2D show no significant reductions in risk of stroke, possibly due to insufficient power [6-8]. Moreover, the effects of T2D or hyperglycemia on etiological stroke subtypes (large artery stroke, cardioembolic stroke, small vessel stroke, intracerebral hemorrhage) remain elusive.

Currently available anti-diabetic medications act by either directly lowering glucose levels or by targeting two major mechanisms that contribute to hyperglycemia: insulin resistance or pancreatic β-cell dysfunction [9]. Observational data suggest that markers of insulin resistance, β-cell dysfunction, and hyperglycemia influence the risk of cardiovascular disease independently of each other [10,11]. However, data on stroke and its etiological subtypes are lacking. Moreover, there is a risk of confounding and reverse causation in observational studies. Developing targeted strategies for stroke prevention in patients at risk or suffering from T2D would require disentangling these relationships.

Mendelian randomization (MR) may help to clarify these associations. MR uses genetic variants as instruments for traits of interest and is not prone to confounding and reverse causation [12]. As such, MR has been proven a powerful methodology for inferring causality [13,14]. The availability of large-scale genome-wide association studies (GWAS) with detailed

phenotyping of cases further enables the exploration of etiological stroke subtypes that are typically not considered in observational studies.

Here, we leveraged large-scale data from GWASs and performed MR analyses, with the following aims: (i) to examine the effects of genetic predisposition to T2D on risk of ischemic stroke, ischemic stroke subtypes, and intracerebral hemorrhage; (ii) to explore the effects of genetically predicted measures of hyperglycemia (HbA1c and fasting glucose levels) on these phenotypes; (iii) to examine the associations of genetic predisposition to insulin resistance and β-cell dysfunction with major stroke etiologies; and (iv) to explore associations between diabetic traits and related vascular phenotypes including carotid atherosclerosis, neuroimaging markers of white mater integrity, and brain atrophy.

METHODS

Study design and data sources

This is a two-sample MR study following the guidelines for strengthening the reporting of Mendelian randomization studies (STROBE-MR) [15]. The study is based on publicly available summary statistics from GWAS consortia. Data sources are detailed in **Table 1**. MR uses genetic variants associated with exposures of interest and then explores the associations between the genetic predisposition to this exposure or the genetically predicted levels of the exposure phenotype with disease outcomes. As the genetic predisposition to a trait of interest is not affected by potential confounders, this approach is considered to be less prone to confounding, as compared to traditional observational analyses.

Our study design is depicted in **Figure e-1** and a detailed description of the phenotypes explored as exposures is provided in **Supplemental Table e-1**. We explored associations of genetic predisposition to T2D, measures of hyperglycemia (HbA1c and fasting glucose levels),

as well as markers of insulin resistance and β-cell dysfunction with cerebrovascular disease phenotypes including stroke subtypes, carotid atherosclerosis, white matter (WM) integrity, and brain atrophy. Information on genetic variants used as instruments are presented in **Supplemental Tables e-2 to e-7.**

Genetic instrument selection

Diabetes mellitus type 2. We selected genetic instruments from the latest GWAS meta-analysis for T2D based on 74,124 cases and 824,006 controls of European ancestry from 32 studies included in the DIAGRAM consortium [16]. The analyses were adjusted for age, sex, and population structure. There were 403 distinct genetic variants showing significant associations with T2D in this meta-analysis. We clumped these variants for linkage disequilibrium based on a distance window of 10,000 kB and an $r^2<0.01$ and used the remaining 289 variants as instruments (**Table e-2**). Given the average LD block length of 22,000 kB,[17] we used a 10,000 kB clumping window, with the notice that we cannot rule out very long-range LD effects.

Hyperglycemia. We selected genetic instruments for HbA1c levels (per 1%-increment) based on two different GWASs that we performed on individuals of White British ancestry in the UK Biobank (UKB) [18]. In the primary analysis, we explored HbA1c levels across the entire range of its values among both diabetic and non-diabetic individuals (n= 421,923). In this analysis, we only excluded individuals on anti-diabetic medications or insulin at the start of the study (n=5,468), as these medications affect HbA1c levels beyond genetic influence. In a secondary analysis, we explored HbA1c levels in the pre-diabetic range among diabetes-free individuals. In this analysis, we excluded individuals with self-reported history of physician-diagnosed diabetes, use of oral antidiabetic drugs or insulin, HbA1c level >6.5%, or random glucose levels >200 mg/dl (n=400,989). In both analyses, we also excluded 17,534 individuals that were included in the GWAS analysis for imaging phenotypes (see below) to avoid population

overlap between exposure and outcome datasets. We adjusted for age, sex, genotyping platform array, assessment center, and the first 20 principal components of the population structure and performed the analyses using BOLT-LMM with correction for relatedness and subtle population stratification. For fasting glucose levels (per 1-SD increment), we used the most recent GWAS meta-analysis (adjusted for age, sex, and population structure) by the MAGIC consortium on 133,010 diabetes-free individuals of European ancestry [19]. For both HbA1c and fasting glucose, we selected as instruments genetic variants reaching genome-wide significance ($p<5x10^{-8}$) after clumping at an $r^2<0.01$ threshold (clumping window 10,000 kB). We identified 333 instruments for HbA1c among both diabetic and non-diabetic individuals, 543 instruments for HbA1c levels among diabetes-free individuals, and 21 for fasting glucose levels among diabetes-free individuals (**Tables e-3** to **e-5**).

As several variants may influence HbA1c levels through effects on erythrocyte biology and not by inducing hyperglycemia [20], to isolate the effects of the hyperglycemia-related genetic component of HbA1c levels, we performed sensitivity analyses excluding those variants reported to be associated at $p<0.001$ with erythrocyte-related traits (hemoglobin concentration, red blood cell count, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin concentration, mean corpuscular hemoglobin, red cell distribution width, reticulocyte count, reticulocyte fraction of red cells, immature fraction of reticulocytes, high light scatter percentage of red cells, high light scatter reticulocyte count) in Phenoscanner [21].

Insulin resistance and β-cell dysfunction. As instruments for insulin resistance we used 53 genetic variants identified in a multi-trait GWAS to associate with the three components of this phenotype (fasting insulin levels, triglycerides and HDL-cholesterol; **Table e-6**) [22]. All three GWASs that were used to perform the multi-trait GWAS were based exclusively on European individuals. We weighted the instruments based on their effects on fasting insulin levels (per 1-log increment) in a GWAS meta-analysis of 108,557 diabetes-free European individuals [19]. In

accordance with existing literature, we proxied β-cell dysfunction based on fasting proinsulin levels (per 1 log-increment) [23,24]. We used summary statistics from a GWAS meta-analysis of 16,378 diabetes-free European individuals and identified 21 genetic instruments (at $p<5x10^{-8}$, $r^2<0.01$; clumping window 10,000 kB; **Table e-7**) [23]. The GWAS for fasting insulin levels was adjusted for age, sex, and population structure [19], whereas the GWAS for pro-insulin was additionally adjusted for fasting insulin levels [23].

We further used T2D-associated genetic variants previously grouped into clusters of diabetic endophenotypes; three clusters of insulin resistance (related to obesity, fat distribution, or lipid metabolism) and two clusters of β-cell dysfunction both associated with reduced levels of fasting insulin, but with opposing effects on fasting proinsulin [25]. We used the clusters of the variants and the respective weights per variant and cluster, as described by Udler *et al.* (**Table e-8)** [25].

Proportion of explained variance

For all genetic variants used as instruments, we estimated the proportion of explained variance for the respective phenotypes (**Tables e-2 to e-7**). We estimated the variance explained by each genetic variant for T2D based on the method by So *et al.* for binary phenotypes [26] and for the continuous traits we used a previously described formula based on summary statistics [27]. For the estimations regarding T2D, we used a prevalence rate of 8.5%, according to the 2015 estimate of the global prevalence of the disease by the International Diabetes Federation [28].

Associations with outcomes

We then examined associations of the selected instruments with ischemic stroke, ischemic stroke subtypes, and intracerebral hemorrhage (ICH) as the primary outcomes of interest. For

ischemic stroke, we used summary GWAS data from MEGASTROKE, mainly consisting of European individuals (70%) [29,30]. We extracted summary GWAS statistics for any ischemic stroke (60,341 cases, 451,210 controls) and for the major ischemic stroke subtypes: large artery stroke (6,688 cases, 238,513 controls), cardioembolic stroke (9,006 cases, 352,852 controls), and small vessel stroke (11,710 cases, 287,067 controls). The major ischemic stroke subtypes in MEGASTROKE were defined according to the TOAST criteria [31]. In sensitivity analyses, we also restricted our analyses to solely individuals of European ancestry. GWAS data for ICH were derived from the International Stroke Genetics Consortium (ISGC) GWAS meta-analysis including 1,545 cases and 1,481 controls of European ancestry [32].

Presence of carotid plaque, markers of WM tract integrity (WM hyperintensities (WMH) volume, mean diffusivity, fractional anisotropy), and markers of brain atrophy (grey matter volume, total brain volume) were explored as secondary outcomes. Carotid plaque data were derived from a GWAS meta-analysis (21,540 cases, 26,894 controls of European ancestry) from the CHARGE consortium.[33] As detailed in this meta-analysis, carotid plaques across the individual studies was defined by atherosclerotic thickening of the common carotid artery wall or the proxy measure of luminal stenosis greater than 25% [33]. For the imaging phenotypes (WMH volume, mean diffusivity, fractional anisotropy, grey matter volume, total brain volume), we undertook GWAS analyses in the UK Biobank neuroimaging dataset including 17,534 individuals of White British ancestry based on the MRI sequences [34]. In this analysis, we excluded study participants who reported having received a diagnosis of dementia, Alzheimer's disease, Parkinson's disease or any other chronic degenerative neurological problem, demyelinating diseases, brain cancer, nervous system infection, brain abscess, encephalitis, cerebral palsy, head or neurological injury/trauma, brain hemorrhage, cerebral aneurysm, or stroke (N= 388). We performed linear regression analyses (additive models) for ln-transformed WMH volume, the first principal components of all measurements of mean

diffusivity and fractional anisotropy across the different white matter tracts in the diffusion sequences, and for normalized grey matter and total brain volumes. Adjustments were made for age, sex, mean resting and task functional MRI head motion, the genotype platform array, and the first 10 principal components of the population structure.

Statistical analysis

All analyses were performed in R (v3.5.0; The R Foundation for Statistical Computing) using the MendelianRandomization, TwoSampleMR, and the MR-PRESSO packages.

Main analyses. We applied two-sample MR using association estimates derived from the abovementioned sources. Following extraction of the SNP-specific association estimates between the instruments and the outcomes, and harmonization of the direction of estimates by effect alleles, we computed MR estimates for each instrument with the Wald estimator. We calculated standard errors with the Delta method. We then pooled individual MR estimates using random-effects inverse-variance weighted (IVW) meta-analyses [35]. For the main analyses, we corrected for multiple comparisons with the false discovery rate (FDR) approach and set statistical significance at q-value<0.05. Associations not reaching this threshold, but showing an unadjusted p<0.05 were considered of nominal significance.

Assessment of pleiotropy and sensitivity analyses. MR estimates derived from the IVW approach could be biased in the presence of directional horizontal pleiotropy. As a measure of overall pleiotropy, we assessed heterogeneity across the SNP-specific MR estimates in the IVW MR analyses with the Cochran's Q statistic (statistical significance set at p<0.05) [36]. We further applied alternative MR methods which are more robust to pleiotropic variants. The weighted median estimator allows the use of invalid instruments as long as at least half of the instruments used in the MR analysis are valid [37]. The MR-Egger regression allows for the

estimation of an intercept term that can be used as an indicator of unbalanced directional pleiotropy [38]. MR-Egger provides less precise estimates and relies on the assumption that the strengths of potential pleiotropic instruments are independent of their direct associations with the outcome [38]. The intercept obtained from MR-Egger regression was used as a measure of unbalanced pleiotropy ($p<0.05$ indicated significance) [38]. Finally, MR-PRESSO regresses the SNP-outcome estimates against the SNP-exposure estimates to test for outlier SNPs [39]. Outliers are detected by sequentially removing all variants from the analyses and comparing the residual sum of squares as a global measure of heterogeneity ($p<0.05$ for detecting outliers); outliers are then removed and outlier-corrected estimates are provided. MR-PRESSO still relies on the assumption that at least half of the variants are valid instruments [39]. Finally, when significant results were found, we also applied bidirectional MR analyses to test for any inverse associations using diabetes and glucose-related traits as outcomes and stroke subtypes as exposures. For these analyses, due to the low number of SNPs associated with stroke or stroke subtypes, we lowered our p-value threshold for selecting genetic instruments at $p<10^{-6}$.

Primary research question/ Classification of evidence

Is genetic predisposition to T2D and hyperglycemia associated with the risk of stroke subtypes? This study provides Class II evidence that genetic predisposition to T2D and higher HbA1c levels are associated with a higher risk of large artery ischemic stroke (OR per 1-log-increment in T2D odds: 1.22, 95%CI: 1.17-1.28; OR per 1%-increment in HbA1c levels: 2.06, 95%CI: 1.60-2.66), and small vessel ischemic stroke (OR per 1-log-increment in T2D odds: 1.18, 95%CI: 1.13-1.23; OR per 1%-increment in HbA1c levels: 1.85, 95%CI: 1.50-2.27).

Standard Protocol Approvals, Registrations, and Patient Consents

This study, conducted in accordance with the STROBE-MR criteria[15] was based on publicly available summary statistics from GWAS meta-analyses of individual studies that had already obtained ethical review board approvals and that had obtained written informed consent from all included patients or their guardians.

Data availability statement

This study was based on summary statistics. Data sources are detailed in **Table 1**. The data from the GWAS studies for ischemic stroke, ICH, and glycemic traits are publicly available and may be accessed through the MEGASTROKE,[40] the ISGC [41], and the MAGIC [42] websites, respectively. Data from the UK Biobank GWAS for the neuroimaging traits may be accessed through an application to the UK Biobank. Data for the carotid plaque phenotype may be accessed through an application to the CHARGE Consortium. The detailed information on the genetic variants used as instruments to produce the presented results are available as Supplementary material (**Tables e-2** to **e-8**).

RESULTS

The 289 genetic variants used as genetic instruments for T2D explained 12.7% of the variance in T2D prevalence (**Table e-2**), whereas variants used as instruments for the continuous hyperglycemia traits, insulin resistance (proxied by fasting insulin levels), and β-cell dysfunction (proxied by fasting proinsulin), explained lower proportions of variance: 2.6% for HbA1c among both diabetic and non-diabetic individuals, 1.9% for HbA1c among non-diabetic individuals, 1.5% for fasting glucose, 0.7% for insulin resistance, and 4.5% for β-cell dysfunction (**Tables e-1 to e-5**).

Genetic predisposition to type 2 diabetes mellitus and risk of stroke

In the primary IVW MR analyses, genetic predisposition to T2D (1-log-increment=2.72-fold higher odds) was significantly associated with a higher risk of any ischemic stroke (OR: 1.11, 95%CI: 1.08-1.13), large artery stroke (OR: 1.22, 95%CI: 1.17-1.28), and small vessel stroke (OR: 1.18, 95%CI: 1.13-1.23; **Figure 1A**). In addition, there was an association of nominal significance with higher risk of cardioembolic stroke (OR: 1.05, 95%CI: 1.01-1.09), but no significant association with ICH (OR: 1.09, 95%CI: 0.97-1.23; **Figure 1A**). With the exception of ICH, there was evidence of significant heterogeneity in all of the main analyses ($p<0.05$; **Table e-9**), but no evidence of unbalanced pleiotropy, as assessed by the Egger intercept p-values (all $p>0.05$; **Table e-10**). Across sensitivity analyses based on alternative MR methods (weighted median, MR-Egger, outlier-corrected MR-PRESSO), all effects remained directionally consistent and all estimates stable with $p<0.05$ for any ischemic stroke, large artery stroke, and small vessel stroke (**Table e-10**). Similar results were also obtained when restricting the analyses to the European population of MEGASTROKE (**Table e-10**). Bidirectional MR analyses showed no effect of genetic predisposition to any ischemic stroke, large artery stroke, or small vessel stroke on risk of T2D (**Table e-11**).

Genetic predisposition to measures of hyperglycemia and risk of stroke

In analyses of hyperglycemia traits we found that genetically predicted HbA1c levels (per 1%-increment) were significantly associated with risk of any ischemic stroke (OR: 1.36, 95%CI: 1.21-1.53), large artery stroke (OR: 2.06, 95%CI: 1.60-2.66), and small vessel stroke (OR: 1.85, 95%CI: 1.50-2.27; **Figure 1B**). There was evidence of heterogeneity in the analyses for HbA1c levels (**Table e-8**) and in some alternative MR analyses the effect estimates for any ischemic stroke, large artery stroke, and small vessel stroke were smaller (**Table e-8**). However, in sensitivity analyses that excluded SNPs influencing HbA1c levels through

erythrocyte-related traits, the association estimates were even larger (ischemic stroke, OR: 1.53, 95%CI: 1.35-1.75; large artery stroke, OR: 2.83, 95%CI: 2.06-3.89; small vessel stroke, OR: 2.26, 95%CI: 1.72-2.97; **Table e-10**) and there was no evidence of heterogeneity (all p>0.10). Similar results were obtained when restricting analyses for stroke subtypes to the European population of MEGASTROKE, as well as when performing analyses for HbA1c in the non-diabetic range among diabetes-free individuals (**Figure e-2**; **Table e-10**). In bidirectional MR analyses genetic predisposition to any ischemic stroke, large artery stroke, or small vessel stroke was not associated with HbA1c levels (**Table e-11).** In contrast, we found no significant associations between genetically predicted fasting glucose levels among diabetes-free individuals and risk of stroke subtypes (**Figure e-2**; **Table e-10**).

Genetic predisposition to insulin resistance, β-cell dysfunction, and risk of stroke

We next selected genetic variants as instruments for insulin resistance and β-cell dysfunction, the two primary underlying mechanisms contributing to the development of hyperglycemia and T2D. Among diabetes-free individuals, we found genetic predisposition to insulin resistance (1-log increment in fasting insulin levels) to be associated with a higher risk for ischemic stroke (OR: 1.33, 95%CI: 1.13-1.57), large artery stroke (OR: 1.60, 95%CI: 1.12-2.31), and small vessel stroke (OR: 1.63, 95%CI: 1.21-2.20; **Figure 2A**). Genetic predisposition to β-cell dysfunction (1-log increment in fasting proinsulin levels) was further associated with a higher risk for small vessel stroke (OR: 1.38, 95%CI: 1.17-1.63) and ICH (OR: 1.75, 95%CI: 1.21-2.52). Furthermore, there was an association of nominal significance between genetic predisposition to β-cell dysfunction and the risk of cardioembolic stroke (OR: 1.18, 95%CI: 1.03-1.35). There was no heterogeneity in these analyses (**Table e-9**) and the results were consistent in alternative MR analyses, as well as in analyses restricted to individuals of European ancestry (**Table e-10**).

To gain additional insights in the relationship between insulin resistance, β-cell dysfunction, and etiological stroke subtypes, we further explored the effects of T2D-associated variants clustered in five different mechanisms of action. These included three clusters for insulin resistance (mediated by obesity, fat distribution, lipid metabolism) and two clusters related to β-cell dysfunction (associated with high or low proinsulin). In multivariable analyses including all clusters and also adjusting for their effects on HbA1c, we found significant effects of genetic predisposition to β-cell dysfunction related to high proinsulin on risk of ischemic stroke and small vessel stroke (**Figure 2B**). We further found genetic predisposition to insulin resistance mediated through altered fat distribution to be associated with higher risk of small vessel stroke. Genetic predisposition to insulin resistance mediated through obesity showed associations of nominal significance with large artery and cardioembolic stroke.

Genetic predisposition to type 2 diabetes and glycemis traits and associations with etiologically related cerebrovascular phenotypes

Table 2 presents the MR associations of genetic predisposition to T2D, measures of hyperglycemia, insulin resistance, and β-cell dysfunction, with carotid plaque, as well as with neuroimaging traits related to white matter integrity and brain atrophy. Genetic predisposition to T2D and genetically elevated HbA1c levels were associated with carotid plaque. We further found a significant association between genetic predisposition to T2D and lower fractional anisotropy, a diffusion imaging marker of impaired white matter tract integrity, as well as significant associations with lower grey matter and total brain volumes (**Table 2**). Genetic predisposition to β-cell dysfunction (1-log increment in fasting proinsulin levels) was further associated with lower grey matter volume (beta: -0.13, 95%CI: -0.20 to -0.07) and total brain volume (beta: -0.17, 95%CI: -0.23 to -0.11; **Table 2**). These results remained stable in sensitivity analyses (**Table e-10**).

DISCUSSION

Levaraging large-scale GWAS data in MR analyses, we investigated the causal associations between T2D, glycemic traits, and cerebrovascular disease. We found genetic predisposition to T2D and hyperglycemia (elevated HbA1c levels) to be associated with a higher risk of ischemic stroke, particularly large artery and small vessel stroke. Independently of hyperglycemia, genetic predisposition to insulin resistance but not β-cell dysfunction was associated with higher risk of large artery stroke, whereas genetic predisposition to both insulin resistance and β-cell dysfunction was associated with small vessel stroke. Genetic determinats for T2D and hyperglycemia further showed significant effects on carotid plaque and fractional anisotropy, a WM neuroimaging marker related to cerebral small vessel disease, as well as neuroimaging markers of brain atrophy. Furthermore, genetic predisposition to β-cell dysfunction was associated with intracerebral hemorrhage and neuroimaging markers of brain atrophy.

Our MR results provide genetic evidence for a causal effect of T2D, and also hyperglycemia on risk of ischemic stroke. While T2D is among the established risk factors for stroke and vascular disease in general [4], primary prevention trials focusing on intensive glucose control or specific oral anti-diabetic agents showed inconsistent effects on stroke risk [6,8]. Previous Mendelian randomization studies were underpowered to detect effects of hyperglycemia (HbA1c or fasting glucose levels) on stroke risk [43,44]. Here, by using data from >400,000 individuals from the UK Biobank, we were able to show that genetically elevated HbA1c levels are associated with a higher risk of ischemic stroke, thus suggesting that preventive strategies focusing on long-term HbA1c-lowering will result in risk reductions for ischemic stroke. The lack of significant effects in previous trials might relate to insufficient power due to the low number of incident

stroke events, short follow-up periods, and differences in the efficacy profiles of the individual treatments [45].

We found the effects of genetic predisposition to T2D and hyperglycemia to be specific for large artery and small vessel stroke. In accordance with these results, we found genetic predisposition to T2D to be associated with carotid plaque, an atherosclerotic phenotype, and fractional anisotropy, a marker of WM integrity associated with small vessel disease. Thus, our findings provide evidence for a causal involvement of T2D and hyperglycemia in both large artery atherosclerosis and cerebral small vessel disease. The discordant effects between genetically predicted HbA1c and fasting glucose levels might relate to the fact that HbA1c levels are a more accurate marker of average glucose levels and less prone to between-measurement variability than single measurements of fasting glucose. Differences in sample sizes between the GWASs, as well as the inclusion of non-diabetic patients in the analysis for HbA1c levels might also partly explain this discordance. On the contrary, we found no significant effects of T2D or other diabetic traits on cardioembolic stroke. Differences in the magnitude of the effects between stroke subtypes might in part explain the heterogeneity in the effects of glucose-lowering treatments across previous clinical trials.[45] On the basis of our findings, future trials testing glucose-lowering approaches should account for stroke subtypes.

As another finding, we show that genetic predisposition to insulin resistance and β-cell dysfunction influences the risk of stroke. This could have clinical implications for oral anti-diabetic medications. While all anti-diabetic agents lower glucose levels, some drug classes primarily target insulin sensitivity whereas others primarily target β-cell function.[9] Specifically, metformin and thiazolidinediones primarily act by improving insulin sensitivity, whereas drug classes like, α-glucosidase inhibitors, sulfonylureas, and GLP1 receptor agonists primarily act by increasing insulin secretion from the β-cells.[9] How these drug classes influence risk of the different stroke subtypes should be further explored in future research.

Our study has several methodological strengths. The large sample size (898,130 individuals for diabetic traits and up to 514,791 individuals for stroke) and nature of our datasets provided the power to detect differential effects of diabetes on etiological stroke subtypes and to perform multiple sensitivity analyses for testing the validity of the MR assumptions, thus minimizing the possibility of biased results. While the genetic determinants of HbA1c might influence its levels via both erythrocyte and glycemic biology, we provided support for the latter, as the effects were stronger when focusing on variants not associated with erythrocyte traits. Incorporating insulin resistance and β-cell dysfunction on top of hyperglycemia in the analyses offered deeper insights into the pathophysiological mechanisms linking diabetes with the different stroke subtypes. Finally, the exploration of additional cerebrovascular disease traits enabled us to triangulate our findings for stroke subtypes by showing similar associations for etiologically related phenotypes.

Our study also has limitations. First, by design MR examines the effects of lifetime exposure to the traits of interest, which might differ from the effects of clinical interventions (e.g. glucose-lowering approaches) applied for shorter time periods later in life. Second, T2D was analyzed as a binary trait and this might violate the monotonicity assumption of MR because only a fraction of individuals with increased genetic liability to T2D will actually get the disease. Thus, genetic liability to T2D that is used as an exposure in our analyses might capture a combination of underlying mechanisms including hyperglycemia, insulin resistance, and β-cell dysfunction. Third, the MR analyses for insulin resistance were weighted based on the effects of the genetic variants on fasting insulin adjusting for BMI and the analyses for β-cell dysfunction based on the effects of the variants on fasting pro-insulin adjusting for fasting insulin. These adjustments in the original GWASs might increase the risk for collider bias in MR analyses [46], which should be considered when interpreting our findings. Fourth, the analyses for HbA1c and fasting glucose that were restricted to non-diabetic individuals might

also introduce collider bias in the analyses, which might bias the association estimates to the null. Yet, the results for HbA1c in the entire population of both diabetic ad non-diabetic individuals showed similar results. Fifth, the variance explained by the genetic instruments used for hyperglycemic traits, insulin resistance, and β-cell dysfunction was very low, which might have limited the power of our analyses. However, despite the low proportion of variance explained, the instruments were sufficiently strong, thus ruling out potential weak instrument bias. Sixth, there was high heterogeneity in the majority of the MR analyses performed for this study. While the results from alternative MR methods were consistent, we cannot entirely rule out the possibility of bias in the derived effect estimates due to pleiotropic effects of the genetic instruments. Seventh, ischemic stroke subtypes were defined according to the TOAST classification system, which although widely used, might still inherently lead to misclassifications, especially in cases of mixed stroke etiology. Eighth, many of our exposure phenotypes like HbA1c levels, fasting glucose, and fasting insulin are time-dependent and might change with age, disease stage, and behavioral factors, as well as by epigenetic factors. However, our MR analyses are inherently limited in not taking such effects into account. Novel methods in addressing the time-varying effects [47] of these phenotypes on stroke subtypes should be examined in the future using datasets with available data. Finally, our analyses were primarily based on datasets involving individuals of European ancestry and might thus not be applicable to other ethnicities.

In conclusion, our results suggest causal associations of T2D and hyperglycemia with a higher risk for ischemic stroke, particularly large artery and small vessel stroke. Against findings from secondary analyses of clinical trials, our results support that therapeutic approaches aimed at lowering HbA1c have the potential to decrease the risk of ischemic stroke.

Appendix 1. Authors.

Name	Location	Role	Contribution
Marios K. Georgakis, MD, MSc	LMU Munich, Germany	Author	Concept and design; data acquisition, analysis, and interpretation of data; statistical analysis; drafting of the manuscript; critical revision of the manuscript for intellectual content
Eric L Harshfield, PhD	Cambridge University, UK	Author	Concept and design; data acquisition, analysis, and interpretation of data; critical revision of the manuscript for intellectual content
Rainer Malik PhD	LMU Munich, Germany	Author	Data acquisition, analysis, and interpretation of data; statistical analysis; critical revision of the manuscript for intellectual content
Nora Franceschini, MD, MPH	UNC Gillings, NC, USA	Author	Data acquisition, analysis, and interpretation of data; critical revision of the manuscript for intellectual content
Claudia Langenberg, MD, PhD	Cambridge University, UK	Author	Concept and design; Data acquisition, analysis, and interpretation of data; critical revision of the manuscript for intellectual content
Nicholas J. Wareham, MD, PhD	Cambridge University, UK	Author	Concept and design; data acquisition, analysis, and interpretation of data; critical revision of the manuscript for intellectual content
Hugh S. Markus, DM, F Med Sci	Cambridge University, UK	Author	Concept and design; data acquisition, analysis, and interpretation of data; critical revision of the manuscript for intellectual content
Martin Dichgans, MD	LMU Munich, Germany	Author	Concept and design; data acquisition, analysis, and interpretation of data; critical revision of the manuscript for intellectual content

REFERENCES

1. G. B. D. DALYs and Hale Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016;388:1603-1658.
2. G. B. D. Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016;388:1459-1544.
3. Cosentino F, Grant PJ, Aboyans V, et al. 2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD. Eur Heart J 2019.
4. Meschia JF, Bushnell C, Boden-Albala B, et al. Guidelines for the primary prevention of stroke: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2014;45:3754-3832.
5. Emerging Risk Factors C, Sarwar N, Gao P, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet 2010;375:2215-2222.
6. Fang HJ, Zhou YH, Tian YJ, Du HY, Sun YX, Zhong LY. Effects of intensive glucose lowering in treatment of type 2 diabetes mellitus on cardiovascular outcomes: A meta-analysis of data from 58,160 patients in 13 randomized controlled trials. Int J Cardiol 2016;218:50-58.
7. Action to Control Cardiovascular Risk in Diabetes Study G, Gerstein HC, Miller ME, et al. Effects of intensive glucose lowering in type 2 diabetes. N Engl J Med 2008;358:2545-2559.
8. Ray KK, Seshasai SR, Wijesuriya S, et al. Effect of intensive control of glucose on cardiovascular outcomes and death in patients with diabetes mellitus: a meta-analysis of randomised controlled trials. Lancet 2009;373:1765-1772.
9. Tahrani AA, Barnett AH, Bailey CJ. Pharmacology and therapeutic implications of current drugs for type 2 diabetes mellitus. Nat Rev Endocrinol 2016;12:566-592.
10. Zethelius B, Byberg L, Hales CN, Lithell H, Berne C. Proinsulin is an independent predictor of coronary heart disease: Report from a 27-year follow-up study. Circulation 2002;105:2153-2158.
11. Gast KB, Tjeerdema N, Stijnen T, Smit JW, Dekkers OM. Insulin resistance and risk of incident cardiovascular events in adults without diabetes: meta-analysis. PLoS One 2012;7:e52036.
12. Holmes MV, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. Nature reviews Cardiology 2017;14:577-590.
13. O'Donnell CJ, Sabatine MS. Opportunities and Challenges in Mendelian Randomization Studies to Guide Trial Design. JAMA Cardiol 2018;3:967.
14. Georgakis MK, Gill D, Rannikmae K, et al. Genetically Determined Levels of Circulating Cytokines and Risk of Stroke. Circulation 2019;139:256-268.
15. Davey Smith G, Davies NM, Dimou N, et al. STROBE-MR: Guidelines for strengthening the reporting of Mendelian randomization studies. . PeerJ Preprints 2019;7:e27857v27851.
16. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet 2018;50:1505-1513.
17. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. Science 2002;296:2225-2229.
18. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12:e1001779.
19. Scott RA, Lagou V, Welch RP, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat Genet 2012;44:991-1005.
20. Wheeler E, Leong A, Liu CT, et al. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. PLoS Med 2017;14:e1002383.
21. Kamat MA, Blackshaw JA, Young R, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 2019;35:4851-4853.
22. Heid IM, Winkler TW. A multitrait GWAS sheds light on insulin resistance. Nat Genet 2016;49:7-8.
23. Strawbridge RJ, Dupuis J, Prokopenko I, et al. Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. Diabetes 2011;60:2624-2634.
24. Li M, Feng D, Zhang K, Gao S, Lu J. Disproportionately Elevated Proinsulin Levels as an Early Indicator of beta-Cell Dysfunction in Nondiabetic Offspring of Chinese Diabetic Patients. Int J Endocrinol 2016;2016:4740678.
25. Udler MS, Kim J, von Grotthuss M, et al. Type 2 diabetes genetic loci informed by multi-trait associations point to disease mechanisms and subtypes: A soft clustering analysis. PLoS Med 2018;15:e1002654.

26. So HC, Gui AH, Cherny SS, Sham PC. Evaluating the heritability explained by known susceptibility variants: a survey of ten complex diseases. Genet Epidemiol 2011;35:310-317.
27. Shim H, Chasman DI, Smith JD, et al. A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. PLoS One 2015;10:e0120758.
28. Ogurtsova K, da Rocha Fernandes JD, Huang Y, et al. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes Res Clin Pract 2017;128:40-50.
29. Malik R, Chauhan G, Traylor M, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet 2018;50:524-537.
30. Malik R, Rannikmae K, Traylor M, et al. Genome-wide meta-analysis identifies 3 novel loci associated with stroke. Ann Neurol 2018;84:934-939.
31. Adams HP, Jr., Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke 1993;24:35-41.
32. Woo D, Falcone GJ, Devan WJ, et al. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. Am J Hum Genet 2014;94:511-521.
33. Franceschini N, Giambartolomei C, de Vries PS, et al. GWAS and colocalization analyses implicate carotid intima-media thickness and carotid plaque loci in cardiovascular outcomes. Nat Commun 2018;9:5141.
34. Cox SR, Lyall DM, Ritchie SJ, et al. Associations between vascular risk factors and brain MRI indices in UK Biobank. Eur Heart J 2019;40:2290-2300.
35. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013;37:658-665.
36. Bowden J, Hemani G, Davey Smith G. Invited Commentary: Detecting Individual and Global Horizontal Pleiotropy in Mendelian Randomization-A Job for the Humble Heterogeneity Statistic? Am J Epidemiol 2018;187:2681-2685.
37. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol 2017;46:1985-1998.
38. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512-525.
39. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018;50:693-698.
40. MEGASTROKE: Multi-ancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes [online]. Available at: https://www.megastroke.org/download.html. Accessed 26 Aug 2020.
41. International Stroke Genetics Consortium (ISGC) [online]. Available at: http://www.kp4cd.org/dataset_downloads/stroke. Accessed 26 Aug 2020.
42. MAGIC (the Meta-Analyses of Glucose and Insulin-related traits Consortium) [online]. Available at: https://www.magicinvestigators.org/downloads/. Accessed 26 Aug 2020.
43. Larsson SC, Scott RA, Traylor M, et al. Type 2 diabetes, glucose, insulin, BMI, and ischemic stroke subtypes: Mendelian randomization study. Neurology 2017;89:454-460.
44. Liu J, Rutten-Jacobs L, Liu M, Markus HS, Traylor M. Causal Impact of Type 2 Diabetes Mellitus on Cerebral Small Vessel Disease: A Mendelian Randomization Analysis. Stroke 2018;49:1325-1331.
45. Bonnet F, Scheen AJ. Impact of glucose-lowering therapies on risk of stroke in type 2 diabetes. Diabetes Metab 2017;43:299-313.
46. Holmes MV, Davey Smith G. Problems in interpreting and using GWAS of conditional phenotypes illustrated by 'alcohol GWAS'. Mol Psychiatry 2019;24:167-168.
47. Georgakis MK, Gill D, Malik R, Protogerou AD, Webb AJS, Dichgans M. Genetically Predicted Blood Pressure Across the Lifespan: Differential Effects of Mean and Pulse Pressure on Stroke Risk. Hypertension 2020;76:953-961.

Table 1. Data sources that were used in the analyses for the current study.

Phenotype	Source	N (Total or Cases/Controls	Imputation reference panel	Ancestry	Adjustments
Diabetes mellitus type 2	DIAGRAM Consortium [16]	74,124/824,006	HRC	European	age, sex, 6 PCs
HbA1c	UK Biobank [18]	421,923	HRC + UK10K	White British	age, sex, 20 PCs, genotyping platform array, assessment center
Fasting glucose levels	MAGIC Consortium [19]	133,010	HapMap	European	age, sex
Insulin resistance (fasting insulin levels)	Multi-trait GWAS and MAGIC Consortium [19]	108,557	HapMap	European	age, sex, BMI
β-cell dysfunction (fasting proinsulin levels)	MAGIC Consortium [23]	16,378	1000 Genomes	European	age, sex, fasting insulin
Any ischemic stroke	MEGASTROKE Consortium [29]	60,341/454,450	1000 Genomes	Trans-ethnic (70% European)	age, sex, population structure up to 20 PCs
Large artery stroke	MEGASTROKE Consortium [29]	6,688/454,450	1000 Genomes	Trans-ethnic (70% European)	age, sex, population structure up to 20 PCs
Cardioembolic stroke	MEGASTROKE Consortium [29]	9,006/454,450	1000 Genomes	Trans-ethnic (70% European)	age, sex, population structure up to 20 PCs
Small vessel stroke	MEGASTROKE Consortium [29]	11,710/454,450	1000 Genomes	Trans-ethnic (70% European)	age, sex, up to 20 PCs
Intracerebral hemorrhage	ISGC meta-analysis [32]	1,545/1,481	1000 Genomes	European	age, sex, 4 PCs
Carotid plaque	CHARGE Consortium [33]	21,540/26,894	1000 Genomes	European	age, sex, up to 10 PCs
WMH volume	UK Biobank imaging database [34]	17,534	HRC + UK10K	White British	age, sex, mean resting and task functional MRI head motion, 10 PCs, genotyping platform array
Mean diffusivity	UK Biobank imaging database [34]	17,534	HRC + UK10K	White British	age, sex, mean resting and task functional MRI head motion, 10 PCs, genotyping platform array
Fractional anisotropy	UK Biobank imaging database [34]	17,534	HRC + UK10K	White British	age, sex, mean resting and task functional MRI head motion, 10 PCs, genotyping platform array
Normalized grey matter volume	UK Biobank imaging database [34]	17,534	HRC + UK10K	White British	age, sex, mean resting and task functional MRI head motion, 10 PCs, genotyping platform array
Normalized total brain volume	UK Biobank imaging database [34]	17,534	HRC + UK10K	White British	age, sex, mean resting and task functional MRI head motion, 10 PCs, genotyping platform array

PC: principal component.

Table 2. Mendelian randomization associations between genetically predicted diabetic traits and etiologically related cerebrovascular phenotypes, as derived from random-effects inverse-variance weighted analyses.

Outcomes	Exposures			
	T2D (1-log-odds increment)	**HbA1c** (1%-increment)	**Insulin resistance** (1 log-increment in fasting insulin levels)	**β-cell dysfunction** (1 log-increment in fasting proinsulin levels)
Carotid atherosclerosis	odds ratios (95%CI)			
Carotid plaque	**1.06 (1.03-1.10)**	1.21 (1.03-1.42) [a]	0.93 (0.83-1.05)	1.10 (0.80-1.50)
White matter integrity	beta coefficients (95%CI)			
WMH volume	0.003 (-0.010, 0.019)	-0.002 (-0.081, 0.077)	0.094 (-0.062, 0.251)	0.062 (-0.021, 0.146)
Mean diffusivity	0.005 (-0.016, 0.026)	--0.086 (-0.171, -0.002) [a]	0.146 (-0.056, 0.347)	0.048 (-0.017, 0.114)
Fractional anisotropy	**-0.028 (-0.048, -0.006)**	-0.008 (-0.118, 0.101)	-0.181 (-0.380, 0.019)	-0.048 (-0.115, 0.020)
Brain atrophy	beta coefficients (95%CI)			
Grey matter volume	**-0.031 (-0.048, -0.013)**	-0.074 (-0.143, -0.005) [a]	-0.039 (-0.220, 0.142)	**-0.130 (-0.195, -0.065)**
Total brain volume	**-0.027 (-0.047, -0.008)**	**-0.181 (-0.272, -0.089)**	-0.087 (-0.285, 0.112)	**-0.170 (-0.232, -0.108)**

Odds Ratios are presented for binary traits (carotid plaque) and beta coefficients (standardized based on the SD of the measure) for the continuous imaging traits.

Bold indicates statistical significance at an FDR-adjusted p-value<0.05.

[a] Associations reaching nominal significance (unadjusted p<0.05).

FIGURE LEGENDS

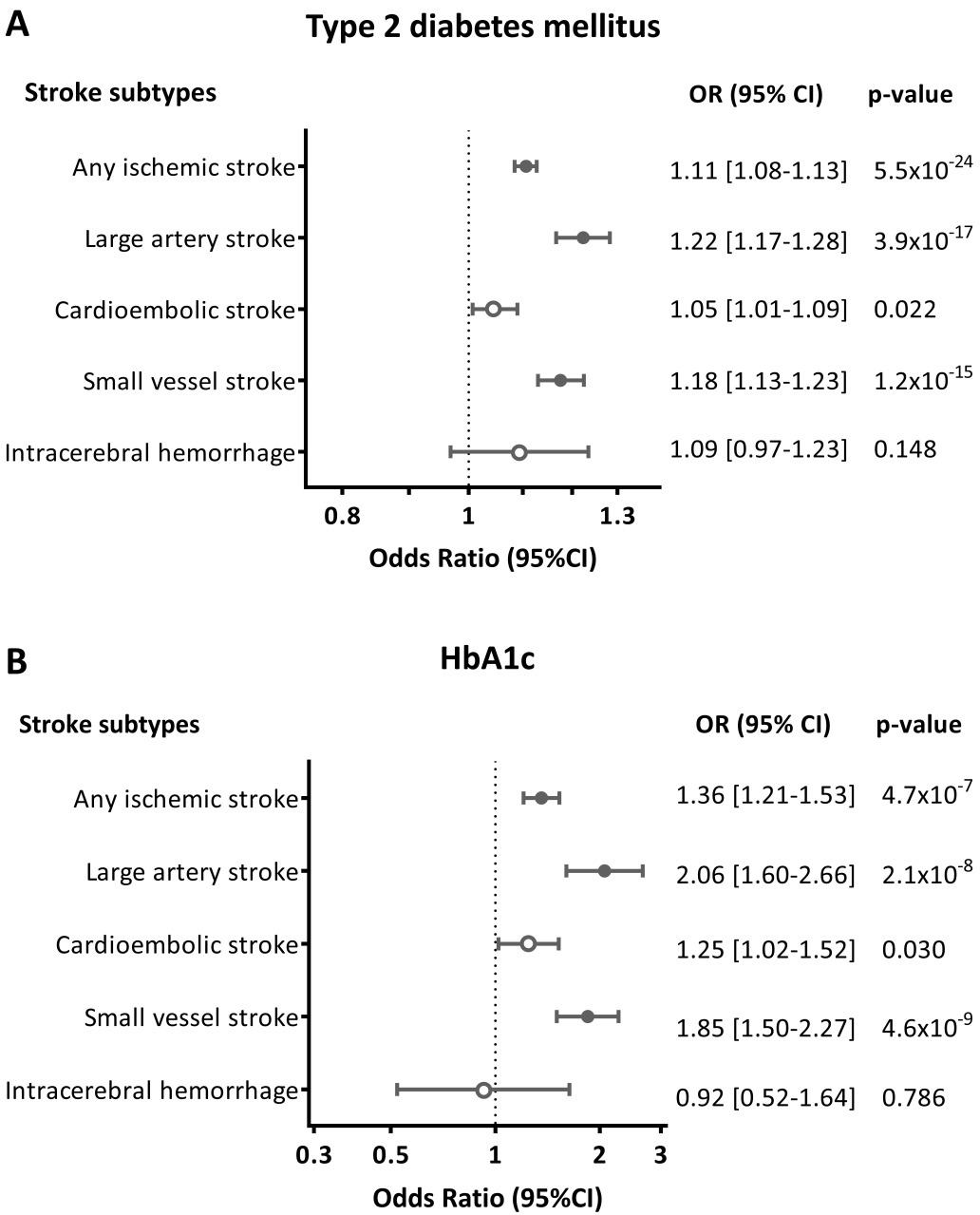
Figure 1. Mendelian Randomization associations of genetic predisposition to (A) type 2 diabetes mellitus, and (B) HbA1c levels among both diabetic and non-diabetic individuals. Results derived from random-effects inverse-variance weighted analyses.

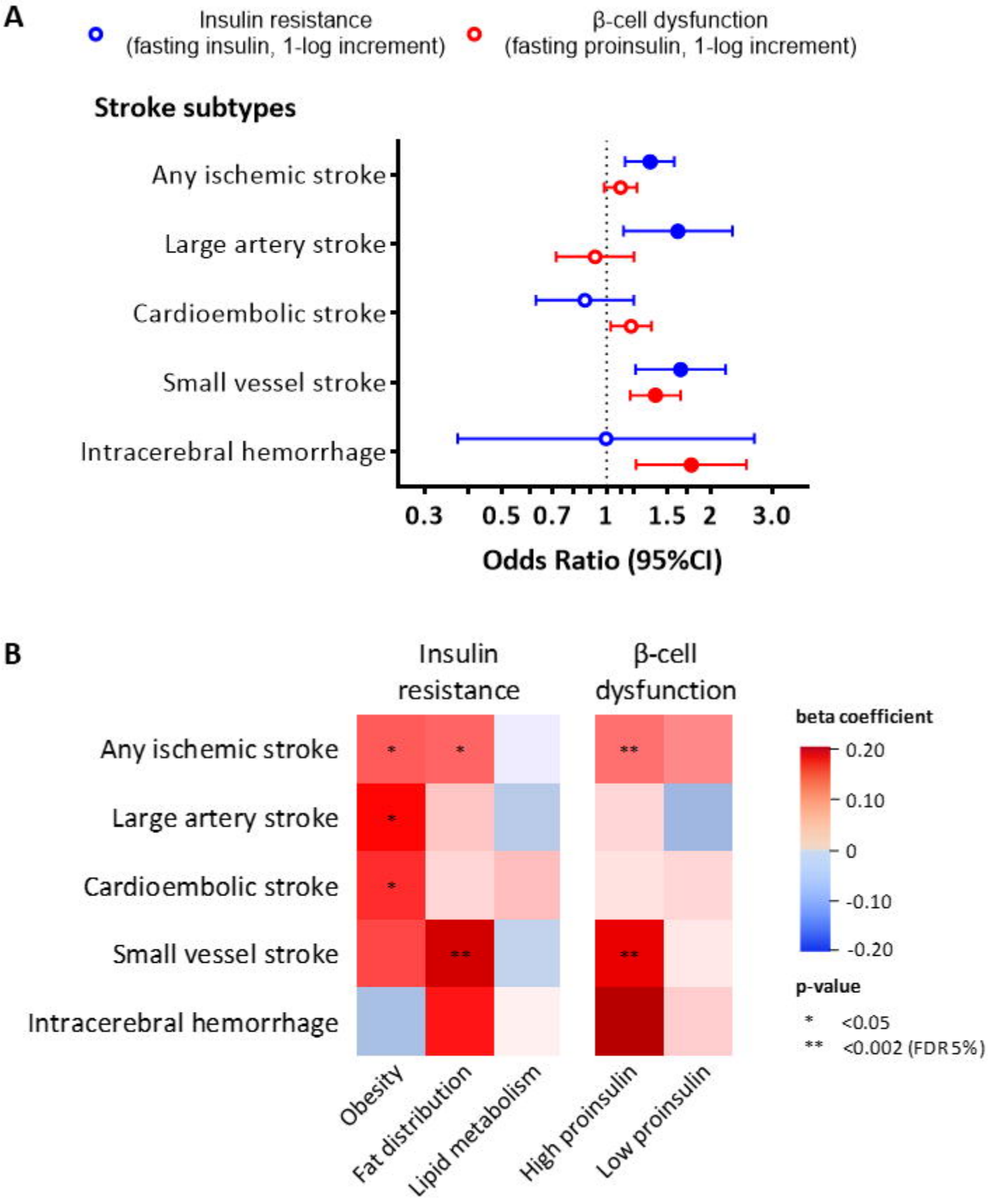
Full circles correspond to statistically significant association estimates at an FDR-adjusted p-value<0.05.

Abbreviations. HbA1c, Glycated hemoglobin.

Figure 2. Mendelian Randomization associations of genetically predicted insulin resistance and β-cell dysfunction with stroke subtypes. (A) Results derived from random-effects inverse-variance weighted analyses. (B) Heatmap of the associations between clusters of diabetic endophenotypes related to β-cell dysfunction and insulin resistance with the risk of stroke subtypes.

Full colored circles in panel A correspond to statistically significant association estimates at an FDR-adjusted p-value<0.05.


A
Type 2 diabetes mellitus
Stroke subtypes
OR (95% CI)
p-value
Any ischemic stroke
1.11 [1.08-1.13]
5.5x10-24
Large artery stroke
1.22 [1.17-1.28]
3.9x10-17
Cardioembolic stroke
1.05 [1.01-1.09]
0.022
Small vessel stroke
1.18 [1.13-1.23]
1.2x10-15
Intracerebral hemorrhage
1.09 [0.97-1.23]
0.148
0.8
1
1.3
Odds Ratio (95%CI)
B
HbA1c
Stroke subtypes
OR (95% CI)
p-value
Any ischemic stroke
1.36 [1.21-1.53]
4.7x10-7
Large artery stroke
2.06 [1.60-2.66]
2.1x10-8
Cardioembolic stroke
1.25 [1.02-1.52]
0.030
Small vessel stroke
1.85 [1.50-2.27]
4.6x10-9
Intracerebral hemorrhage
0.92 [0.52-1.64]
0.786
0.3
0.5
1
2
3
Odds Ratio (95%CI)


A
Insulin resistance
(fasting insulin, 1-log increment)
β-cell dysfunction
(fasting proinsulin, 1-log increment)
Stroke subtypes
Any ischemic stroke
Large artery stroke
Cardioembolic stroke
Small vessel stroke
Intracerebral hemorrhage
0.3
0.5
0.7
1
1.5
2
3.0
Odds Ratio (95%CI)
B
Insulin resistance
β-cell dysfunction
Any ischemic stroke
Large artery stroke
Cardioembolic stroke
Small vessel stroke
Intracerebral hemorrhage
Obesity
Fat distribution
Lipid metabolism
High proinsulin
Low proinsulin
beta coefficient
0.20
0.10
0
-0.10
-0.20
p-value
* <0.05
** <0.002 (FDR 5%)