Gut Microbiome and Stroke: a Bidirectional Mendelian Randomisation Study in East Asian and European Populations

Shiyao Cheng, Hao Zheng, Yuanan Wei, Xingchen Lin, Yuqin Gu, Xinxin Guo, Zhe Fan, Hao Li, Si Cheng, Siyang Liu

ABSTRACT

Background and aims Observational studies have implicated the involvement of gut microbiome in stroke development. Conversely, stroke may disrupt the gut microbiome balance, potentially causing systemic infections exacerbated brain infarction. However, the causal relationship remains controversial or unknown. To investigate bidirectional causality and potential ethnic differences, we conducted a bidirectional two-sample Mendelian randomisation (MR) study in both East Asian (EAS) and European (EU) populations.

Methods Leveraging the hitherto largest genome-wide association study (GWAS) summary data from the MiBioGen Consortium (n=18,340, EU) and BGI (n=2,524, EAS) for the gut microbiome, stroke GWAS data from the GIGASTROKE Consortium (264,655 EAS and 1,308,460 EU), we conducted bidirectional MR and sensitivity analyses separately for the EAS and EU population.

Results We identified nominally significant associations between 85 gut microbiome taxa in EAS and 64 gut microbiomes taxa in EU with stroke or its subtypes. Following multiple testing, we observed that genetically determined 1 SD increase in the relative abundance of species Bacteroides pectinophilus decreased the risk of cardioembolic stroke onset by 28% (OR 0.72 (95% CI 0.62 to 0.84); p=4.22e−5), and that genetically determined 1 SD increase in class Negativicutes resulted in a 0.76% risk increase in small vessel stroke in EAS. No significant causal association was identified in the EU population and the reverse MR analysis.

Conclusion Our study revealed subtype-specific and population-specific causal associations between gut microbiome and stroke risk among EAS and EU populations. The identified causality holds promise for developing a new stroke prevention strategy, warrants further mechanistic validation and necessitates clinical trial studies.

INTRODUCTION

Stroke is the leading cause of death and disability globally, particularly in East Asian areas like China. Ischaemic stroke accounts for 87% of all stroke cases, with most of them resulting from arterial occlusion. Risk factors for stroke can be categorised into modifiable factors, including hypertension (the most significant risk factor), hyperlipidaemia, smoking (accounting for 90% of stroke cases) and unmodifiable factors, such as genetic predisposition. Despite this knowledge, the current prevention and treatment strategies for stroke remain limited. Reperfusion, considered the most effective treatment for ischaemic stroke, is often impractical for the majority of patients due to the restricted time window. Therefore, there is a need to investigate risk factors that can be treated across different populations, accounting for genetic differences.

Recently, the role of the gut microbiome in disease progression has gained attention, including its association with cardiovascular disease, immune system disorders and neurological conditions, such as stroke. Research has shown that the gut microbiota...
exerts its effects by producing bioactive metabolites such as trimethylamine N-oxide, which has been linked to atherosclerosis, a significant risk factor for stroke and its subtypes. Conversely, stroke leads to dysbiosis of the gut microbiome and exacerbates brain infarction by promoting systemic inflammation.

However, the two-way communication theory studies have mainly been studied through observational studies or preclinical studies, with few studies establishing causal relationships. Furthermore, previous studies have highlighted substantial variations in both gut microbiome composition and stroke incidence between East Asian (EAS) and European (EU) populations, emphasising the need for population-specific investigations. Therefore, the existence and nature of causal relationships between the gut microbiome and stroke remain controversial and unknown.

To address this, we applied the bidirectional two-sample Mendelian randomisation (MR) method, which uses genetic variation as instrumental variables (IVs), to explore the two-way causality between the gut microbiome and the onset of stroke and its primary subtypes in EU and EAS populations. By employing the summary statistic method, specifically the inverse-variance weighted meta-analysis using summarised data from the recently published, largest genome-wide association study (GWAS) summary statistics of microbiome and stroke subtypes, we enhanced the statistical power of our study.

**METHODS**

**Data sources**

For the gut microbiome, we used two large-scale population-specific GWAS summary datasets for EU and EAS individually. For the EU population, we employed the MiBioGen Consortium’s GWAS meta-analysis study, which included 18,473 individuals from 24 population-based cohorts of EU ancestry. For the EAS population, we used the GWAS summary data from the metagenome-GWAS, which involved 25,45 Chinese individuals (1539 for the discovery cohort, 1006 for the replication cohort). In total, there were 499 gut microbiome traits in EAS and 211 traits in EU. Further details about the data source, including the URLs to download the summary statistics are provided in online supplemental table S1.

For stroke, we used the hitherto largest GWAS meta-analysis study, conducted by the GIGASTROKE consortium, which included more than one million participants and identified triple the number of novel loci compared with previous studies. Further details of the case-control design and the data source in the EAS and EU population are described in figure 1 and online supplemental table S1.

According to the cohort information disclosed in the original paper, there is no sample overlap among the GWAS summary data sources. Ethical approval and participants’ consent were provided in the original studies.

**Study design**

Based on the newly-released guidelines for MR analysis, we employed the bidirectional two-sample MR approach to evaluate the causal relationship between the gut microbiome and stroke in both EAS and EU populations, separately. To ensure the validity of potential causal effects, the MR analyses ought to fulfil the three assumptions:

1. The genetic variants selected as IVs are strongly associated with the exposure.
2. The genetic variants are not associated with any confounder in the exposure-outcome association, and
3. The genetic variants do not affect the outcome except through the association with the exposure. The overall study design of our study is illustrated in figure 1.

**IVs selection**

To choose SNPs from the GWAS summary data that fulfilled the three assumptions and served as IVs, we used the genome-wide conditional and joint association analysis (GCTA-COJO). This approach ensured the independence of multiple variants at a given locus, and we used these independent SNPs as genetic instruments in our subsequent MR analyses. For the GCTA-COJO analysis, we used in-house 10,000 Chinese high-depth sequencing data as a reference panel for East Asian population data, and the UK Biobank high-depth sequencing data as a reference panel for the EU population data. We included variants with a p value less than $1\times10^{-5}$ and a linkage disequilibrium $r^2$ less than 0.01 (gcta64 -bfile reference_panel_file -cojo-file gwas_file -cojo-slcet -cojo-p 1e-5 -cojo-collinear 0.01). We calculated the phenotype variance explained (PVE) and F-statistics of every SNP and phenotype to assess instrumental strength. We exclude the IVs and phenotypes with F-statistics less than 10. Details about the IVs are presented in GitHub (https://github.com/liusylab/GutMicrobiome_STROKE_2SMR).

The analyses of IV selection were performed using the GCTA software tool, V.1.93.2 and R V.4.2.2.

**Statistical analyses**

Before performing MR, we used the harmonise_data() function from the TwoSampleMR package to harmonise the exposure and outcome data. Subsequently, we conducted bidirectional two-sample univariable MR analyses, employing five statistical methods, including inverse variance weighted (IVW), weighted median, weighted mode, simple mode, and MR Egger, to assess causal relationships using multiple IVs. In cases where only one IV was selected, we used the Wald ratio method to assess causality. The IVW method, which is robust in the absence of directional pleiotropy, was used as the primary method to estimate the causal relationship. To mitigate the risk of multiple testing bias, we applied Bonferroni correction based on a categorisation scheme considering biological and aetiological characteristics. To deal with collinearity between the cross-taxonomy traits, the number of tests was determined by the product of the number of phylum/class/order/family/genus/species microbial
traits at various taxonomic levels and the number of independent stroke outcomes.

To avoid the risk of multiple testing bias, we conducted Bonferroni correction based on a categorisation scheme considering the biological and aetiological characteristics. We defined the number of tests as a product of the number of phylum/class/order/family/genus/species microbial traits and the number of independent stroke outcome. Specifically, in EAS, among the 499 gut microbiome traits, 480 containing IVs were classified into 9 phylum traits, 3 class traits, 13 order traits, 31 family traits, 92 genus traits, 233 species traits and 99 microbial features traits. In EU, among the 211 traits, 186 with IVs were classified into 9 phylum, 15 class, 19 order, 34 family and 109 species traits. The five stroke subtypes were treated as three independent categories, aligning with the categorisation scheme used in the GIGASTROKE study.

Supplementary and sensitivity analyses

We used the MR-Egger, MR-PRESSO, weighted median, simple mode, and weighted mode methods as sensitivity analysis. As for the horizontal pleiotropy, we conducted MR-Egger regression and used the intercept as a parameter to evaluate it. We also employed funnel plots to detect directional pleiotropy and assess horizontal pleiotropy.23 Heterogeneity was quantified by Cochran’s Q statistic in the IVW method or MR-Egger regression. Additionally, we performed the leave-one-out analysis to identify SNPs with strong effects and test the sensitivity of the bidirectional MR analysis results. MR Steiger analysis was used to test the sensitivity of the causal direction. To better validate our primary analysis, we introduced positive and negative controls in the sensitivity analysis. In addition, we did MR power analysis based on formula (3) from Brion et al., IJE, 2013.24 All statistical analyses, including pleiotropy and sensitivity analyses, were performed using the TwoSampleMR package in R V.4.2.2.

RESULT

To explore the causal relationship between the gut microbiome and stroke, we conducted a total of 2400 two sample MR tests in the EAS population, examining 480 microbiome taxa and five stroke outcomes (online supplemental table S2). Additionally, we performed 930 tests in the EU population, involving 186 microbiome taxa and 5 stroke outcomes (online supplemental table S3). Using the IVW method, we identified 85 significant MR analyses in the EAS population and 64 significant analyses (12 by Wald Ratio) in the EU population, with a
nominal p significance threshold of p<0.05. An overview of these findings is presented in figure 2.

Following rigorous multiple testing correction, we identified two gut microbiome traits in the EAS population that exhibited causal associations with stroke. Specifically, we observed 1 genetically predicted SD increase in the relative abundance of species *Bacteroides pectinophilus* will decrease 28% risk of cardioembolic stroke (CES) onset (OR 0.72 (95% CI 0.62 to 0.84); p=4.22e−5), and 1 genetically predicted SD decrease in the relative abundance of class *Negativicutes* will decrease 0.76% risk of small vessel stroke (SVS) onset (OR 1.076 (95% CI 1.02 to 1.13); p=4.20e−3). The two significant results remained insignificant in the reverse MR analysis (online supplemental Fig S1, Fig S2, Fig S3). The MR analysis power for species *B. pectinophilus* on CES (PVE=0.161, power=0.93) and class *Negativicutes* on SVS (PVE=0.163, power=0.62) were more than 0.6 and more than 0.9 for species *B. pectinophilus* on CES, indicating the validity of the MR analysis (online supplemental Fig S1, Fig S2, Fig S3). The MR analysis power for species *B. pectinophilus* on CES, indicating the validity of the MR analysis (online supplemental Fig S1, Fig S2, Fig S3). However, in the EU population, after applying multiple testing corrections, we did not identify significant associations. The causal relationship between species *B. pectinophilus* and CES was supported by the Weighted Median method, the Sample Mode method, and the Simple Mode methods, with consistent effect direction across all methods (figure 3). The causal relationship between class *Negativicutes* on SVS was supported by the Weighted Median and Sample Mode methods at the nominal significance threshold, with a consistent mean effect direction across all methods (figure 3). To enhance result robustness, we introduced positive controls through two-sample MR using the same parameters and methods as our analysis, and successfully replicated the causal effect of atrial fibrillation on CES in EAS population (online supplemental table S4). Negative control was selected from an MR investigation from a previous study on the causal effects of the microbiome on stroke outcome in EU population with about half of the stroke outcome sample size in our study. We found highly consistent results between the prior and our study (online supplemental table S5-S8). All the negative result in the previous study remains negative in our study after Bonferroni correction (online supplemental Fig S8, Fig S9, Fig S10, Table S8, online supplemental table S8).

We further employed the MR-Egger intercept from MR-Egger regression in conjunction with the funnel plot, to assess the presence of horizontal pleiotropy. We found no evidence of horizontal pleiotropy in species *B. pectinophilus* to CES and class *Negativicutes* to SVS (p>0.05) (figure 3, online supplemental fig S4, fig S5). Cochran’s Q test showed that all p values were greater than 0.05, indicating no evidence of heterogeneity. In the leave-one-out analysis, the results remained largely unchanged when excluding any single SNP, thus indicating the robustness of the findings (online supplemental fig S11, fig S12). The MR Steiger filtering test showed that the sensitivity ratio of Vz 0 /Vz 1 was significantly bigger than 1 (species *B. pectinophilus* to CES: −1162.6, class *Negativicutes* to SVS was −98.159), further approving the reverse MR analysis results (figure 3, online supplemental table S2). Notably, the *B. pectinophilus* species not only demonstrated a protective effect on CES but also SVS and the overall risk of AIS (figure 4).

Figure 2 The overview of the causality between gut microbiome and stroke. Results are based on the inverse variance weighted (IVW) method and Wald Ratio method. (A) The result for the gut microbiome to stroke in East Asian (EAS) population (species level), (B) the result for the gut microbiome to stroke in EAS population (class, family, genus, microbial features (MF), order, phylum) and part (C) the result for the gut microbiome to stroke in European (EU) population. Only the gut microbiomes associated with stroke are listed in figure 2.
above. To enhance the validation of our findings, we use IVs of species \textit{B. pectinophilus} and transethic stroke and its subtypes from MEGASTROKE consortium.\(^\text{26}\) The protective effects persisted in the MEGASTROKE, although with reduced strength for CES and SVS (online supplemental table S5). To better understand the nature of the species \textit{B. pectinophilus} to CES and class \textit{Negativicutes} to SVS, we also searched the related function of SNPs used as IVs in phenoscan V2.\(^\text{27,28}\) The annotation of these IVs did not suggest correlated pleiotropy, except that a proxy SNP of rs78852205 (R\(^2\geq0.8\)) has been related with vascular or heart problems diagnosed by doctor (online supplemental table S9). However, when removing rs78852205, we still observe a significant protective effect (online supplemental fig S1).

In the reverse MR analysis, we conducted 2495 (5×499, EAS, online supplemental table S10) and 1081 (5×216, EU, online supplemental table S11) tests to investigate the causal relationship between stroke and gut microbiome. Similar to the results when considering the gut microbiome as exposures, we found 93 (EAS) and 36 (EU, 2 by Wald ratio) potential causal relationships by the IVW or Wald ratio method passing the nominal significance threshold (online supplemental fig S1). However, after adjusting for multiple testing, no associations remain significant in either the EAS or EU population (online supplemental table S10-S11).

**DISCUSSION**

In our study, we conducted a bidirectional two-sample MR analysis to explore the causal relationship between the gut microbiome and stroke in EAS and EU populations, respectively. To ensure statistical power, we have used the largest stroke and gut microbiome cohorts with genetic information. We identified a set of gut microbiome traits

---

**Figure 3** Mendelian randomisation (MR) and sensitivity test results for species \textit{Bacteroides pectinophilus} to cardioembolic stroke (CES) and class \textit{Negativicutes} to small vessel stroke (SVS) in East Asian populations.

**Figure 4** The causal relation between species \textit{Bacteroides pectinophilus} and stroke and its subtypes.
that demonstrated nominal significance with stroke or its subtypes (p<0.05). After adjusting for multiple testing, two of these associations remained significant. In the EAS population, we observed a causal reduction in the risk of cardioembolic stroke (CES) associated with the species \textit{B. pectinophilus}, while the class \textit{Negativicutes} was found to be a risk factor for SVS. However, in the EU population, no robust genetic causal relationship was found after multiple testing, which is consistent with previous investigations in the EU population. In addition, we did not identify a causal impact of stroke on the gut microbiome.

Changes in the relative abundance of the gut microbiome are implicated in the onset of stroke and vary among different ancestral populations. Previous observational and MR studies have provided some insights into the associations and causal relationship between the gut microbiome and stroke, but the results have been conflicting and inconclusive and primarily focused on the EU population. Some studies have failed to pass multiple testing, suggesting a weak effect size. No previous studies have specifically examined the causal relationship between the gut microbiome and stroke in the EAS population. Our study found a weak association between certain lactate-producing bacteria, such as the species \textit{Bifidobacterium breve}, and stroke risk, which is consistent with an observational study conducted in the EAS population. In addition, we observed a protective role of opportunistic pathogens, such as the family \textit{Enterobacteriaceae}, in CES, which differed from previous observational studies. By addressing common limitations of epidemiological studies, such as unavoidable confounders and time-consuming designs, we found several gut microbiome traits that showed a potential causal relationship with stroke. The effects varied at different taxonomic levels and among different ancestral backgrounds, and their effects changed with different stroke and its subtypes, highlighting the importance of personalised precision medicine.

The causal relationship between gut microbiome abundance to stroke and its underlying mechanism is intricate. Through the application of the MR method and multiple testing, we found a genetic causal association between the abundance of species \textit{B. pectinophilus} and CES, and between class \textit{Negativicutes} and SVS. These associations remained significant even after correcting for multiple testing and exhibited a considerable effect. \textit{B. pectinophilus} is an anaerobic, rod-shaped bacteria found in the human intestinal tract that can deesterify pectin, an insoluble dietary fibre, via extracellular pectinesterase in the human intestinal tract, it is reasonable to deduce that an increased intake of food rich in insoluble dietary fibre may aid in stroke prevention, particularly CES and SVS. This hypothesis paves the way for a genetically informed understanding of dietary approaches for stroke prevention, such as the DASH (Medication and Dietary Approaches to Stop Hypertension) pattern. Based on the PhenoScanner (a database of human genotype–phenotype association), we found rs78852205 related with vascular or heart problems diagnosed by doctor, which was the selected IV in the causality between species \textit{B. pectinophilus} on CES. The \textit{Negativicutes} class comprises Gram-negative bacteria characterised by double membranes and encompasses 31 genera. Research has indicated a higher abundance of \textit{Negativicutes} class bacteria in smokers than non-smokers. However, no studies have offered insights into the mechanisms linking the \textit{Negativicutes} class to SVS. Future investigations employing experimental approaches are warranted to delve into the pathophysiological impacts of gut microbiome composition on stroke.

According to a few observational or preclinical studies, stroke can cause gut microbiome dysbiosis, leading to a neuroinflammatory response and exacerbating brain infarction. Our study found that stroke was associated with an increased abundance of opportunistic pathogens, such as \textit{Megasphaera}, at the genus level, which was consistent with previous observational studies in most species levels (species \textit{Megasphaera elsdenii} increased after AIS, species unclassified \textit{Megasphaera sp. UPII 135-E increased after LAS, but species \textit{Megasphaera micronuciformis} decreased after AIS). In addition, we found that stroke was related to the overgrowth of \textit{Bacteroides} at some species levels in the EAS but not in the EU population, which aligns with previous studies in mice models or populations of EAS ancestry. Insight into the brain–gut axis reveals potential mechanisms linking stroke, the gut microbiome and stroke recovery. After stroke onset, intestinal ischaemia and microbiome dysbiosis can lead to intestinal inflammation and the release of proinflammatory cells that enter the brain through the compromised blood-brain barrier, further exacerbating stroke and impairing brain recovery. However, our study did not establish new genetic causal relationships between stroke and gut microbiome in either the EAS or EU population, which suggests that the causal relationship between stroke onset and specific gut microbiome may not exist due to the complexity of the underlying biological mechanisms and unknown interactions within subgroups, or the effects may be too subtle to detect in the current study.

We incorporated several strengths into our study. First, we used the largest available GWAS summary data for both gut microbiome and stroke, which increased the statistical power and allowed us to explore potential
casual relationships. Moreover, we considered the cross-ancestry factor, which could affect not only genetics but also lifestyle factors such as diets and exercise, enabling us to capture population-specific associations between the gut microbiome and stroke. Finally, we took into account the role of stroke in gut microbiome dysbiosis, which helped to rule out reverse causality and elucidate the causal pathway simultaneously. However, our study also has limitations. First, unlike other traits, the heritability of gut microbiome taxa abundance is only 20% on average, and it exhibits high heterogeneity and inter-individual variability. Owning to these characteristics, we used a locus-wide significance level (pval=1e-5) instead of a genome-wide significance level (pval=5e-8) to select IVs. We calculated $F$ statistics to address weak instrumental bias, and all IVs had $F$ statistics greater than 10. Second, the species-level taxonomic resolution was challenging due to the limitations of the 16S rRNA analysis method and differences in gut microbiome assessment across cohorts. Therefore, we lack species-level taxonomic resolution in the EU population. Third, despite using the largest available cohort with sequencing data in EAS, the sample size of gut microbiome data remained relatively small compared with other traits, emphasising the need for a larger cohort with standardised protocols and sequencing data in future studies. Fourth, while the MR investigation conducted here provides a comprehensive view of the potential causal relationship between microbiome and stroke subtypes, the discovered causal relationship will require further validation to understand the nature of the observed statistical significance.

CONCLUSION

In conclusion, our MR study provides evidence supporting a causal relationship between the gut microbiome and stroke and identified several potential genetic causal relationships, as well as a promising protective effect of species *B. pectinophilus* on the onset of CES. Our findings strengthen the understanding of the causal relationship between the gut microbiome and stroke in ancestry-specific populations. Considering the growing clinical implantation of the gut microbiome, like faecal microbiome transplantation, in stroke-related areas, the identified causality between the gut microbiome and stroke may be valuable for developing a new stroke prevention strategy and warrants further mechanistic validation and clinical trial studies.

Author affiliations

1 School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen 518107, Guangdong, China
2 China National Clinical Research Center for Neurological Diseases, Beijing 100070, China
3 Changping Laboratory, Beijing 100000, China
4 Center of excellence for Omics Research (CORe), Beijing Tiantan Hospital, Capital Medical University, Beijing 100070, China
5 Department of Neurology, Beijing Tiantan Hospital, Capital Medical University, Beijing 100070, China
6 Clinical Center for Precision Medicine in Stroke, Capital Medical University, Beijing 100069, China

Contributors Study was conceived and designed by SL, ShiyaoC, SiC and HL. GWAS summary data collecting and IVs selection were conducted by ShiyaoC, YG and XL. MR analysis was done by ShiyaoC and YG. Manuscript was drafted by ShiyaoC and XG. Manuscript was revised by SL, ShiyaoC, XL, SiC, ZF and HZ. Responsible for the overall content as the guarantor: SL.

Funding This study is supported by grants from National Key R&D Program of China (2022YFC2002400, 2022YFC2002402, 2022YFE0209600), the National Natural Science Foundation of China (82101359, 82111530203, U20A20358, 31900487), Capital’s Funds for Health Improvement and Research (2020-1-2041), and Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2019-2SMR-5-029).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval The Ethics Committee(s) or Institutional Board(s) involving in this studies have been written and done by the authors of the previous GWAS work. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. All data relevant to the study are included in the article, serving as online supplemental information or uploaded on Github available in a public, open-access repository (https://github.com/lusylab/GutMicrobiome_STROKE_2SMR/).

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID ids Yuandan Wei http://orcid.org/0000-0009-6433-3040
Hao Li http://orcid.org/0000-0002-8591-4105
Si Cheng http://orcid.org/0000-0003-0551-1837
Siyang Liu http://orcid.org/0000-0001-6780-9419

REFERENCES
