Pharmacogenomics and Personalized Medicine

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Role of *CYP19A1* Loci (rs28757157 and rs3751591) with Ischemic Stroke Risk in the Chinese Han Population

Kang Huang[1], Tianyi Ma[1], Qiang Li[1], Zanrui Zhong[1], Ting Qin[1], Yilei Zhou[2], Wei Zhang[1], Shilin Tang[1], Jianghua Zhong[1], Shijuan Lu[1]

[1]Department of Cardiology, Haikou People's Hospital, Central South University Xiangya School of Medicine Affiliated Haikou Hospital, Haikou, Hainan, 570100, People's Republic of China; [2]Jingchu University of Technology, School of Medicine, Jingmen, Hubei, 448000, People's Republic of China

Correspondence: Shijuan Lu, Department of Cardiology, Haikou People's Hospital, Central South University Xiangya School of Medicine Affiliated Haikou Hospital, No. 43, Renmin Avenue, Haidian Island, Haikou, 570100, Hainan, People's Republic of China, Tel +86-13005013999, Email Lu_shijuan@163.com; Jianghua Zhong, Jingchu University of Technology, School of Medicine, No. 33, Xiangshan Avenue, Jingmen, Hubei, 448000, People's Republic of China, Tel +86-13322003882, Email zhong3882@163.com

Introduction: Ischemic stroke (IS) is a multifactorial and polygenic disease, which is affected by genetic factors. In this study, we explored the role of *CYP19A1* single nucleotide polymorphisms (SNPs) in IS in the Chinese population.

Methods: 1302 subjects (651 controls and 651 cases) were recruited in this case–control study. Four candidate SNPs (rs28757157 C/T, rs3751592 C/T, rs3751591 G/A, rs59429575 C/T) of *CYP19A1* were selected by the 1000 genomes project database. The association between *CYP19A1* SNPs and IS risk was assessed using logistic regression analysis with odds ratio (OR) and 95% confidence intervals (CIs). False-positive report probability (FPRP) analysis further verified the positive results. The interaction of SNP-SNP was analyzed by multi-factor dimensionality reduction (MDR) to predict is risk.

Results: In the research, *CYP19A1* loci (rs28757157 and rs3751591) were associated with the occurrence of IS. The two variants conferred an increased susceptibility to IS in the subjects aged over 60 years old, smokers and drinkers. Rs28757157 was related to the risk of IS in females, non-smokers and subjects with BMI less than 24, while rs59429575 was related to the risk of IS in males and subjects with BMI greater than 24.

Conclusion: The study revealed that there is a significant association between *CYP19A1* loci (rs28757157 and rs3751591) and IS risk in the Chinese Han population, providing a theoretical basis for further exploring its specific role in the pathogenesis of IS.

Keywords: ischemic stroke, risk, *CYP19A1*, single nucleotide polymorphisms, case-control study, Chinese Han population

Introduction

Stroke is a multifactorial and polygenic disease, which is known to be the main cause of death and adult disability.[1] In general, most stroke patients cannot fully recover in the later stage. In severe cases, hemiplegia, lack of living ability, and inability to take care of themselves will occur, which brings great mental and economic burdens to their families. Stroke is divided into ischemic and hemorrhagic, of which 85% of stroke patients are due to ischemic.[2] Moreover, hemorrhagic stroke may occur in ischemic brain, but it is not clear how much bleeding begins from ischemic stroke (IS). In China, stroke has become the leading cause of death.[3] Up to 2013, stroke was the leading cause of death in 27 out of 33 provinces in China.[4] The incidence of stroke in China is increasing year by year, but its pathogenesis is complex and has not yet been elucidated. Through recent years of research, factors such as hypertension,[5] diabetes,[6] obesity,[7] hypercholesterolemia,[8] smoking,[9] and alcohol consumption[10] were found to be likely to influence IS burden in China. However, numerous scientific studies have shown that IS was highly influenced by gene polymorphisms. Therefore, loci should also be paid attention to in the study of the pathogenesis of IS.

Received: 9 January 2023
Accepted: 11 April 2023
Published: 29 May 2023

Cytochrome P450s are important oxidative metabolic enzymes that play a well-established role in the metabolism of endogenous compounds and the metabolic clearance of drugs and other exogenous substances.[11] Cytochrome P450s enzymes can catalyze the metabolism of arachidonic acid into epoxyeicosatrienoic acids, dihydroxyeicosatetraenoic acids, and hydroxyeicosahexaenoic acids (HETEs). Advances in many disciplines have clarified the emerging role of cytochrome P450 enzymes and their metabolic substrates and end-products in the pathogenesis of central nervous system diseases, including Alzheimer's disease, Parkinson's disease, and stroke.[12] Now, studies have shown that CYP candidate genes and variant loci played a non-negligible role in the pathogenesis of IS.[13,14] Cytochrome P450 family 19 subfamily A member 1 (CYP19A1), one of the members of cytochrome P450 genes, which encodes an aromatase enzyme that can catalyze the conversion of C19 androgen, androstenedione, and testosterone to C18 estrogen, estrone, and estradiol, respectively. In an article reported by Zheng et al,[15] *CYP19A1* variants were associated with Alzheimer's disease, but the mechanism of *CYP19A1* loci in IS remains unclear.

In this research, we used to explore the association between *CYP19A1* variants and the incidence of IS in the Chinese population by logistic-regression analysis with OR and 95%, to provide the theoretical basis for the role of *CYP19A1* in ischemic stroke.

Methods

Study Subjects

In the present study, we used G*power 3.1.9.7 software to estimate the sample size of the case group and the control group through an independent-sample *t*-test. The parameters were set as follows: Tail = 2, Effect size = 0.20, $\alpha = 0.05$, Power = 0.95, Allocation ratio N2/N1=1. Finally, 1302 subjects (651 controls and 651 cases) were recruited from Haikou People's Hospital. All participants were Han Chinese and signed informed consent. The study was approved by the ethics committee of Haikou People's Hospital.

Based on the World Health Organization's diagnostic criteria, all patients with IS underwent a series of diagnostic tests, including standardized blood tests, as well as computed tomography (CT) and/or magnetic resonance imaging (MRI). And each patient's diagnosis was confirmed by at least two independent neurologists. Patients with a history of transient ischemic attack, systemic inflammatory disease, or tumor were excluded from this study. In the controls, they were examined at the physical examination center of the hospital at the same time. They had no history of hypertension, diabetes, coronary heart disease, cerebrovascular disease, etc.

SNPs Selection and Genotyping

Based on the 1000-genome project (https://www.internationalgenome.org/) in the global population with minor allele frequencies (MAFs) >0.05 and Hardy-Weinberg equilibrium (HWE) >0.01, we selected variants of *CYP19A1* gene. In addition, we applied online software-Agena MassARRAY Assay Designer 3.1 to design amplification primers and extension primers (Supplementary Table 1). By removing non-specific primers and ineffective primers, we finally selected four loci (rs28757157, rs3751592, rs3751591 and rs59429575) on *CYP19A1* gene. Using our previously extracted genomic DNA as a template, we completed SNPs genotyping for *CYP19A1* gene on the basis of the Agena MassARRAY nucleic acid mass spectrometry platform. Data sorting and analysis of the genotyping results were finished by Agena software Bioscience TYPER, version 4.0.

Statistical Analyses

With the support of SPSS 19.0 and Microsoft Excel 22.0, we completed the data analysis. Among them, each SNP in the control group was detected using the chi-square test to compare the expected and actual genotype frequencies of the SNPs to test whether these loci were consistent with HWE. Subsequently, the allele and genotype frequencies of the selected SNPs were analyzed in the case group and the control group using the chi-square test. The association between *CYP19A1* polymorphisms and IS risk was assessed by logistic regression analysis with odds ratios (OR) and 95% confidence intervals (CIs). FPRP tested for significant positive results at the power level (OR = 1.5) and a prior probability level "0.25, 0.1, 0.01, 0.001, 0.0001". Moreover, MDR software predicted the correlation between

CYP19A1 polymorphisms and IS risk from the perspective of SNP interaction (synergy or antagonism). $p < 0.05$ indicated statistical significance.

Results

The Information of Study Subjects and Variants

Totally, 651 controls (437 males/214 females) and 651 cases (434 males/217 females) were randomly recruited in the study. In the Supplementary Table 2, the basic information of subjects' characteristics was shown in the table. The average age of the controls and cases was 60.55 ± 0.23 years old and 63.76 ± 0.39 years old, respectively. Although the groups were not matched in age ($p = 0.000$), they were matched in gender ($p = 0.860$). In addition, the average height, weight and BMI of the controls (168.51 ± 0.29, 65.43 ± 0.43 and 23.01 ± 0.13) and cases (168.56 ± 0.28, 65.65 ± 0.38 and 23.09 ± 0.12) were displayed in the table, respectively. Also, there was no significant difference in height, weight, BMI index, smoking status and drinking status between the two groups ($p = 0.910$, $p = 0.700$, $p = 0.633$, $p = 0.121$ and $p = 0.579$).

In Table 1, the basic information of the selected variants (rs28757157, rs3751592, rs3751591 and rs59429575) located in *CYP19A1*, including the chromosome, position, allele, minor allele frequency (MAF) in cases and controls, consequence, HWE *p*-value and the function of loci predicted by HaploReg v4.1.

The Association Between *CYP19A1* Polymorphisms and Ischemic Stroke Risk (Overall Analysis)

In Table 2 and Figure 1, we performed the correlation between *CYP19A1* polymorphisms and IS risk shown in the overall analysis. In the allele model, rs28757157C and rs59429575C were associated with an increased risk of IS (OR: 1.23, $p_{adj} = 0.030$; OR: 1.27, $p_{adj} = 0.020$). A significant association of the rs28757157 was also observed in the heterozygous C/T model with OR: 1.42 and $p_{adj} = 0.004$, the dominant model with OR: 1.39 and $p_{adj} = 0.005$ and the log-additive model with OR: 1.26 and $p_{adj} = 0.018$. Meanwhile, rs28757157 and rs59429575 were providing risk for IS in the codominant (OR: 1.42 and $p_{adj} = 0.004$; OR: 1.47 and $p_{adj} = 0.003$), dominant (OR: 1.39 and $p_{adj} = 0.005$; OR: 1.45 and $p_{adj} = 0.003$) and log-additive (OR: 1.26 and $p_{adj} = 0.018$; OR: 1.34 and $p_{adj} = 0.006$) models. Subsequently, FPRP analysis was completed to verify the significant above results in the allele model (rs28757157: Power = 0.982, FPRP values = 0.080, 0.206; rs59429575: Power = 0.944, FPRP values = 0.067, 0.178), the heterozygous model (rs28757157: Power = 0.671, FPRP values = 0.020, 0.058; rs59429575: Power = 0.564, FPRP values = 0.011, 0.033) and the dominant model (Power = 0.747, FPRP values = 0.016, 0.047; Power = 0.610, FPRP values = 0.011, 0.032) at the prior probabilities are 0.25 and 0.1 (Supplementary Table 3).

Association Between *CYP19A1* Polymorphisms and Ischemic Stroke Risk Stratified by Age and Gender

In Table 3, we did the age and gender stratified analysis to assess the association between *CYP19A1* polymorphisms and IS risk. In people older than 60 years old, the variant-rs28757157 showed a significant association with IS risk with OR:

Table 1 The Basic Information of the Selected Variants Located in *CYP19A1*

SNP-ID	Gene	Chr: Position	Consequence	Allele	MAF		HWE *p*-value	HaploReg
					Case	Control		
rs28757157	CYP19A1	15:51253204	Intron_variant	C/T	0.241	0.206	0.402	Motifs changed, Selected eQTL hits
rs3751592	CYP19A1	15:51314381	2KB_upstream_variant	C/T	0.137	0.118	0.851	DNAse, Motifs changed, Selected eQTL hits
rs3751591	CYP19A1	15:51314513	2KB_upstream_variant	G/A	0.143	0.143	0.202	DNAse, Motifs changed
rs59429575	CYP19A1	15:51314880	2KB_upstream_variant	C/T	0.190	0.156	0.370	DNAse, Proteins bound, Motifs changed

Notes: *p*-value was calculated from Person's chi-square test.
Abbreviations: SNP, single nucleotide polymorphism; MAF, minor allele frequency; HWE, Hardy–Weinberg equilibrium; A, minor alleles; B, major alleles.

Table 2 The Association Between *CYP19A1* Polymorphisms and Ischemic Stroke Risk (Overall Analysis)

SNP-ID	Model	Genotype	Frequency		With Adjustment	
			Case	Control	OR (95% CI)	*p*-value
rs28757157	Allele	C	988	1034	1	
		T	314	268	1.23 (1.02–1.48)	**0.030**
	Codominant	C/C	369	414	1	
		C/T	250	206	1.42 (1.12–1.81)	**0.004**
		T/T	32	31	1.16 (0.69–1.96)	0.575
	Dominant	C/C	369	414	1	
		C/T-T/T	282	237	1.39 (1.10–1.74)	**0.005**
	Recessive	C/C-C/T	619	620	1	
		T/T	32	31	1.02 (0.61–1.71)	0.935
	Log-additive	–	–	–	1.26 (1.04–1.53)	**0.018**
rs3751592	Allele	T	1124	1149	1	
		C	178	153	1.19 (0.94–1.50)	0.141
	Codominant	T/T	489	506	1	
		C/T	146	137	1.14 (0.87–1.49)	0.359
		C/C	16	18	2.15 (0.89–5.19)	0.088
	Dominant	T/T	489	506	1	
		C/T-C/C	162	145	1.19 (0.92–1.55)	0.192
	Recessive	C/T-T/T	635	643	1	
		C/C	16	8	2.10 (0.87–5.05)	0.099
	Log-additive	–	–	–	1.22 (0.96–1.54)	0.102
rs3751591	Allele	A	1116	1113	1	
		G	186	185	1.00 (0.80–1.25)	0.981
	Codominant	A/A	477	473	1	
		G/A	162	167	0.98 (0.76–1.27)	0.879
		G/G	12	9	1.56 (0.64–3.80)	0.333
	Dominant	A/A	477	473	1	
		G/G-G/A	174	176	1.01 (0.78–1.30)	0.952
	Recessive	G/A-A/A	639	640	1	
		G/G	12	9	1.56 (0.64–3.81)	0.326
	Log-additive	–			1.04 (0.83–1.30)	0.756
rs59429575	Allele	C	1054	1099	1	
		T	248	203	1.27 (1.04–1.56)	**0.020**
	Codominant	C/C	425	467	1	
		C/T	204	165	1.47 (1.14–1.88)	**0.003**
		T/T	22	19	1.32 (0.70–2.49)	0.397
	Dominant	C/C	425	467	1	
		T/T-C/T	226	184	1.45 (1.14–1.84)	**0.003**
	Recessive	C/T-C/C	629	632	1	
		T/T	22	19	1.18 (0.63–2.22)	0.608
	Log-additive	–	–	–	1.34 (1.09–1.65)	**0.006**

Notes: *p*-value was calculated by logistic regression analysis with adjustments for age and gender. Bold values indicated that the *p*-value was statistically significant.

Abbreviations: SNP, single nucleotide polymorphism; OR, odds ratio; 95% CI, 95% confidence interval.

1.34 and p_{adj} = 0.023 (allele), OR: 1.46 and p_{adj} = 0.029 (codominant), OR: 1.45 and p_{adj} = 0.028 (dominant). As for rs59429575, an increased association of the variant was also observed in the dominant model with OR: 1.43 and p_{adj} = 0.043 and the log-additive model with OR: 1.46 and p_{adj} = 0.034. The two variants were not found to be associated with IS risk in people younger than 60 years old.

In the results of gender stratification, the allelic association of rs28757157 showed a significant association with IS risk in females, with an observed OR of 1.46 (1.02–1.98 at 95% CI) and p_{adj} = 0.038 (allele). In the codominant and dominant models, the

SNP-ID	Model	Genotype	Case	Control	OR (95% CI)	*p*-value
rs28757157	allele	C	988	1034	1	
		T	314	268	1.23(1.02-1.48)	**0.030**
	codominant	C/C	369	414	1	
		C/T	250	206	1.42(1.12-1.81)	**0.004**
		T/T	32	31	1.16(0.69-1.96)	0.575
	dominant	C/C	369	414	1	
		C/T-T/T	282	237	1.39(1.10-1.74)	**0.005**
	recessive	C/C-C/T	619	620	1	
		T/T	32	31	1.02(0.61-1.71)	0.935
	log-additive	–	–	–	1.26(1.04-1.53)	**0.018**
rs3751592	allele	T	1124	1149	1	
		C	178	153	1.19(0.94-1.50)	0.141
	codominant	T/T	489	506	1	
		C/T	146	137	1.14(0.87-1.49)	0.359
		C/C	16	8	2.15(0.89-5.19)	0.088
	dominant	T/T	489	506	1	
		C/T-C/C	162	145	1.19(0.92-1.55)	0.192
	recessive	C/T-T/T	635	643	1	
		C/C	16	8	2.10(0.87-5.05)	0.099
	log-additive	–	–	–	1.22(0.96-1.54)	0.102
rs3751591	allele	A	1116	1113	1	
		G	186	185	1.00(0.80-1.25)	0.981
	codominant	A/A	477	473	1	
		G/A	162	167	0.98(0.76-1.27)	0.879
		G/G	12	9	1.56(0.64-3.80)	0.333
	dominant	A/A	477	473	1	
		G/G-G/A	174	176	1.01(0.78-1.30)	0.952
	recessive	G/A-A/A	639	640	1	
		G/G	12	9	1.56(0.64-3.81)	0.326
	log-additive	–	–	–	1.04(0.83-1.30)	0.756
rs59429575	allele	C	1054	1099	1	
		T	248	203	1.27(1.04-1.56)	**0.020**
	codominant	C/C	425	467	1	
		C/T	204	165	1.47(1.14-1.88)	**0.003**
		T/T	22	19	1.32(0.70-2.49)	0.397
	dominant	C/C	425	467	1	
		T/T-C/T	226	184	1.45(1.14-1.84)	**0.003**
	recessive	C/T-C/C	629	632	1	
		T/T	22	19	1.18(0.63-2.22)	0.608
	log-additive	–	–	–	1.34(1.09-1.65)	**0.006**

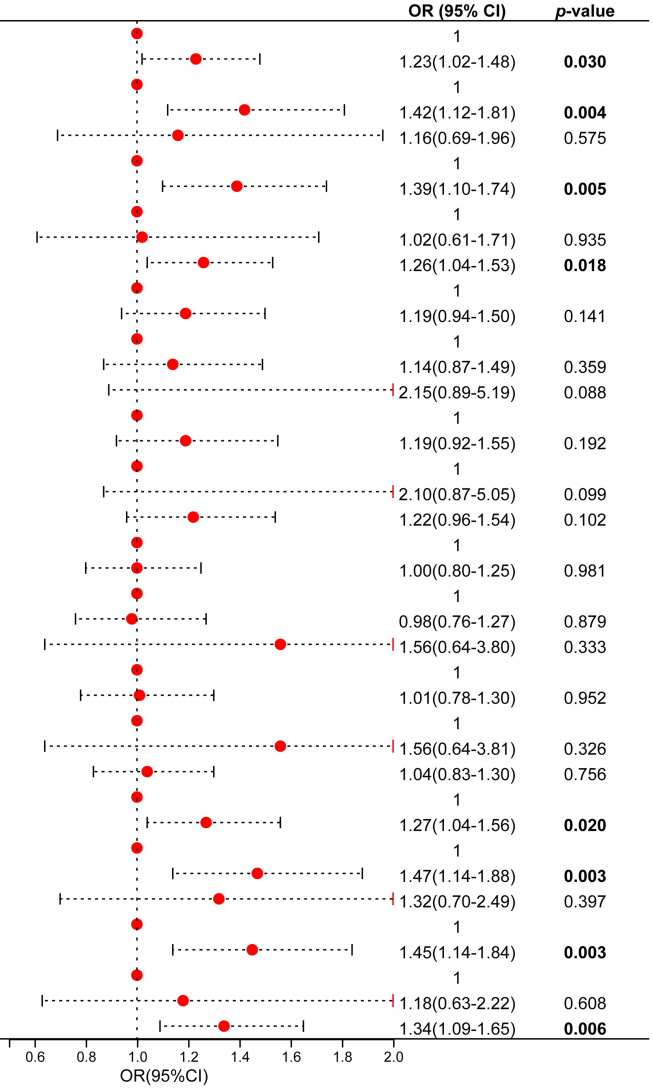

Figure 1 The forest map showed the association between CYP19A1 polymorphisms and ischemic stroke risk (overall analysis).

locus provided an increased risk for IS in females (codominant: OR: 1.70, p_{adj} = 0.015; dominant: OR: 1.62, p_{adj} = 0.020). In males, a significant association between rs59429575 and IS risk were found in the allele (OR: 1.33, p_{adj} = 0.025), codominant (OR: 1.69, p_{adj} = 0.001), dominant (OR: 1.64, p_{adj} = 0.001) and log-additive (OR: 1.44, p_{adj} = 0.006) models.

Association Between *CYP19A1* Polymorphisms and Ischemic Stroke Risk Stratified by BMI

In Table 4, BMI stratification analysis was also completed to evaluate the association between *CYP19A1* loci and IS risk. In subjects with BMI greater than 24, variant-rs59429575 was found to be associated with IS risk in the allele (OR: 1.45, p_{adj} = 0.041), codominant (OR: 1.87, p_{adj} = 0.008), dominant (OR: 1.84, p_{adj} = 0.007) and log-additive (OR: 1.59, p_{adj} = 0.014) models. In subjects with BMI less than 24, the locus showed a significant association with IS risk in the allele model with OR: 1.27 and p_{adj} = 0.035, the codominant model with OR: 1.45 and p_{adj} = 0.012, the dominant model with OR: 1.41 and p_{adj} = 0.015, the log-additive model with OR: 1.29 and p_{adj} = 0.037.

Association Between *CYP19A1* Polymorphisms and Ischemic Stroke Risk Stratified by Smoking and Drinking

In Table 5, smoking and drinking stratification analysis was also completed to evaluate the association between *CYP19A1* loci and IS risk. In the results of smoking stratification, variant rs28757157 was found to be significantly associated with

Table 3 The Association Between *CYP19A1* Polymorphisms and Ischemic Stroke Risk (Age and Gender Stratified Analysis)

SNP	Model	Genotype	>60		≤60		Female		Male	
			OR (95% CI)	*p*-value	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value
rs28757157	Allele	C	1		1		1		1	
		T	1.34 (1.04–1.73)	**0.023**	1.15 (0.87–1.52)	0.315	1.42 (1.02–1.98)	**0.038**	1.15 (0.92–1.44)	0.224
	Codominant	C/C	1		1		1		1	
		C/T	1.46 (1.04–2.06)	**0.029**	1.27 (0.88–1.84)	0.208	1.70 (1.11–2.62)	**0.015**	1.29 (0.97–1.72)	0.084
		T/T	1.33 (0.62–2.84)	0.467	1.04 (0.45–2.41)	0.925	1.20 (0.51–2.87)	0.675	1.14 (0.59–2.21)	0.703
	Dominant	C/C	1		1		1		1	
		C/T-T/T	1.45 (1.04–2.00)	**0.028**	1.24 (0.87–1.78)	0.238	1.62 (1.08–2.43)	**0.020**	1.27 (0.96–1.68)	0.091
	Recessive	C/C-C/T	1		1		1		1	
		T/T	1.16 (0.55–2.47)	0.691	0.95 (0.41–2.16)	0.896	1.01 (0.43–2.39)	0.977	1.03 (0.54–1.98)	0.928
	Log-additive	–	1.32 (1.00–1.74)	0.048	1.16 (0.86–1.56)	0.346	1.38 (0.99–1.92)	0.059	1.19 (0.94–1.51)	0.144
rs59429575	Allele	C	1		1		1		1	
		T	1.28 (0.97–1.70)	0.082	1.34 (0.99–1.82)	0.057	1.17 (0.83–1.65)	0.382	1.33 (1.04–1.72)	**0.025**
	Codominant	C/C	1		1		1		1	
		C/T	1.38 (0.97–1.98)	0.076	1.45 (0.99–2.14)	0.060	1.20 (0.78–1.84)	0.402	1.69 (1.24–2.32)	**0.001**
		T/T	1.96 (0.74–5.20)	0.176	1.14 (0.44–2.97)	0.788	1.88 (0.59–6.02)	0.289	1.23 (0.57–2.68)	0.596
	Dominant	C/C	1		1		1		1	
		T/T-C/T	1.43 (1.01–2.03)	**0.043**	1.42 (0.97–2.06)	0.069	1.25 (0.83–1.89)	0.291	1.64 (1.21–2.21)	**0.001**
	Recessive	C/T-C/C	1		1		1		1	
		T/T	1.78 (0.68–4.71)	0.242	1.02 (0.39–2.63)	0.973	1.78 (0.56–5.64)	0.331	1.06 (0.49–2.29)	0.878
	Log-additive	–	1.46 (1.04–2.06)	**0.034**	1.29 (0.94–1.77)	0.118	1.25 (0.87–1.80)	0.218	1.44 (1.11–1.87)	**0.006**

Notes: *p*-value was calculated by logistic regression analysis with adjustments for age and gender. Bold values indicated that the *p*-value was statistically significant.
Abbreviations: SNP, single nucleotide polymorphism; OR, odds ratio; 95% CI, 95% confidence interval.

Table 4 The Association Between *CYP19A1* Polymorphisms and Ischemic Stroke Risk (BMI Stratified Analysis)

SNP	Model	Genotype	≥24		<24	
			OR (95% CI)	*p*-value	OR (95% CI)	*p*-value
rs28757157	Allele	C	1		1	
		T	1.13 (0.81–1.56)	0.476	1.27 (1.02–1.60)	**0.035**
	Codominant	C/C	1		1	
		C/T	1.36 (0.87–2.12)	0.176	1.45 (1.09–1.93)	**0.012**
		T/T	1.26 (0.54–2.95)	0.595	1.16 (0.59–2.29)	0.672
	Dominant	C/C	1		1	
		C/T-T/T	1.34 (0.88–2.04)	0.169	1.41 (1.07–1.87)	**0.015**
	Recessive	C/C-C/T	1		1	
		T/T	1.14 (0.49–2.63)	0.757	1.00 (0.51–1.96)	0.996
	Log-additive	–	1.23 (0.88–1.71)	0.225	1.29 (1.02–1.63)	**0.037**
rs59429575	Allele	C	1		1	
		T	1.45 (1.02–2.07)	**0.041**	1.19 (0.93–1.53)	0.162
	Codominant	C/C	1		1	
		C/T	1.87 (1.18–2.98)	**0.008**	1.33 (0.98–1.79)	0.067
		T/T	1.63 (0.58–4.61)	0.358	1.19 (0.52–2.71)	0.688
	Dominant	C/C	1		1	
		T/T-C/T	1.84 (1.18–2.86)	**0.007**	1.31 (0.98–1.76)	0.067
	Recessive	C/T-C/C	1		1	
		T/T	1.37 (0.49–3.85)	0.548	1.09 (0.48–2.48)	0.837
	Log-additive	–	1.59 (1.10–2.30)	**0.014**	1.24 (0.96–1.61)	0.096

Notes: *p*-value was calculated by logistic regression analysis with adjustments for age and gender. Bold values indicated that the *p*-value was statistically significant.

Abbreviations: SNP, single nucleotide polymorphism; OR, odds ratio; 95% CI, 95% confidence interval.

IS risk in smokers and non-smokers. Meanwhile, a significant association was observed between rs59429575 and IS risk in the allele (OR: 1.48, p_{adj} = 0.011), codominant (OR: 1.97, p_{adj} = 0.000), dominant (OR: 1.94, p_{adj} = 0.000) and log-additive (OR: 1.70, p_{adj} = 0.001) models.

In the results of drinking stratification, rs28757157 was related to the risk of IS (allele C: OR: 1.36, p_{adj} = 0.026; Heterozygous C/T: OR: 1.36, p_{adj} = 0.026; dominant C/T-C/C: OR: 1.45, p_{adj} = 0.025; log-additive: OR: 1.35, p_{adj} = 0.042) in drinkers. The variant-rs59429575 provided an increased risk for IS in the allele (OR: 1.48, p_{adj} = 0.010), codominant (OR: 1.87, p_{adj} = 0.001), dominant (OR: 1.87, p_{adj} = 0.001) and log-additive (OR: 1.65, p_{adj} = 0.002) models. However, the two variants were not found to be associated with IS risk in drinkers and non-drinkers.

Association Between *CYP19A1* Polymorphisms and Ischemic Stroke Risk (Clinical Information Analysis)

In Table 6, we further performed comparative analyses of hypertensive patients versus non-hypertensive patients, coronary heart disease patients versus non-coronary heart disease patients, and diabetic patients versus non-diabetic patients in the case group. In the patients with and without diabetes, we found that rs28757157 was associated with an increased risk of IS in the allele (OR: 0.66, p_{adj} = 0.025), dominant (OR: 0.65, p_{adj} = 0.047) and log-additive (OR: 0.65, p_{adj} = 0.023) models. However, the loci did not show a significant association with IS risk in patients with hypertension or coronary heart disease.

SNP-SNP Interaction Models of Candidate SNPs Analyzed by the MDR Software

Then, MDR software was used to analyze the impact of potential SNP–SNP interaction on IS risk (Supplementary Table 4). The three-locus model containing *CYP19A1* loci (rs28757157, rs3751591 and rs59429575) was regarded as the

Table 5 The Association Between *CYP19A1* Polymorphisms and Ischemic Stroke Risk (Smoking and Drinking Stratified Analysis)

SNP	Model	Genotype	Smoking		No-Smoking		Drinking		No-Drinking	
			OR (95% CI)	*p*-value	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value
rs28757157	Allele	C	1		1		1		1	
		T	1.15 (0.88–1.50)	0.293	1.31 (1.01–1.69)	**0.043**	1.36 (1.04–1.80)	**0.026**	1.13 (0.88–1.45)	0.345
	Codominant	C/C	1		1		1		1	
		C/T	1.27 (0.90–1.79)	**0.174**	1.50 (1.07–2.10)	**0.018**	1.36 (0.60–3.08)	**0.026**	1.27 (0.91–1.76)	0.155
		T/T	1.25 (0.58–2.68)	0.568	1.11 (0.53–2.32)	0.783	1.20 (0.51–2.87)	0.463	1.03 (0.51–2.08)	0.934
	Dominant	C/C	1		1		1		1	
		C/T-T/T	1.27 (0.91–1.76)	0.159	1.45 (1.05–1.99)	**0.025**	1.48 (1.05–2.08)	**0.024**	1.24 (0.90–1.69)	0.188
	Recessive	C/C-C/T	1		1		1		1	
		T/T	1.15 (0.54–2.45)	0.718	0.95 (0.46–1.97)	0.896	1.19 (0.53–2.67)	0.679	0.94 (0.47–1.87)	0.860
	Log-additive	–	1.20 (0.91–1.58)	0.192	1.28 (0.98–1.68)	0.069	1.35 (1.01–1.80)	**0.042**	1.15 (0.88–1.48)	0.307
rs59429575	Allele	C	1		1		1		1	
		T	1.48 (1.09–1.99)	**0.011**	1.13 (0.86–1.50)	0.384	1.48 (1.10–2.01)	**0.010**	1.12 (0.85–1.48)	0.410
	Codominant	C/C	1		1		1		1	
		C/T	1.97 (1.35–2.88)	**0.000**	1.18 (0.84–1.67)	0.338	1.87 (1.27–2.74)	**0.001**	1.25 (0.89–1.75)	0.203
		T/T	1.71 (0.67–4.36)	0.264	0.95 (0.39–2.33)	0.910	1.89 (0.76–4.68)	0.171	0.91 (0.36–2.33)	0.848
	Dominant	C/C	1		1		1		1	
		T/T-C/T	1.94 (1.35–2.79)	**0.000**	1.16 (0.83–1.61)	0.385	1.87 (1.30–2.70)	**0.001**	1.21 (0.87–1.68)	0.248
	Recessive	C/T-C/C	1		1		1		1	
		T/T	1.42 (0.56–3.60)	0.460	0.90 (0.37–2.20)	0.819	1.59 (0.64–3.90)	0.316	0.85 (0.34–2.17)	0.737
	Log-additive	–	1.70 (1.24–2.33)	**0.001**	1.11 (0.83–1.47)	0.497	1.65 (1.21–2.26)	**0.002**	1.14 (0.86–1.53)	0.362

Notes: *p*-value was calculated by logistic regression analysis with adjustments for age and gender. Bold values indicated that the *p*-value was statistically significant.
Abbreviations: SNP, single nucleotide polymorphism; OR, odds ratio; 95% CI, 95% confidence interval.

Table 6 The Association Between *CYP19A1* Polymorphisms and Ischemic Stroke Risk (Clinical Information Analysis)

SNP	Model	Genotype	Hypertension vs Non-Hypertension		Coronary Heart Disease vs Non-Coronary Heart Disease		Diabetes vs Non-Diabetes	
			OR (95% CI)	*p*-value	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value
rs28757157	Allele	C	1		1		1	
		T	0.86 (0.66–1.13)	0.291	1.03 (0.75–1.42)	0.834	0.66 (0.46–0.95)	**0.025**
	Codominant	C/C	1		1		1	
		C/T	0.76 (0.54–1.08)	0.128	1.04 (0.68–1.58)	0.856	0.71 (0.46–1.09)	0.117
		T/T	1.01 (0.45–2.26)	0.984	1.31 (0.55–3.12)	0.541	0.27 (0.06–1.14)	0.075
	Dominant	C/C	1		1		1	
		C/T-T/T	0.79 (0.56–1.10)	0.164	1.07 (0.72–1.60)	0.737	0.65 (0.42–0.99)	**0.047**
	Recessive	C/C-C/T	1		1		1	
		T/T	1.13 (0.51–2.49)	0.770	1.29 (0.55–3.02)	0.557	0.30 (0.07–1.29)	0.106
	Log-additive	–	0.86 (0.65–1.14)	0.289	1.09 (0.78–1.51)	0.619	0.65 (0.44–0.94)	**0.023**
rs59429575	Allele	C	1		1		1	
		T	0.99 (0.73–1.33)	0.930	1.08 (0.77–1.52)	0.662	0.81 (0.55–1.19)	0.289
	Codominant	C/C	1		1		1	
		C/T	0.99 (0.68–1.42)	0.942	1.33 (0.87–2.04)	0.193	0.98 (0.63–1.53)	0.939
		T/T	0.89 (0.35–2.23)	0.796	0.94 (0.31–2.92)	0.921	0.23 (0.03–1.74)	0.154
	Dominant	C/C	1		1		1	
		T/T-C/T	0.98 (0.69–1.39)	0.893	1.29 (0.85–1.95)	0.236	0.90 (0.58–1.39)	0.638
	Recessive	C/T-C/C	1		1		1	
		T/T	0.89 (0.35–2.23)	0.802	0.86 (0.28–2.63)	0.790	0.23 (0.03–1.74)	0.155
	Log-additive	–	0.97 (0.71–1.32)	0.843	1.18 (0.83–1.68)	0.355	0.83 (0.57–1.23)	0.362

Notes: *p*-value was calculated by logistic regression analysis with adjustments for age and gender. Bold values indicated that the *p*-value was statistically significant.
Abbreviations: SNP, single nucleotide polymorphism; OR, odds ratio; 95% CI, 95% confidence interval.

best model for predicting the impact of SNP–SNP interactions on IS risk (cross-validation consistency 10/10, testing balanced accuracy 54.07%, OR: 1.37 (1.33–2.07 at 95% CI), $p < 0.000$). In Figure 2, the interaction between rs28757157 and rs3751591 was synergistic, with the information gain value 0.35%.

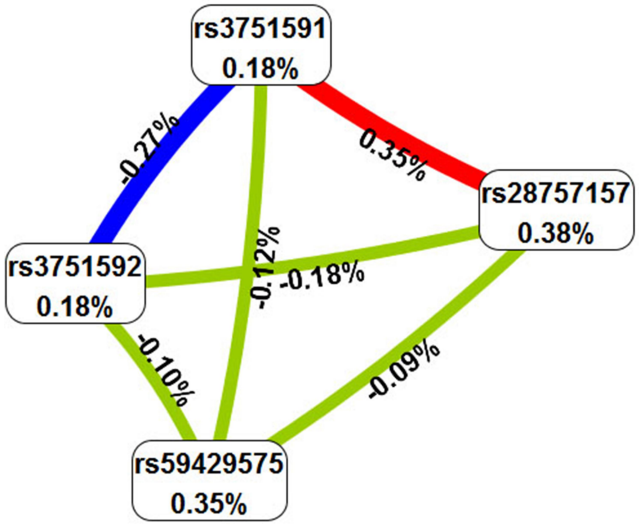

Figure 2 The circle graph showed that the interaction between rs28757157 and rs3751591 was synergistic, with the information gain value 0.35%.

Discussion

In the research, *CYP19A1* loci (rs28757157 and rs3751591) were associated with the occurrence of IS. The two variants conferred an increased susceptibility to IS in the subjects aged over 60 years old, smokers and drinkers. Rs28757157 was related to the risk of IS in females, non-smokers and subjects with BMI less than 24, while rs59429575 was related to the risk of IS in males and subjects with BMI greater than 24. The findings suggest that *CYP19A1* loci (rs28757157 and rs3751591) may affect IS risk in individuals from China.

Previous researchers have detected the expression of CYP450 aromatase in the human cerebral cortex,[16] and even found that it can regulate the metabolism of androgen and estrogen. For example, aromatase encoded by CYP19A1 played an important role in androgen metabolism.[17] Gaella et al[18] found that Cyp19a1 was post-transcriptionally upregulated in Celf1$^{-/-}$ testis, resulting in high aromatase activity. And excess aromatase partly caused defects in spermatogenesis. In addition, aromatase encoded by CYP19A1 converts testosterone to estradiol. There is increasing evidence that local estrogen synthesized by aromatase can also prevent cerebral ischemic injury. Among them, the aromatase encoded by CYP19A1 was thought to be beneficial to the injured brain.[12,19] Compared with young male rodents, young female rodents were less possibility to develop cerebral injury, because estrogen played a major role in neuroprotection.[20,21] Also, *CYP19A1* SNPs were also related to the occurrence of nervous system diseases, such as Alzheimer's disease[15] and astrocytoma.[22] In addition, Zhang et al[23] observed the relationship between *CYP19A1* loci (rs4646 and rs17601876) and a reduced risk of stroke. Although the above loci were not found to be associated with the risk of stroke in our study, we first found that *CYP19A1* loci (rs28757157 and rs3751591) conferred an increased susceptibility to IS.

Previous studies have found that the prevalence of stroke gradually increases with age and that stroke patients are more common in the elderly population.[24] Gender is another key factor in the occurrence of stroke. According to the subgroup analysis results of age, we observed that rs28757157 and rs59429575 showed an increased association with IS risk in people older than 60 years old. However, the two loci were not found to be related to IS stroke in people younger than 60 years old. Besides, our research showed that rs28757157 was related to the risk of IS in females, while rs59429575 was related to the risk of IS in males. Therefore, our study suggested that the effect of *CYP19A1* polymorphisms (rs28757157 and rs59429575) on IS risk was age- and sex-dependent, but the underlying mechanism remains unclear.

Smoking is a risk factor for ischemic stroke. Previous studies have shown a strong dose-dependent relationship between smoking and IS risk.[25] As cigarette consumption increased, the incidence of stroke increased linearly. According to the stratified analysis of smoking and alcohol consumption, *CYP19A1* loci (rs28757157 and rs3751591) were associated with the occurrence of IS in smokers. Combined with the results of gender stratification, we found that rs28757157 was related to the risk of is in females, while rs59429575 was related to the risk of is in males. We believed that this may be closely related to participant's living habits, such as smoking. With the change of people's life concept, drinking has become an indispensable way of communication in the process of interpersonal communication. Alcohol consumption is also the risk factor for ischemic stroke. Ischemic stroke risk increased with increasing frequency of alcohol consumption.[26] According to the stratified analysis of alcohol consumption, *CYP19A1* loci (rs28757157 and rs3751591) were associated with the occurrence of IS in drinkers. The above correlation obtained was dependent on drinking and smoking. Thus, this may be the reason why we strongly advocate people to actively quit smoking and alcohol.

In the present study, we observed that *CYP19A1* loci (rs28757157 and rs3751591) may be a risk factor for ischemic stroke. However, this study has some shortcomings. First, the selected samples are mainly from one hospital, and the sample size is not large. We further expand the sample size from different hospitals in the follow-up study to make the obtained results credible. Second, the role of *CYP19A1* and its sites in ischemic stroke is unknown. Molecular experiments need to be further explored to provide a theoretical basis for the mechanism of *CYP19A1* in ischemic stroke.

Data Sharing Statement

The data used to support the findings of this study are available from the corresponding author upon request.

Human and Animal Ethics

This study was approved by the ethics committee of Haikou People's Hospital. All participants signed informed consent forms before participating in this study.

Ethics Approval and Consent to Participate

This study was approved by the ethics committee of Haikou People's Hospital and conformed to the ethical principles of the Declaration of Helsinki. All participants signed informed consent forms before participating in this study.

Acknowledgments

We thank all authors for their contributions and supports. We are also grateful to all subjects for providing blood samples. Kang Huang and Tianyi Ma are the co-first authors. Jianghua Zhong and Shijuan Lu are the co-corresponding authors.

Funding

The study was supported by the specific research fund of The Innovation Platform for Academicians of Hainan Province (YSPTZX202032).

Disclosure

The authors have no conflict of interest.

References

1. Feigin VL, Forouzanfar MH, Krishnamurthi R, et al. Global and regional burden of stroke during 1990–2010: findings from the Global Burden of Disease Study 2010. *Lancet*. 2014;383(9913):245–254. doi:10.1016/S0140-6736(13)61953-4
2. Della-Morte D, Guadagni F, Palmirotta R, et al. Genetics of ischemic stroke, stroke-related risk factors, stroke precursors and treatments. *Pharmacogenomics*. 2012;13(5):595–613. doi:10.2217/pgs.12.14
3. Liu L, Wang D, Wong KS, Wang Y. Stroke and stroke care in China: huge burden, significant workload, and a national priority. *Stroke*. 2011;42(12):3651–3654. doi:10.1161/STROKEAHA.111.635755
4. Zhou M, Wang H, Zhu J, et al. Cause-specific mortality for 240 causes in China during 1990–2013: a systematic subnational analysis for the Global Burden of Disease Study 2013. *Lancet*. 2016;387(10015):251–272. doi:10.1016/S0140-6736(15)00551-6
5. Li W, Gu H, Teo KK, et al. Hypertension prevalence, awareness, treatment, and control in 115 rural and urban communities involving 47,000 people from China. *J Hypertens*. 2016;34(1):39–46. doi:10.1097/HJH.0000000000000745
6. Xu Y, Wang L, He J, et al. Prevalence and control of diabetes in Chinese adults. *JAMA*. 2013;310(9):948–959. doi:10.1001/jama.2013.168118
7. Zhang YX, Wang ZX, Zhao JS, Chu ZH. Trends in overweight and obesity among rural children and adolescents from 1985 to 2014 in Shandong, China. *Eur J Prev Cardiol*. 2016;23(12):1314–1320. doi:10.1177/2047487316643830
8. Barkas F, Elisaf M, Milionis H. Statins decrease the risk of stroke in individuals with heterozygous familial hypercholesterolemia: a systematic review and meta-analysis. *Atherosclerosis*. 2015;243(1):60–64. doi:10.1016/j.atherosclerosis.2015.08.038
9. Han J, Chen X. A meta-analysis of cigarette smoking prevalence among adolescents in China: 1981–2010. *Int J Environ Res Public Health*. 2015;12(5):4617–4630. doi:10.3390/ijerph120504617
10. Li C, Mao Z, Yu C. The effects of smoking, regular drinking, and unhealthy weight on health care utilization in China. *BMC Public Health*. 2021;21(1):2268. doi:10.1186/s12889-021-12309-z
11. Kelly E, Nakano M, Rohatgi P, Yarov-Yarovoy V, Rettie A. Finding homes for orphan cytochrome P450s: CYP4V2 and CYP4F22 in disease states. *Mol Interv*. 2011;11(2):124–132. doi:10.1124/mi.11.2.10
12. Liu M, Hurn PD, Alkayed NJ. Cytochrome P450 in neurological disease. *Curr Drug Metab*. 2004;5(3):225–234. doi:10.2174/1389200043335540
13. Yi XY, Liao DX, Wang C, Cheng W, Fu XQ, Zhang B. Cytochrome P450 genetic variants and their metabolite levels associated with plaque stability in patients with ischemic stroke. *J Atheroscler Thromb*. 2016;23(3):330–338. doi:10.5551/jat.31120
14. Li C, Jia W, Li J, Li F, Ma J, Zhou L. Association with CYP2C19 polymorphisms and Clopidogrel in treatment of elderly stroke patients. *BMC Neurol*. 2021;21(1):104. doi:10.1186/s12883-021-02127-6
15. Zheng J, Yan H, Shi L, et al. The CYP19A1 rs3751592 variant confers susceptibility to Alzheimer disease in the Chinese Han population. *Medicine*. 2016;95(35):e4742. doi:10.1097/MD.0000000000004742
16. Yague JG, Muñoz A, de Monasterio-Schrader P, Defelipe J, Garcia-Segura LM, Azcoitia I. Aromatase expression in the human temporal cortex. *Neuroscience*. 2006;138(2):389–401. doi:10.1016/j.neuroscience.2005.11.054
17. Boon WC, Chow JD, Simpson ER. The multiple roles of estrogens and the enzyme aromatase. *Prog Brain Res*. 2010;181:209–232.
18. Boulanger G, Cibois M, Viet J, et al. Hypogonadism associated with Cyp19a1 (Aromatase) posttranscriptional upregulation in Celf1 knockout mice. *Mol Cell Biol*. 2015;35(18):3244–3253. doi:10.1128/MCB.00074-15
19. Garcia-Segura LM, Veiga S, Sierra A, Melcangi RC, Azcoitia I. Aromatase: a neuroprotective enzyme. *Prog Neurobiol*. 2003;71(1):31–41. doi:10.1016/j.pneurobio.2003.09.005
20. Liu M, Kelley MH, Herson PS, Hurn PD. Neuroprotection of sex steroids. *Minerva Endocrinol*. 2010;35(2):127–143.
21. Manwani B, McCullough LD. Sexual dimorphism in ischemic stroke: lessons from the laboratory. *Womens Health*. 2011;7(3):319–339. doi:10.2217/whe.11.22

22. Wang T, Sun Y, Xiong Z, et al. Association of ST6GAL1 and CYP19A1 polymorphisms in the 3'-UTR with astrocytoma risk and prognosis in a Chinese Han population. *BMC Cancer*. 2021;21(1):391. doi:10.1186/s12885-021-08110-1
23. Cai Q, Zheng J, Bai M, et al. Genetic variations of CYP19A1 gene and stroke susceptibility: a case-control study in the Chinese Han population. *J Cardiovasc Pharmacol*. 2020;75(4):344–350. doi:10.1097/FJC.0000000000000793
24. Centers for Disease Control and Prevention. Prevalence of stroke–United States, 2006–2010. *MMWR Morb Mortal Wkly Rep*. 2012;61 (20):379–382.
25. Bhat VM, Cole JW, Sorkin JD, et al. Dose-response relationship between cigarette smoking and risk of ischemic stroke in young women. *Stroke*. 2008;39(9):2439–2443. doi:10.1161/STROKEAHA.107.510073
26. Cho IY, Yoo JE, Han K, et al. Frequent drinking is more predictive of ischemic stroke than binge drinking, but not of myocardial infarction. *Atherosclerosis*. 2022;350:65–72. doi:10.1016/j.atherosclerosis.2022.04.027