

## Original Article

# Genetic predisposition of stroke: understanding the evolving landscape through meta-analysis

Hai-Dong Huang<sup>1\*</sup>, Chun-Min Yang<sup>2\*</sup>, Hai-Feng Shu<sup>1</sup>, Yong-Qin Kuang<sup>1</sup>, Tao Yang<sup>1</sup>, Wei-Qi He<sup>1</sup>, Kai Zhao<sup>1</sup>, Xun Xia<sup>1</sup>, Jing-Min Cheng<sup>1</sup>, Yuan Ma<sup>1</sup>, Jian-Wen Gu<sup>1</sup>

Departments of <sup>1</sup>Neurosurgery, <sup>2</sup>Medical Imaging, General Hospital of The People's Liberation Army Chengdu Military Region, Sichuan 610083, China. \*Equal contributors.

Received September 25, 2014; Accepted December 18, 2014; Epub January 15, 2015; Published January 30, 2015

**Abstract:** Stroke, either ischemic or hemorrhagic, is the leading cause of death and morbidity worldwide. Identifying the risk factors is a prerequisite step for stroke prevention and treatment. It is believed that a major portion of the currently unidentified risk factors is of genetic origin. Consistent with this idea, numerous potential risk alleles for stroke have been reported, however, the genetic evidence so far is not conclusive. The major goal of this review is to update the current knowledge about the genetic predisposition to the common multifactorial stroke, and to provide a bird's-eye view of this fast moving field. We selectively review and meta-analyze the related English literatures in public domain (PubMed) from 2000 onward, including the original reports and meta-analyses, to evaluate the genetic risk factors of common multifactorial stroke. The results indicated that we reviewed and meta-analyzed original reports and existing meta-analyses that studied the genetic predisposition to the common multifactorial stroke. Some original reports and meta-analyses were specific for ischemic stroke and others were for hemorrhagic stroke only. We also evaluated the major evolving issues in this field and discussed the future directions. In conclusion, strong evidences suggest that genetic risk factors contribute to common multifactorial stroke, and many genetic risk genes have been implicated in the literatures. However, not a single risk allele has been conclusively approved.

**Keywords:** Genetic predisposition/risk, ischemic stroke, hemorrhagic stroke

## Introduction

A stroke, either ischemic or hemorrhagic, is commonly defined as a sudden loss of neurological function resulting from focal disturbance of cerebral blood flow [1], while World Health Organization defines stroke as "rapidly developing clinical signs of focal (or global) disturbance of cerebral function, with symptoms lasting 24 hours or longer or leading to death, with no apparent cause other than of vascular origin" [2]. Regardless the semantic differences of the definitions, it is commonly accept that stroke is itself a syndrome cause by a number of different disease processes, not a single homogeneous disease.

Numerous epidemiological studies have identified many risk factors for stroke [3, 4], including medical factors such as previous stroke, ischemic heart disease [5], atrial fibrillation [6], hypertension and glucose intolerance [7]. Other risk factors could be inherent biological

traits such as age and sex [8], physiological characteristics such as serum cholesterol [9], and fibrinogen [10]; behaviors such as smoking, diet, alcohol consumption and physical inactivity [11]; social characteristics such as education, social class and ethnicity; or geographical factors such as altitude and temperature [12, 13]. Overall, more than 300 risk factors have been associated with stroke, and these factors are generally classified as either modifiable (e.g., environmental and behavioral factors) or non-modifiable (e.g., genetic and age).

Many of the common modifiable risk factors of stroke have been identified [14], however, the identified factors so far explain only about 60% of the attributable risk, and up to 40% of stroke risk can be attributed to currently unknown factors [4]. A major portion of these unknown factors is believed to be of genetic origin. Consistently, evidence of genetic risk factors comes from numerous studies. For example, it has been noticed for centuries that stroke

tends to run in families [15, 16]. A three-fold increased incidence has been reported in the risk of stroke in men whose mothers had died of stroke in comparison with men without a maternal history of stroke [17]. It is also reported that stroke in a first-degree relative increases a subject's odds by as much as 2- to 6-fold [18]. The contribution of genetic factors also has been increasingly recognized from findings in twin studies [19]. In addition, a number of monogenic disorders cause stroke, such as cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) [20], even though common stroke is thought to be a polygenic multifactorial disease. Animal model studies also provided strong genetic evidence that certain lines of animals have predisposition to stroke [21, 22]. Nevertheless, the exact loci or alleles that are responsible for the common polygenic stroke risk are still largely unknown.

Numerous studies have been carried out to investigate the genetic risk genes, but many confounding factors contributed to the profound confusion in the field, such as, 1) some investigators chose to investigate only genetic risk for the hemorrhagic stroke (HS), but the majority of reports so far studied the genetic risk of ischemic stroke (IS); 2) the method of genetic risk factor identification evolved with technical progresses, 3) each study followed its own unique design and protocol, even its own classification. For these reasons, a single meta-analysis of the pooled data from different original studies is often not accurate because of the confounding effects.

Furthermore, specific meta-analyses that focused exclusively on certain specific risk alleles can't provide the overall picture of the field. Therefore, in this study, we try to address this issue through a combinational approach, i.e., through meta-analyzing both the existing original reports and meta-analyses. We argue that this approach may provide a balanced view and minimize the inherited confounding effects of single meta-analysis.

Our major intention is to integrate and update knowledge in the field and provide a better overall understanding of the evolving landscape of this fast moving field, not just on certain risk alleles. To accomplish this seemingly ambitious goal in a regular review, we have to dependent heavily on the existing meta-analyses. We argue that even though it is impractical to exhaustively re-analyze all the original re-

ports, it is possible to re-analyze the existing meta-analyses. This approach may also help to integrate the existing data and provide a better overview of the field.

The other focus of this review would be discussing the major evolving, sometimes thorny issues, in this field, such as classification and methodology. These issues have been noticed for a long period time, but have not been systematically dealt with. Current review hopefully will clarify the profound confusions and put the data in the right perspective. In the end, we will also briefly discuss the future opportunities. We hope this review may foster a healthier growth of the field.

### Materials and methods

#### *Data sources and search strategy*

The literature selection and systematic review was generally performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (<http://www.prisma-statement.org/statement.htm>). Search terms included: genetic predisposition/genetic risk factors, ischemic stroke, hemorrhagic stroke and meta-analyses. Only English reports published in PubMed in full manuscript form were selected.

#### *Inclusion criteria*

- 1) We know there may be over a hundred alleles have been implicated, but in this Meta-analysis, only the candidate alleles that have reasonable supports from literatures were selected.
- 2) For each selected candidate allele, only the solid recent reports, both original studies and the representative meta-analyses were selected.
- 3) From each report, we only extracted the number of participants (cases and controls), the odds ratios along with appropriate 95% CIs and *P* values.

#### *Statistical methods*

Odds ratio (OR) and 95% confidential interval (CIs) were used to assess the association between all kinds of risk alleles and common multifactorial stroke risk. Even though all subgroup analyses followed the same meta-analysis procedure, due to the way we address the question and the nature of the question we would like to address, the selection bias is unavoidable. Therefore, we made use of Begg's

## Genetic predisposition of stroke

**Table 1.** The association studies between selected candidate genes and IS

| Category/gene            | Risk alleles   | Patients/controls         | OR (95% CI) *P<0.05          | References |
|--------------------------|----------------|---------------------------|------------------------------|------------|
| Renin-angiotensin        |                |                           |                              |            |
| ACE                      | DD vs I        | 108/79 (Turkish)          | 0.949 (0.69-1.3)             | [24]       |
|                          | DD vs I        | 10070/22103 (meta)        | 1.37* (1.22-1.53)            | [25]       |
|                          | DD vs I        | 5870/4850 (meta, Chinese) | 2.30* (1.84-2.89)            | [26]       |
|                          | II, ID, DD     | 162/150 (South Indian)    | 1.88* (1.15-3.07) for D      | [27]       |
| AGT (M235T)              | TT vs. MM      | 1298/1094 (meta, Chinese) | 2.65* (1.89-3.71)            | [26]       |
|                          | T vs M         | 3842/12188 (meta)         | 1.447* (1.207-1.735)         | [28]       |
| CYP11B2 (C344T)          | TT vs C        | 332/250 (north Chinese)   | 1.572* (1.095-2.258)         | [29]       |
|                          | T vs C         | 3620/4090 (meta)          | 1.19 (0.95-1.49)             | [30]       |
| Coagulation/fibrinolysis |                |                           |                              |            |
| TPA                      | TT vs C        | 2299/1948 (meta)          | 2.42* (1.07-5.48)            | [31]       |
|                          | TT vs C        | 516/513                   | 1.37 (0.883-2.127)           | [32]       |
| PAI-1                    | 4G/4G          | 127/201 (Brazilian)       | 0.41 (0.24-0.68)             | [33]       |
|                          | 4G/4G          | 388/775 (2 studies)       | 1.79/1.60* (1.01-3.19)       | [34]       |
|                          | 4G/4G          | 600/600                   | 0.84 (0.64 to 1.11)          | [35]       |
| factor VII               | 670C           | 150/150                   | 2.00* (1.29-3.10)            | [36]       |
|                          | IVS7>7         | 150/150                   | 1.93* (1.20-3.09)            |            |
|                          | R353Q          | 81/149                    | 1.78 (0.88-3.66)             | [37]       |
| Antioxidation            |                |                           |                              |            |
| GCH1                     | rs841 GA+AA    | 558/557                   | 1.73* (1.27-2.35)            | [38]       |
| GPx-3                    | H2 haplotype   | 123/123                   | 2.07* (1.03-4.47)            | [39]       |
| Lipid                    |                |                           |                              |            |
| apoE                     | ε4 vs ε3       | 166/192                   | 1.16 (0.68-1.98)             | [40]       |
|                          | ε4 vs ε3       | 3814/3425 (meta )         | 2.34* (1.91-2.86)            |            |
|                          | ε4 vs ε3       | 72/171                    | NA                           | [41]       |
| PON1                     | 192 R vs Q     | 85/71                     | 1.41 (0.8-2.47)              | [42]       |
|                          | 107 TT         | 172/105                   | 1.973* (1.014-3.839)         | [43]       |
| APOA5                    | -1131 CC vs TT | 2294/1858 (meta)          | 4.47* (1.33-15.06)           | [44]       |
|                          | -1131 C vs T   | 302/289                   | 2.1* (1.3-4.7)               | [45]       |
|                          | IVS3+G476A     | 378/131                   | 2.844* (1.320-6.124)         | [46]       |
|                          | -1131 C vs T   | 378/131                   | 2.214* (1.158-4.235)         |            |
| Inflammation             |                |                           |                              |            |
| IL1                      | IL1α-889C/T    | 846/1223 (meta)           | 1.21 (0.86-1.70)             | [47]       |
|                          | IL1β-511C/T    | 133 4/1594 (meta)         | 1.22 (0.85-1.87)             |            |
| IL6                      | IL6-147G/C     | 1879/2092 (meta)          | 1.56 (0.61-3.99)             |            |
| TNFa                     | -308G/A        | 2349/1848 (2 studies)     | 1.30*/1.46* (1.02-1.96)      | [48]       |
|                          | -1031T/C       | 7106/7853 (meta)          | 1.43* (1.21-1.69)            |            |
|                          | -308G/A        | 3515/3949 (meta)          | 0.822 (0.648-1.042)          | [49]       |
| E-selectin               | S128R AC vs AA | 610/610                   | 5.47* (3.25-9.21)            | [50]       |
|                          | A561C          | 314/389                   | 2.73* (1.29-5.76)            | [51]       |
| ALOX5AP                  | SG13S114A/T    | 5361/5676 (meta)          | 1.47* (1.13-1.91)            | [52]       |
|                          | SG13S114A/T    | 1092/781 (meta)           | 1.22* (1.06-1.40)            | [53]       |
| Others                   |                |                           |                              |            |
| NOS1                     | G894T T vs GG  | 6537/6475 (meta)          | 1.60* (1.38-1.79) for Asians | [54]       |
|                          | T-786C         | 2125/2673 (meta)          | 1.14 (0.95-1.37) for Asians  |            |
|                          | 4b/a, a vs bb  | 3459/3951 (meta)          | 1.60* (1.30-1.97) for Asians |            |
|                          | T-786C         | 2836/ 3354 (meta)         | 1.14 (1.02-1.28) for Asians  | [55]       |
| TGFB1                    | -509 T vs C    | 312/5558                  | 1.29* (1.03-1.61)            | [56]       |
|                          | P10L LP vs LL  | 312/5558                  | 1.24* (1.03-1.50)            |            |
|                          | P10L LL vs P   | 271/207                   | 1.63* (1.06-2.49)            | [57]       |
| MTHFR                    | C677T T vs C   | 2223/2936                 | 1.28* (1.17-1.40)            | [58]       |
|                          | A1298C C vs T  | 2133/2572                 | 1.227* (1.062-1.416)         | [59]       |
|                          | C677/1298C     | 92/259                    | 3.463* (1.699-7.058)         | [60]       |

\*P<0.05 and #P<0.05 represent the comparison between the risk alleles.

funnel plot to examine the underlying publication bias. The data analysis was performed using RevMan 5.2 statistical software. \* $P < 0.05$  represents the difference significant.

## Results

### *The current status of genetic risk identification for common stroke*

One of our focuses of this review is to summarize the overall status of genetic risk identification for common stroke. Many studies have combined ischemic and hemorrhagic strokes. We reasoned that it is unlikely that these very different pathological conditions are under the same genetic influences. To reduce the confounding effects, we analyze the genetic risk factors for IS and HS separately.

**Genetic risk factors for IS:** About 80% of all strokes are IS [23]. There is a long list of candidate gene pathways and genes that have been studied for a possible association with IS. The most widely investigated genes are those involved in the inflammation, lipid metabolism, nitric oxide release, and extracellular matrix hemostasis. In most cases, however, findings were either negative or could not be replicated in subsequent studies. Based on our primary study we chose to summarize stroke related genes that are involved in inflammation, the renin-angiotensin system, and atherosclerosis and lipid metabolism. Unsurprisingly, not all associations have been consistently replicated. **Table 1** summarized the major findings about the selected candidate risk alleles.

**Genetic risk factors for HS:** In contrast, to IS, only about 20% of strokes are HS, therefore, few systematic studies have been focused on HS. Yet, through similar approaches, we reviewed the risk alleles in genes of Renin-angiotensin-aldosterone, blood coagulation, lipid metabolism-related, homocysteine metabolism-related, inflammation-related, extracellular matrix (ECM) degradation, and antioxidant systems. Unfortunately, current data for the genetic risk factors for HS are even less conclusive.

The angiotensin-converting enzyme (ACE) gene has an insertion (I)/deletion (D) polymorphism in Intron 16, which has been associated with variations in ACE activity. The DD genotype has been associated with HS in Polish and Indian populations, and a meta-analysis of 744 cases and 1289 controls revealed significant associations of homozygosity for the ACE/I allele with hemorrhagic stroke (1.48\*, 95% 1.20-1.83)

[24]. However, no differences in genotype frequency were observed between controls and subjects with HS in Japanese [25] or Greek cohorts [26].

Fewer studies have specifically addressed the relationship between genes of coagulation/fibrinolysis and HS. For example, frequency of the prothrombin 20210A/G genotype was lower in patients than in controls, whereas there was no significant difference in the prevalence of the V G1691A Leiden mutation [27], however, this conclusion has not been independently repeated. Similarly, association studies of the polymorphisms of Factor XIII [28-30] and Factor VII [27, 31] genes also have generated conflicting results.

Apolipoprotein E (ApoE) [32-34], ApoH [35] and Apo(a) [36] play major roles in lipid transport and metabolism. Studies have produced conflicting results regarding the influence of ApoE alleles on predisposition to HS. Several studies have reported an association between the  $\epsilon 4$  allele and HS risk [32, 37]. However, a meta-analysis of 31 studies (5961 cases, 17,965 controls) showed that the  $\epsilon 2+$  genotype, but not the  $\epsilon 4+$  genotype, was associated with HS [33]. In addition, a prospective study of 5671 patients suggested that both the  $\epsilon 2$  and  $\epsilon 4$  alleles were associated with an increased risk of HS [38].

Similarly, four polymorphisms of ApoH were examined in 140 HS patients in a Chinese population [35]. Frequencies of the A allele of G341A were significantly higher in HS patients than in controls, especially in HS patients with hypertension and a family history of stroke. No differences in the genotype frequencies of the G817T, G1025C, and C1080T polymorphisms were found. Similarly, polymorphisms of Apo(a) has been associated with HS [36]. Unfortunately, the data of ApoH and Apo(a) have not been independently repeated.

Two polymorphisms in the methylenetetrahydrofolate reductase (MTHFR) gene, C677T and A1298C, have been shown to reduce enzyme activity and elevate plasma homocysteine levels. These polymorphisms were also found to be genetic risk factors for hemorrhagic and ischemic stroke, respectively, in a Turkish Caucasian population [23]. In a case-control study of Mongolian patients with HS, the C677T polymorphism TT genotype was more common in patients with HS and was associated with reduced plasma folate levels [39]. However, no

association of the C677T polymorphism with HS was observed in two other studies [40, 41].

Inflammation-related genes have also been implied in literatures, for example, the -572G>C polymorphism of the IL6 gene was examined in 3151 Japanese individuals. Analysis revealed that the -572G>C polymorphism was significantly associated with HS [42]. However, the -74G>C polymorphism of IL6 showed no association with HS [43]. Similarly, associations of spontaneous deep HS with four single-nucleotide polymorphisms (T-1031C, C-863A, C-857T, and G-308A) within the Tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ) gene promoter were examined in a Taiwanese population. The risk was positively associated with minor alleles -1031C and -308A in men, but inversely associated with -863A in females, which indicated that the associations were gender-dependent [44]. However, none of the data have been independently repeated.

### *The evolving issues*

*The issues of methodology:* The method of genetic risk factor identification evolves with technical progress. Most early studies use the genetic lineage or candidate gene approaches, but the high efficient genome-wide association studies (GWAS) and Next generation sequencing (NGS) gained momentum recently. In addition, cutting edge gene expression profiling, proteomic and metabolomics approaches offer new hope for risk assessment.

1) Genetic linkage approach. This technique examines genetic variants through multiple generations of the same family and examines incidence with disease status. This is a powerful technique for Mendelian diseases where a single gene controls the phenotype, but it has had relatively little success in identifying the risk factors of stroke due to the late age of onset and lack of large pedigrees with multiple affected individuals. Thus, this approach offers little help in identify the genetic risk factors for polygenic multifactorial disease, such as stroke.

2) Candidate gene approach. This is a strictly hypothesis driven approach.

For above-mentioned limitations, most published studies have taken a candidate gene approach using case-control methodologies. Candidate genes are typically chose on the basis of biological plausibility for the disease in question, and then the genetic variants (poly-

morphisms) at that one locus are subjected to classic association test. Many potential risk polymorphisms have been identified though this approach. However, further study with this approach is greatly limited by the lack of deep understanding the underlying molecular mechanisms of disease process, and this approach is also incompetent to identify the potential gene-gene interactions. In addition, any given genetic variant has a low pre-test probability of being truly associated with a phenotype. To be reliably detected, small relative risks require large sample sizes, probably in the order of 1,000 patients or more. Few studies have achieved such numbers. This difficulty, together with the mediocre success of candidate-gene studies provides support to alternative systematic hypothesis-free approaches, such as genome-wide linkage and genome-wide association studies (GWASs).

3) The Genome Wide Association study (GWAS) approach is not strictly hypothesis driven, thus it is not limited by the lack of understanding of the disease process. GWAS studies could be retrospective case-control, prospectively collected cohorts or family-based association studies. This technique looks at multiple genetic variants (typically up to 1 million) at a time. The genetic variants are spread throughout the genome at random and allow systematic unbiased investigation of a large number of regions in cases and controls. Although complex and expensive, this technique provides unprecedented power to identify multiple risk alleles in one experiment. Over the past several years, GWAS have succeeded in identifying hundreds of genetic markers associated with common diseases. Furthermore, at least theoretically, this approach is competent to identify the potential gene-gene interaction. However, the limitations of this approach are also obvious: similar to candidate approach, individuals are not related, thus requiring large, typically several thousand, cases and controls for comparison, and due to the large population, it requires strong efforts for quality assurance concerning patient recruiting, subtyping and data processing. Any inconsistency in these procedures may cause undesired noises, which can easily dilute the weak signals from relatively rare risk alleles. Consistent with this idea, the first genome-wide association study in stroke in 2007 found no genetic locus specifically and robustly associated with the disease [45].

Similarly, genome-wide linkage studies utilize family structure and large numbers of tagging

SNPs to track the inheritance of stroke risk with the transmission of the SNP alleles. Although genome-wide linkage studies have the ability to detect single risk loci with relatively large effect, success has been limited. For example, a whole genome linkage scan of 109 families from a genetically homogeneous region of Northern Sweden failed to identify any new major loci for ischemic stroke [46].

4) Next generation sequencing (NGS) approach. NGS is currently the cutting edge technology available for direct sequencing of DNA, allowing determination of the entire exome (coding portion) or the entire human genome in a single experiment [47]. Because of its unparalleled efficiency and resolution, far fewer individuals are required for identifying the potential risk factors. In fact, at least theoretically, NGS can exhaust all potential genetic risk factors in selected populations in one experiment. Thus, NGS has been used predominantly to examine the Rare Variant, Common Disease (RVCD) hypothesis. The only limitation is that high-coverage whole-genome sequencing remains costly and time-consuming, even though current advances in multiplexing of samples by labeling with genetic tags before sequencing allows a reduction in cost.

5) Gene expression profiling, proteomic and metabolomics approaches. Studies of mRNA expression and candidate proteomic approaches have also shown promise for risk assessment and the development of signatures for stroke classification. mRNA profiling of the peripheral blood has been used to investigate stroke risk factors. For example, using peripheral blood mononuclear cells, a validated and replicated gene expression signature of ischemic stroke within the first 72 h has been identified [21], i.e., a panel of 22 genes was about 80% accurate for the detection of ischemic stroke in an external sample [21]. Another group found this panel to be over 85% accurate, using whole-blood profiling [20].

Overall, the evolving of the methodology is a normal, self-perfection process, and it brings more options and opportunities to further probe the genetic risk factors, however, this natural process did create a problem for meta-analysis, because it is impossible to simply pool the data from different studies that using different methods together.

*The issues of classification:* The remarkable heterogeneity of stroke highlights the impor-

tance of the accurate classification. The careful classification has been also driven by the needs of both clinical trials and epidemiological studies, and the ultimate goal of classification should be to categorize the heterogeneous population into a manageable number of discrete, more or less homogenous subtypes/subpopulations, because different subtypes may differ significantly in their causes, pathophysiology, treatments, and outcomes.

The ideal classification should not only be consistent, precise, reliable, universally accepted, and ease to use, but also integrate the clinical features, diagnostic tests, knowledge about potential etiologic factors. Unfortunately, even though there are many different classification systems and they are evolving rapidly over time, none of them have met the above-mentioned criteria yet.

For example, the representative classification systems include the Stroke Data Bank classification (published in 1978), the Lausanne Stroke Registry Classification (published in 1988), the Oxfordshire Community Stroke Project Subtype Classification (OSCP, published in 1991), the Trial of Org 10172 in acute stroke treatment (TOAST, published in 1997) classification, the GÉNIC classification (published in 2000), the causative classification system (CCS, published in 2007), the A-S-C-O (Atherosclerosis, Small vessel disease, Cardiac source, Other cause) phenotypic classification (published in 2009), and the recent Chinese ischemic stroke subclassification (CISS, published in 2011). As expected, none of them are perfect, and each has its own limitations.

Even though it is not our intention to extensively comparing the advantages and shortcomings of each individual classification system, this review does want to summarize some of the general trend, i.e., 1) the old classification systems are generally easier to use and tend to be more reliable; 2) but the more recent classification systems can accommodate the new findings from the emerging new imaging and diagnostic technologies. 3) therefore, the more recent classification systems allow more detailed classification, however the reliability of the new classification systems have not been extensively tested.

Overall, even though the improvement of classification systems is a necessary, self-perfection process, we believe that the co-existing of multiple, competing classification systems is

currently a major contribution to the confusion in the field. Furthermore, without a universally accepted standard classification system, it is impossible to directly compare the data from different studies. In fact, we believe that a substantial portion of the current conflicting data may simply reflect the inconsistency in the criteria of stroke classification in different study. For this reason, a universally accepted, precise classification of stroke subtypes is urgently needed.

These studies in stroke genetics have been largely disappointing, especially the inconsistency in replicating the initial findings. The potential reasons are many, and some are seemingly legitimate reasons, but others are not. For example, the sample sizes, the different sex and ethnic origin, and the variable robustness of different alleles are all legitimate reasons of the conflicting results. In contrast, inconsistent in genotyping, phenotypes and classification are the major illegitimate reasons that will create noises that will eventually mask the correlation.

The ultimate goal of genetic risk identification is to establish the correlation of specific genetic trait (polymorphism) with certain disease phenotype. Any future genetic study, whether hypothesis driven or non-hypothesis driven, should address the above-mentioned issues, especially the proposed illegitimate reasons have to be controlled carefully. Power calculations demonstrating the number of cases required for confirmation or refutation of a finding should be included to allow an estimate of the significance and robustness of the findings presented.

Increasing evidence suggests genetic risks differ depending on stroke subtype. Future genetic studies should therefore include reference to subtypes and subtype specific risks. These measures will, of course, depend heavily on a consistent and accurate classification system, and may lead to increased cost and complexity of studies, but we argue that only through such robust experimental procedures that we will truly begin to understand the genetic risks of stroke.

In the long-term, to fully understand the risk factors of stroke, it is necessary to further understand how different polymorphisms differentially influence the gene-gene and gene-environment interactions, and eventually lead to predisposition to stroke. Unfortunately, the study of allele specific gene-gene and gene-

environment interactions, in vivo or in vitro, is still in its infancy, and there is currently no easy and valid way to dissect and gauge the predisposition of specific allele in the context of gene-gene and gene-environment interactions.

## Disclosure of conflict of interest

None.

**Address correspondence to:** Dr. Jian-Wen Gu, Department of Neurosurgery, General Hospital of The People's Liberation Army Chengdu Military Region, Sichuan 610083, China. E-mail: gujianwen-sic@yeah.net

## References

- [1] Nie SW, Wang XF, Tang ZC. Correlations between MMP-2/MMP-9 promoter polymorphisms and ischemic stroke. *Int J Clin Exp Med* 2014; 7: 400-404.
- [2] Hatano S. Experience from a multicentre stroke register: a preliminary report. *Bull World Health Organ* 1976; 54: 541-553.
- [3] Cole JW, Meschia JF. Stroke Genetics Update: 2011. *Curr Cardiovasc Risk Rep* 2011; 5: 533-541.
- [4] Montagna JC, Santos BC, Battistuzzo CR, Loureiro AP. Effects of aquatic physiotherapy on the improvement of balance and corporal symmetry in stroke survivors. *Int J Clin Exp Med* 2014; 7: 1182-1187.
- [5] Kovacic M, Madarasz S, Kral M, Veverka T, Herzig R, Kanovsky P. Risk factors associated with ischemic heart disease occurrence in acute ischemic stroke patients. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* 2013; 157: 168-171.
- [6] Ahmad Y, Lip GY, Lane DA. Recent developments in understanding epidemiology and risk determinants of atrial fibrillation as a cause of stroke. *Can J Cardiol* 2013; 29: S4-S13.
- [7] Yong H, Foody J, Linong J, Dong Z, Wang Y, Ma L, Meng HJ, Shiff S, Dayi H. A systematic literature review of risk factors for stroke in China. *Cardiol Rev* 2013; 21: 77-93.
- [8] Touze E, Rothwell PM. Sex differences in heritability of ischemic stroke: a systematic review and meta-analysis. *Stroke* 2008; 9: 16-23.
- [9] Demarin V, Lisak M, Morovic S, Cengic T. Low high-density lipoprotein cholesterol as the possible risk factor for stroke. *Acta clinica Croatica* 2010; 49: 429-439.
- [10] Chitsaz A, Mousavi SA, Yousef Y, Mostafa V. Comparison of changes in serum fibrinogen level in primary intracranial hemorrhage (ICH) and ischemic stroke. *ARYA atherosclerosis* 2012; 7: 142-145.
- [11] Norrving B, Kissela B. The global burden of stroke and need for a continuum of care. *Neurology* 2013; 80: S5-S12.

- [12] Al Mamun M, Rumana N, Specogna AV, Turin TC. Studying effects of weather parameters on acute stroke: considering the contribution of "risk factors" and "triggering factors" together. *J Stroke Cerebrovasc Dis* 2014; 23: 194-195.
- [13] Xie F, Xu R, Song X, Zhu H, Wang X, Zhu J. Joint protective effect of exogenous neuroglobin and hemin in rat focal ischemic brain tissues. *Int J Clin Exp Med* 2014; 7: 2009-2016.
- [14] Romero JR, Morris J, Pikula A. Stroke prevention: modifying risk factors. *Ther Adv Cardiovasc Dis* 2008; 2: 287-303.
- [15] Pezzini A. Genetic determinants of juvenile stroke. *Thromb Res* 2012; 129: 330-335.
- [16] Polychronopoulos P, Gioldasis G, Ellul J, Metallinos IC, Lekka NP, Paschalis C, Papapetrooulos T. Family history of stroke in stroke types and subtypes. *J Neurol Sci* 2002; 195: 117-122.
- [17] Munshi A, Kaul S. Genetic basis of stroke: an overview. *Neurol India* 2010; 58: 185-190.
- [18] Woo D, Sauerbeck LR, Kissela BM, Khoury JC, Szaflarski JP, Gebel J, Shukla R, Pancioli AM, Jauch EC, Menon AG, Deka R, Carrozzella JA, Moomaw CJ, Fontaine RN, Broderick JP. Genetic and environmental risk factors for intracerebral hemorrhage: preliminary results of a population-based study. *Stroke* 2002; 33: 1190-1195.
- [19] Bak S, Gaist D, Sindrup SH, Skyttke A, Christensen K. Genetic liability in stroke: a long-term follow-up study of Danish twins. *Stroke* 2002; 33: 769-774.
- [20] Tournier-Lasserre E, Joutel A, Melki J, Weissenbach J, Lathrop GM, Chabriat H, Mas JL, Cabanis EA, Baudrimont M, Maciazek J. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy maps to chromosome 19q12. *Nat Genet* 1993; 3: 256-259.
- [21] Aslan C, Melikoglu C, Ocal I, Saglam G, Sutcu R, Hosnuter M. Effect of epigallocatechin gallate on ischemia-reperfusion injury: an experimental study in a rat epigastric island flap. *Int J Clin Exp Med* 2014; 7: 57-66.
- [22] Jeffs B, Clark JS, Anderson NH, Gratton J, Brosnan MJ, Gauguier D, Reid JL, Macrae IM, Dominiczak AF. Sensitivity to cerebral ischaemic insult in a rat model of stroke is determined by a single genetic locus. *Nat Genet* 1997; 16: 364-367.
- [23] Ay H. Advances in the diagnosis of etiologic subtypes of ischemic stroke. *Curr Neurol Neurosci Rep* 2010; 10: 14-20.
- [24] Peck G, Smeeth L, Whittaker J, Casas JP, Hingorani A, Sharma P. The genetics of primary haemorrhagic stroke, subarachnoid haemorrhage and ruptured intracranial aneurysms in adults. *PLoS One* 2008; 3: e3691.
- [25] Nakata Y, Katsuya T, Rakugi H, Takami S, Sato N, Kamide K. Polymorphism of angiotensin converting enzyme, angiotensinogen, and apolipoprotein E genes in a Japanese population with cerebrovascular disease. *Am J Hyperten* 1997; 10: 1391-1395.
- [26] Dardiotis E, Jagiella J, Xiromerisiou G, Dardiotis M, Vogiatzi C, Urbanik A, Paterakis K, Komnos A, Fountas KN, Slowik A, Hadjigeorgiou GM. Angiotensin-converting enzyme tag single nucleotide polymorphisms in patients with intracerebral hemorrhage. *Pharmacogenet Genomics* 2011; 21: 136-141.
- [27] Corral J, Iniesta JA, Gonzalez-Conejero R, Villalon M, Vicente V. Polymorphisms of clotting factors modify the risk for primary intracranial hemorrhage. *Blood* 2001; 97: 2979-2982.
- [28] Gemmati D, Serino ML, Ongaro A, Tognazzo S, Moratelli S, Resca R, Moretti M, Scapoli GL. A common mutation in the gene for coagulation factor XIII-A (VAL34Leu): a risk factor for primary intracerebral hemorrhage is protective against atherothrombotic diseases. *Am J Hematol* 2001; 67: 183-188.
- [29] Cho KH, Kim BC, Kim MK, Shin BA. No association of factor XIII Val34Leu polymorphism with primary intracerebral hemorrhage and healthy controls in Korean population. *J Korean Med Sci* 2002; 17: 249-253.
- [30] Slowik A, Dziedzic T, Pera J, Figlewicz DA, Szczudlik A. Coagulation factor XIII Val34Leu polymorphism in patients with small vessel disease or primary intracerebral hemorrhage. *Cerebrovasc Dis* 2005; 19: 165-170.
- [31] Greisenegger S, Weber M, Funk M, Endler G, Lang W, Ferrari J, Lalouschek W, Mannhalter C. Polymorphisms in the coagulation factor VII gene and risk of primary intracerebral hemorrhage. *Eur J Neurol* 2007; 14: 1098-1101.
- [32] Sinha SK, Mishra V, Nagwani S, Rajavashisth TB. Effects of G-CSF on serum cholesterol and development of atherosclerotic plaque in apolipoprotein E-deficient mice. *Int J Clin Exp Med* 2014; 7: 1979-1989.
- [33] Sudlow C, Martinez Gonzalez NA, Kim J, Clark C. Does apolipoprotein E genotype influence the risk of ischemic stroke, intracerebral hemorrhage, or subarachnoid hemorrhage? Systematic review and meta-analyses of 31 studies among 5961 cases and 17,965 controls. *Stroke* 2006; 37: 364-370.
- [34] Seifert T, Lechner A, Flooh E, Schmidt H, Schmidt R, Fazekas F. Lack of association of lobar intracerebral hemorrhage with apolipoprotein E genotype in an unselected population. *Cerebrovasc Dis* 2006; 21: 266-270.
- [35] Xia J, Yang QD, Yang QM, Xu HW, Liu YH, Zhang L, Zhou YH, Wu ZG, Cao GF. Apolipoprotein H gene polymorphisms and risk of primary cerebral hemorrhage in a Chinese population. *Cerebrovasc Dis* 2004; 17: 197-203.
- [36] Sun L, Li Z, Zhang H, Ma A, Liao Y, Wang D, Zhao B, Zhu Z, Zhao J, Zhang Z, Wang W, Hui R. Pentanucleotide TTTTA repeat polymorphism of apolipoprotein(a) gene and plasma lipoprotein(a) are associated with ischemic and hemorrhagic stroke in Chinese: a multicenter case-

- control study in China. *Stroke* 2003; 34: 1617-1622.
- [37] Kokubo Y, Chowdhury AH, Date C, Yokoyama T, Sobue H, Tanaka H. Age-dependent association of apolipoprotein E genotypes with stroke subtypes in a Japanese rural population. *Stroke* 2000; 31: 1299-1306.
- [38] Tzourio C, Arima H, Harrap S, Anderson C, Godin O, Woodward M, Neal B, Bousser MG, Chalmers J, Cambien F, MacMahon S. APOE genotype, ethnicity, and the risk of cerebral hemorrhage. *Neurology* 2008; 70: 1322-1328.
- [39] Fang X, Namba H, Akamine S, Sugiyama K. Methylenetetrahydrofolate reductase gene polymorphisms in patients with cerebral hemorrhage. *Neurol Res* 2005; 27: 73-76.
- [40] Sancakdar E, Guven AS, Uysal EB, Kaya A, Deveci K, Karapinar H, Akkar I. Evaluation of cytokines as Th1/Th2 markers in pathogenesis of children with Crimean-Congo hemorrhagic fever. *Int J Clin Exp Med* 2014; 7: 751-757.
- [41] Somarajan BI, Kalita J, Mittal B, Misra UK. Evaluation of MTHFR C677T polymorphism in ischemic and hemorrhagic stroke patients. A case-control study in a Northern Indian population. *J Neurol Sci* 2011; 304: 67-70.
- [42] Yamada Y, Metoki N, Yoshida H, Satoh K, Ichihara S, Kato K, Kameyama T, Yokoi K, Matsuo H, Segawa T, Watanabe S, Nozawa Y. Genetic risk for ischemic and hemorrhagic stroke. *Arterioscler Thromb Vasc Biol* 2006; 26: 1920-1925.
- [43] Strand M, Soderstrom I, Wiklund PG, Hallmans G, Weinehall L, Soderberg S, Olsson T. Polymorphisms at the osteoprotegerin and interleukin-6 genes in relation to first-ever stroke. *Cerebrovasc Dis* 2007; 24: 418-425.
- [44] Chen YC, Hu FJ, Chen P, Wu YR, Wu HC, Chen ST, Lee-Chen GJ, Chen CM. Association of TNF-alpha gene with spontaneous deep intracerebral hemorrhage in the Taiwan population: a case control study. *BMC Neurol* 2010; 10: 41.
- [45] Matarin M, Brown WM, Scholz S, Simon-Sanchez J, Fung HC, Hernandez D, Gibbs JR, De Vrieze FW, Crews C, Britton A. A genome-wide genotyping study in patients with ischaemic stroke: initial analysis and data release. *Lancet Neurol* 2007; 6: 414-420.
- [46] Nilsson-Ardnor S, Janunger T, Wiklund PG, Lackovic K, Nilsson AK, Lindgren P, Escher SA, Stegmayr B, Asplund K, Holmberg D. Genome-wide linkage scan of common stroke in families from northern Sweden. *Stroke* 2007; 38: 34-40.
- [47] Kato K. Impact of the next generation DNA sequencers. *Int J Clin Exp Med* 2009; 2: 193-202.
- [48] Cui G, Wang H, Li R, Zhang L, Li Z, Wang Y. Polymorphism of tumor necrosis factor alpha (TNF-alpha) gene promoter, circulating TNF-alpha level, and cardiovascular risk factor for ischemic stroke. *J Neuroinflamm* 2012; 9: 235.
- [49] Gu L, Wu G, Long J, Su L, Yan Y, Chen Q. The role of TNF-alpha 308G>A polymorphism in the risk for ischemic stroke. *Am J Med Sci* 2013; 345: 227-33.
- [50] Roy S, Das S, Danaboina R, Sharma V, Kaul S, Jyothy A. Association of E-selectin Gene Polymorphism (S128R) with Ischemic Stroke and Stroke Subtypes. *Inflammation* 2013; 37: 599-603.
- [51] Zhao DX, Feng J, Cong SY, Zhang W. Association of E-selectin gene polymorphisms with ischemic stroke in a Chinese Han population. *J Neurosci Res* 2012; 90: 1782-1787.
- [52] Ye F, Liu NN, Zheng YQ, Zhang WJ, Wang CM, Xu Y. Three polymorphisms of ALOX5AP and risk of ischemic stroke in Chinese: Evidence from a meta-analysis. *J Neurol Sci* 2013; 336: 93-98.
- [53] Domingues-Montanari S, Fernandez-Cadenas I, del Rio-Espinola A, Corbeto N, Krug T, Manso H. Association of a genetic variant in the ALOX5AP with higher risk of ischemic stroke: a case-control, meta-analysis and functional study. *Cerebrovasc Dis* 2010; 29: 528-537.
- [54] Yao YS, Chang WW, Jin YL, He LP. An updated meta-analysis of endothelial nitric oxide synthase gene: three well-characterized polymorphisms with ischemic stroke. *Gene* 2013; 528: 84-92.
- [55] Liu R, Geng P, Ma M, Yu S, Wang X, Zhang W. Association between Endothelial nitric oxide synthase gene polymorphism (T-786C) and ischemic stroke susceptibility: A meta-analysis. *Int J Neurosci* 2013; 124: 642-651.
- [56] Sie MP, Uitterlinden AG, Bos MJ, Arp PP, Breteler MM, Koudstaal PJ. TGF-beta 1 polymorphisms and risk of myocardial infarction and stroke: the Rotterdam Study. *Stroke* 2006; 37: 2667-2671.
- [57] Kim Y, Lee C. The gene encoding transforming growth factor beta 1 confers risk of ischemic stroke and vascular dementia. *Stroke* 2006; 37: 2843-2845.
- [58] Li P, Qin C. Methylenetetrahydrofolate reductase (MTHFR) gene polymorphisms and susceptibility to ischemic stroke: A meta-analysis. *Gene* 2013; 535: 359-364.
- [59] Kang S, Wu Y, Liu L, Zhao X, Zhang D. Association of the A1298C polymorphism in MTHFR gene with ischemic stroke. *J Clin Neurosci* 2013; 21: 198-202.
- [60] Sazci A, Ergul E, Tuncer N, Akpinar G, Kara I. Methylenetetrahydrofolate reductase gene polymorphisms are associated with ischemic and hemorrhagic stroke: Dual effect of MTHFR polymorphisms C677T and A1298C. *Brain Res Bull* 2006; 71: 45-50.