Evaluating the Genetic Component of Ischemic Stroke Subtypes
A Family History Study

Paula Jerrard-Dunne, MRCPI; Geoffrey Cloud, MRCP; Ahamad Hassan, MRCP; Hugh S. Markus, FRCP

Background and Purpose—Twin and family history studies support a role for genetic factors in stroke risk. Because the etiology of ischemic stroke is heterogeneous, genetic factors may vary by etiologic subtype. We determined the familial aggregation of stroke risk in different stroke phenotypes and used the results to model estimated sample size requirements for case-control studies.

Methods—One thousand consecutive white subjects with ischemic stroke and 800 white controls matched for age and sex were recruited. A first-degree family history of stroke and myocardial infarction was obtained by structured interview. Stroke subtype was determined with the use of modified Trial of Org 10172 in Acute Stroke Treatment (TOAST) criteria.

Results—A family history of stroke at ≤65 years was a significant risk factor for large-vessel disease (odds ratio [OR], 2.24; 95% CI, 1.49 to 3.36; \( P < 0.001 \)) and for small-vessel disease (OR, 1.93; 95% CI, 1.25 to 2.97; \( P = 0.003 \)). When only cases aged ≥65 years were considered, these ORs increased to 2.93 (95% CI, 1.68 to 5.13) (\( P < 0.001 \)) and 3.15 (95% CI, 1.81 to 5.50) (\( P < 0.001 \)), respectively. No significant associations were seen for cardioembolic stroke or stroke of undetermined etiology.

Conclusions—A family history of vascular disease is an independent risk factor for both large-vessel atherosclerosis and small-vessel disease, especially in cases presenting before age 65 years. The estimated sample sizes for case-control studies illustrate how candidate gene studies for ischemic stroke might be made more effective by focusing on these specific phenotypes, in which the genetic component of the disease appears to be strongest. (Stroke. 2003;34:1364-1369.)

Key Words: cerebral infarction ■ epidemiology ■ genetics ■ risk factors ■ stroke classification

Both twin and family history studies support a role for genetic factors in stroke risk.1-4 To date, candidate gene studies investigating the molecular basis of polygenic ischemic stroke have yielded inconsistent or inconclusive results.5 One explanation for this is the heterogeneity of pathophysiological mechanisms causing ischemic stroke.

Recent advances in neuroimaging have made it easier to identify underlying disease mechanisms and appropriately classify stroke. It is likely that the genetic risk profile at the molecular level differs according to ischemic stroke subtype. For example, the factors leading to carotid atherothrombosis may be different from those that predispose to cerebrovascular small-vessel disease or to cardioembolism.

While certain genes, such as those that predispose to cerebral ischemia, may increase ischemic stroke risk in all the major subtypes, other genetic factors that predispose to underlying disease mechanisms, such as carotid atherosclerosis or microangiopathy, may be relatively more important in specific etiologic subtypes. An example of this is seen in the monogenic stroke disorder CADASIL (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy), in which the notch 3 mutation predisposes specifically to lacunar infarction.5

Future candidate gene studies investigating the genetic basis of ischemic stroke may be more successful if they focus on selected patient groups in which genetic factors are most important. In particular, it has been suggested that familial factors are more important in individuals presenting with stroke at a young age5 and in certain stroke subtypes.5 Using a positive family history as a marker of increased genetic risk, in this study we sought to determine the familial component of different ischemic stroke subtypes, specifically focusing on younger age groups that are likely to have the greatest genetic component.

Subjects and Methods
One thousand consecutive white subjects presenting to 2 south London hospitals with a diagnosis of ischemic stroke or transient ischemic attack were recruited. During the same period, 800 white controls, free of clinical cerebrovascular disease and matched for age and sex (within 5 years), were recruited by random sampling of family physician practices from the same geographic area as the stroke cases. A physician trained in stroke medicine interviewed...
cases and controls. A complete first-degree family history of stroke and myocardial infarction (MI) was obtained by a structured interview. Subjects were asked whether parents and siblings were alive or dead, whether they had been affected by stroke or MI, and the age at onset in affected family members. A positive family history was defined as a reported history of stroke or MI in a first-degree relative (ie, biological parent or sibling). Family history was classified as negative in cases in which the cause of death was unknown. When subjects were unable to provide an adequate family history, a collateral history from a relative was sought whenever possible.

Risk factors documented in both cases and controls included age, sex, smoking status, hypertension, diabetes, hyperlipidemia, and ischemic heart disease. Smoking status was categorized as current smoker (patient admission or ≥1 cigarette per day in the past 12 months), ex-smoker, or never-smoker. Hypertension was defined as pharmacological treatment for hypertension or systolic blood pressure ≥160 mm Hg and/or diastolic blood pressure ≥95 mm Hg persisting >7 days after the acute event (World Health Organization classification). Diabetes was defined as reported or medical notes of either diet-controlled, oral hypoglycemic–treated, or insulin-treated diabetes. Hyperlipidemia was defined as pharmacological treatment or total serum cholesterol ≥6.5 mmol/L. Ischemic heart disease was defined as self-reported or hospital record of MI or angina. Details were corroborated by patients’ medical records. Relatives’ medical records were not available to validate family history data. Informed consent was obtained from all subjects, and the local ethics committee approved the study.

**Stroke Subtyping**

All cases were examined by a neurologist and underwent neuroimaging (CT and/or MRI) and an ECG. Ninety-five percent had extracranial carotid and vertebral duplex ultrasound. Echocardiography (31%) and MR angiography and/or transcranial Doppler ultrasound (9%) were performed when clinically indicated. With the use of clinical, radiological, cardiac, and ultrasound test results, each case was assessed according to modified Trial of Org 10172 in Acute Stroke Treatment (TOAST) criteria to determine stroke subtype. Large-vessel disease was defined as >50% stenosis or occlusion of an appropriate major brain artery or branch cortical artery in the absence of sources of cardiac embolism. Small-vessel disease was defined as a clinical lacunar syndrome with a relevant infarct of <1.5 cm in the absence of a cardioembolic source or carotid stenosis >50%. Cardioembolic stroke was defined as the presence of atrial fibrillation, MI in the past 6 months, or a high-risk source of embolism identified on echocardiogram according to TOAST criteria. Stroke of undetermined etiology was used when no etiologic source could be identified. A tandem (combined) classification was used when >1 etiology was identified. Stroke of other determined etiology included those with carotid dissection, vasculopathies, and hematologic disorders. Subjects with a monogenic cause of stroke, eg, CADASIL, were excluded from the study. Tandem strokes and strokes of other determined etiology accounted for only a small number of ischemic strokes, and therefore analysis by subtype was confined to the 4 major subtypes: large-vessel disease, small-vessel disease, cardioembolic stroke, and ischemic stroke of undetermined etiology.

**Statistical Methods**

Differences between groups were examined with the χ² test to calculate proportions and with Student’s t test for continuous variables. Multivariate odds ratios (ORs) and 95% CIs for a family history of stroke or MI in cases versus controls were calculated with the use of binary logistic regression analysis controlling for age, sex, arterial hypertension, diabetes mellitus, serum cholesterol, and smoking status. ORs were calculated for overall ischemic stroke and for the 4 major subtypes. Because there is some evidence to suggest that genetic factors are more important in younger individuals, a prespecified analysis of families with disease onset at a young age (≤65 years) was also performed. To determine whether the effect of a positive family history was dependent on age, the population was then stratified according to the age of onset of stroke (in 5-year categories), and ORs were recalculated.

To determine potential implications for planning future candidate gene studies, we created a model using the observed multivariate ORs to estimate the sample size requirements for case-control studies with specific stroke subtypes and age groups compared with ischemic stroke overall. Sample sizes were estimated on the basis of standard tests for the difference between proportions based on the anticipated OR. For this model, sample sizes required to detect an effect equal to the magnitude of the observed multivariate OR were calculated, with a 2-tailed significance level of 0.05 and a power of 0.8, with the assumption of candidate gene allele frequencies of 5%, 10%, or 15%. The software used for sample size calculations was DSTPLAN version 4.2 (University of Texas, MD Anderson Cancer Center, 2000).

**Results**

In 22 cases (2.2%) and 11 controls (1.4%), subjects were unable to provide a family history because they were adopted or were not in contact with some or all of their first-degree relatives. In an additional 34 cases (3.4%), subjects were unable to provide an adequate history because of dysphasia or confusion and because no collateral history was available. Therefore, full family history data were available for analysis in 944 cases (94.4%) and 789 controls (98.6%).

There were no significant differences between cases and controls with regard to age, sex, or number of siblings in the overall analysis. Arterial hypertension, diabetes mellitus, ischemic heart disease, hyperlipidemia, and current smoking were all significantly more common in cases than in controls (Table 1). The distribution of stroke subtypes was as follows: large-vessel disease, 262 (26.2%); small-vessel disease, 232 (23.2%); cardioembolic, 118 (11.8%); ischemic stroke of undetermined etiology, 296 (29.6%); tandem, 69 (6.9%); and other determined etiology, 23 (2.3%). Demographic and risk factor profiles according to the 4 major stroke subtypes are given in Table 1. Subjects with cardioembolic stroke were older and had a more equal sex distribution than controls. A greater proportion of subjects with stroke of undetermined etiology were female compared with the other stroke subtypes.

The relationship between family history parameters and ischemic stroke is shown in Table 2. The strongest risk factors for stroke were a family history of stroke at ≤65 years (OR, 1.69; 95% CI, 1.25 to 2.29; P=0.001) and a family history of MI at any age (OR, 1.57; 95% CI, 1.29 to 1.91; P<0.001). Additional adjustment for age, sex, and vascular risk factors only slightly attenuated these ORs, which remained significant (Table 2). The observed relationships were similar when maternal, paternal, or sibling risk was studied individually; for example, the OR for ischemic stroke at any age conferred by a paternal history of stroke at ≤65 years was 1.90 (95% CI, 1.23 to 2.93) compared with 1.63 (95% CI, 1.04 to 2.54) for a maternal history.

Table 3 presents the ORs for both a family history of stroke at ≤65 years and a family history of MI at any age, according to stroke subtype. The strongest associations with family history parameters were seen for large-vessel disease (multivariate OR, 2.32; 95% CI, 1.68 to 3.21; P<0.001 for a family history of MI; multivariate OR, 1.67; 95% CI, 1.08 to 2.66; P=0.021 for a family history of
stroke at ≤65 years). The relationship between large-vessel disease and family history of stroke at a young age was significant regardless of the age group analyzed, although the strength of the relationship increased progressively as the age at which the stroke occurred decreased. An OR of 2.34 (95% CI, 1.21 to 4.52) (P=0.011) was seen for stroke occurring at age ≤65 years (Table 4).

Small-vessel disease was less strongly associated with a family history of MI (multivariate OR, 1.46; 95% CI, 1.05 to 2.03; P=0.025). The association with a family history of stroke at ≤65 years was just outside significance on multivariate analysis (multivariate OR, 1.49; 95% CI, 0.94 to 2.37; P=0.088) (Table 3). For small-vessel disease, a significant interaction was found between age and a family history of stroke at ≤65 years (P for interaction=0.015). The ORs progressively increased as age decreased, and for cases aged ≤65 years the multivariate OR was 2.69 (95% CI, 1.46 to 4.96) (P=0.002). When individuals aged >65 years were included, this association was greatly weakened (Table 4). In contrast to the findings for large- and small-vessel disease, no significant associations were seen for cardioembolic stroke or for ischemic stroke of undetermined etiology. (Table 3).

Table 5 shows estimated sample sizes required to detect an effect equal to the observed multivariate ORs. With the use of this model and assumption of an allele frequency of 10%, a case-control study including all unselected ischemic stroke cases would require a sample size of 1481 cases and 1481 controls. Focusing on specific stroke subtypes considerably reduced the estimated sample sizes. Including either large- or small-vessel disease cases of all ages would reduce the number of cases to 542 and 946, respectively. Further confining the study to subjects with onset of stroke at a young age (≤65 years) would have a major impact on sample size, with n=205 for all ischemic stroke and n=175 and n=123 for large- and small-vessel disease, respectively.

**Discussion**

These data suggest that a positive family history of vascular disease is an independent risk factor for both large-vessel atherosclerosis and small-vessel disease. In addition, a family history of vascular disease was a particularly strong risk factor in subjects presenting with stroke before the age of 65 years. The estimated sample sizes for case-control studies illustrate how candidate gene studies for ischemic stroke might be made more effective by focusing on these specific phenotypes, ie, small-vessel disease and large-vessel disease, and on younger age groups (≤65 years), in whom the genetic component of the disease appears to be strongest.

---

**TABLE 1. Prevalence of Vascular Risk Factors in Controls and Cases Both Overall and in the Major Stroke Subtypes**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Control</th>
<th>Case</th>
<th>LVD</th>
<th>SVD</th>
<th>CE</th>
<th>IU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total n</td>
<td>800 (100)</td>
<td>1000 (100)</td>
<td>262 (26.2)</td>
<td>232 (23.2)</td>
<td>118 (11.8)</td>
<td>296 (29.6)</td>
</tr>
<tr>
<td>Male sex, n (%)</td>
<td>490 (61.3)</td>
<td>580 (58.0)</td>
<td>173 (66.0)</td>
<td>150 (64.7)</td>
<td>60 (50.8)</td>
<td>152 (51.4)</td>
</tr>
<tr>
<td>No. of siblings, median (IQR)</td>
<td>2.0 (3.0)</td>
<td>2.0 (3.0)</td>
<td>3.0 (3.0)</td>
<td>2.0 (3.0)</td>
<td>2.0 (3.0)</td>
<td>2.0 (3.0)</td>
</tr>
<tr>
<td>Mean age, y (SD)</td>
<td>64.4 (8.7)</td>
<td>65.1 (12.8)</td>
<td>66 (9.7)</td>
<td>66 (10.5)</td>
<td>72 (14.2)</td>
<td>63 (14.3)</td>
</tr>
<tr>
<td>Arterial hypertension, n (%)</td>
<td>305 (38.1)</td>
<td>605 (60.5)</td>
<td>161 (61.5)</td>
<td>167 (72.0)</td>
<td>64 (54.2)</td>
<td>151 (51.0)</td>
</tr>
<tr>
<td>Diabetes mellitus, n (%)</td>
<td>45 (5.6)</td>
<td>137 (13.7)</td>
<td>129 (49.2)</td>
<td>86 (37.1)</td>
<td>25 (21.7)</td>
<td>99 (33.4)</td>
</tr>
<tr>
<td>Ischemic heart disease, n (%)</td>
<td>97 (12.1)</td>
<td>198 (19.8)</td>
<td>78 (29.8)</td>
<td>35 (15.3)</td>
<td>29 (24.1)</td>
<td>29 (9.8)</td>
</tr>
<tr>
<td>Hyperlipidemia, n (%)</td>
<td>109 (23.8)</td>
<td>376 (37.6)</td>
<td>129 (49.2)</td>
<td>86 (37.1)</td>
<td>30 (25.4)</td>
<td>99 (33.4)</td>
</tr>
<tr>
<td>Current smokers, n (%)</td>
<td>152 (19.0)</td>
<td>339 (33.9)</td>
<td>118 (45.0)</td>
<td>74 (31.9)</td>
<td>22 (18.6)</td>
<td>92 (31.1)</td>
</tr>
</tbody>
</table>

LVD indicates large-vessel disease; SVD, small-vessel disease; CE, cardioembolic; IU, ischemic stroke of undetermined etiology.

† P<0.05, compared with controls; ‡ P<0.001, compared with controls.

---

**TABLE 2. Family History Parameters and Ischemic Stroke**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Control, n (%)</th>
<th>Case, n (%)</th>
<th>OR (95% CI), Univariate</th>
<th>Multivariate*</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>789</td>
<td>994</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td>235 (30)</td>
<td>319 (34)</td>
<td>1.20 (0.98–1.48)</td>
<td>1.22 (0.90–1.39)</td>
</tr>
<tr>
<td>Stroke &lt;65 y</td>
<td>72 (9)</td>
<td>137 (15)</td>
<td>1.69 (1.25–2.29)‡</td>
<td>1.38 (1.01–1.90)†</td>
</tr>
<tr>
<td>MI</td>
<td>246 (31)</td>
<td>392 (42)</td>
<td>1.57 (1.29–1.91)‡</td>
<td>1.41 (1.14–1.74)†</td>
</tr>
<tr>
<td>MI &lt;65 y</td>
<td>141 (18)</td>
<td>219 (23)</td>
<td>1.39 (1.10–1.76)†</td>
<td>1.24 (0.96–1.59)</td>
</tr>
<tr>
<td>Stroke or MI</td>
<td>418 (53)</td>
<td>585 (62)</td>
<td>1.45 (1.19–1.75)‡</td>
<td>1.34 (1.09–1.64)†</td>
</tr>
<tr>
<td>Stroke or MI ≤65 y</td>
<td>203 (26)</td>
<td>316 (34)</td>
<td>1.46 (1.18–1.79)‡</td>
<td>1.26 (1.01–1.58)†</td>
</tr>
</tbody>
</table>

*Multivariate=adjusted for age, sex, arterial hypertension, diabetes mellitus, serum cholesterol, and smoking status.

†P<0.05; ‡P<0.001.
Previous studies examining the role of a positive family history in ischemic stroke risk have yielded conflicting results. The strongest evidence comes from twin concordance studies, which found a 2- to 4-fold increase in stroke risk in monozygotic versus dizygotic twin pairs. Prospective cohort studies have found a family history of stroke to be predictive of future stroke, and numerous case-control studies have shown positive associations. In contrast, others, including a large prospective cohort study of >13,000 subjects, have been negative. Positive results were more often seen in selected subgroups, for example, in younger subjects or in subjects with a family history of disease onset at a young age, suggesting that the extent to which genetic factors contribute to stroke risk may be age dependent. This is consistent with the findings of this study, in which positive associations were strongest among younger probands, and a positive association was seen for a family history of stroke at ≤65 years but not for a family history of stroke at any age.

While it has been proposed that defining specific subtypes may be key in elucidating the genetic contribution to stroke risk, few studies to date have examined the familial component of different etiologic subtypes. In a sample of 310 subjects with ischemic stroke, Meschia et al found no relationship between a positive family history of stroke and proband stroke subtype. However, this study did not include a control population and may therefore have been underpowered. A more recent case-control study (421 cases and 239 controls) found that both large- and small-vessel strokes were associated with a positive family history of stroke, with no association seen for cardioembolic stroke or for stroke of undetermined etiology, consistent with the findings of our study.

Because stroke is the end result of a number of pathologically different processes, it is possible that many genes, each conferring a small amount of risk, are involved in influencing the end phenotype. Conventional case-control candidate gene studies may not be sufficiently powerful to detect the contribution of an individual disease allele since, to date, this approach has not yielded consistent results. Our modeling data, estimating sample sizes for specific stroke subtypes, suggest that using these phenotypes may be an effective way to increase power in case-control studies. Recent studies examining the role of genetic factors in specific stroke subtypes further support this hypothesis. For example, it has been suggested that a deletion polymorphism in the angiotensin-converting enzyme gene may be a specific risk factor for lacunar but not other stroke subtypes by predisposing to lipohyalinosis. The apolipoprotein E4 allele and the A A genotype of the β-fibrinogen G/A-455 polymorphism were found to be risk factors for large-vessel disease but not other stroke subtypes, possibly mediated via an increased risk of atherosclerosis and atherothrombosis.

Other mechanisms, including familial aggregation of stroke risk factors, cannot be excluded as a potential explanation for the findings of this study. Hypertension, dyslipidemia, and diabetes mellitus have all been shown to aggregate in families affected by stroke. Although adjustment for vascular risk factors attenuated the observed associations, a positive family history remained an independent predictor of large-vessel disease and of small-vessel disease in younger subjects. This suggests that...
aggregation of risk factors cannot completely explain the relationship and that additional genetic factors may be involved.

The familial contribution to vascular disease is likely to reflect both shared genetic load and environmental exposures. The strong association seen between stroke and family history of MI in this study suggests that certain inherited etiologic factors are common to both diseases, for example, a genetic predisposition to arterial atherosclerosis in the case of large-vessel disease or to shared risk factors such as hypertension in the case of small-vessel disease. A trend was also seen toward an association between a family history of MI and cardioembolic stroke, which may reflect a shared familial risk for ischemic heart disease. The nonsignificant OR may have been affected by the lower number of cardioembolic strokes in the cohort compared with other stroke subtypes. The lack of association between a family history of stroke and cardioembolic disease may be a consequence of the heterogeneous etiology of the cardioembolic source, which may, for example, include rheumatic valvular disease, ischemic heart disease, and nonischemic cardiomyopathies.

A strength of this study is that the hospital-based cohort was extensively investigated, which permitted accurate stroke subtyping. However, this was at the expense of the exclusion of community stroke cases. Limitations of this study include possible family information bias, whereby cases are more likely to recall a relative affected by the same illness. Relatives may be more effective if they focus on these specific etiologic subtypes and on younger age groups, which appear to have a particularly strong familial component of their stroke risk.

Acknowledgments

This study was supported by grants from the Stroke Association (UK) to P. Jerrard-Dunne, G. Cloud, and A. Hassan.

References


<table>
<thead>
<tr>
<th>Table 5. Estimated Sample Size Requirements for Studies Using Specific Stroke Subtypes and Age Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR (95% CI) Associated With the Genetic Variant*</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>All ischemic stroke</td>
</tr>
<tr>
<td>All ischemic stroke ≤65</td>
</tr>
<tr>
<td>Large-vessel disease</td>
</tr>
<tr>
<td>Large-vessel disease ≤65</td>
</tr>
<tr>
<td>Small-vessel disease</td>
</tr>
<tr>
<td>Small-vessel disease ≤65</td>
</tr>
</tbody>
</table>

*ORs are based on the observed multivariate OR and 95% CIs for a positive family history of stroke ≤65 years.†Estimated sample sizes required to detect an effect equal to the observed ORs and 95% CIs, assuming allele frequencies of 5%, 10%, or 15%, with a 2-tailed significance level of 0.05 and a power of 0.8.