Ischemic Stroke in Young Women
A Nested Case–Control Study Using the UK General Practice Research Database

Alison L. Nightingale, BSc(Hons); Richard D.T. Farmer, MB, PhD, FFPHM, FFPM

Background and Purpose—Estimates of the incidence of ischemic stroke in young women vary widely from 0.9 to 8.9 per 100 000 per year. This study was conducted to determine the incidence and risk factors for ischemic stroke in young women in the UK.

Methods—Women aged 15 to 49 with a first diagnosis and supporting evidence of ischemic stroke between January 1, 1992, and December 31, 1998, were identified from the UK General Practice Research Database. Age-specific incidence rates were calculated and a nested case-control study was conducted with up to 6 controls randomly selected and matched to each case by year of birth and general practice. Crude and adjusted odds ratios (ORs) were calculated using conditional logistic regression.

Results—The incidence of ischemic stroke was 3.56/100 000 per year. Factors associated with an increased risk were heart disease (OR, 10.5), heavy alcohol consumption (OR, 8.5), previous venous thromboembolism (OR, 6.2), treated diabetes mellitus (OR, 4.7), hypertension (OR, 4.6), migraine (OR, 2.3), and use of combined oral contraceptives (OR, 2.3). Light alcohol consumption was found to be protective (OR, 0.17).

Conclusions—The crude incidence rate was lower than previously reported for the USA and Europe but higher than that reported for the UK Oxford Region. This could be because of an under-representation of mild cases or because of a true lower incidence in the UK compared with the USA and the rest of Europe. The results of the case-control study are consistent with previous studies of ischemic stroke in young women. (Stroke. 2004;35:1574-1578.)

Key Words: cerebrovascular accident ■ incidence ■ risk factors ■ case–control studies ■ longitudinal studies

The incidence of ischemic stroke in young women ranges from 4.3 to 8.9/100 000 per year1,2 in the United States and continental Europe. Incidence increases with age.3,4 “Idiopathic” ischemic stroke among women aged 20 to 44 in the UK is 0.9/100 000 per year.5

Hypertension, heart disease, and diabetes mellitus are risk factors in adults and in younger people.4,6–9 Other factors reported to increase the risk in young women are smoking,4,7,9,10 migraine (particularly classical migraine),4,11,12 antiphospholipid antibodies,13 combined oral contraceptives (COCs),4,6,7,10 and heavy alcohol consumption.4,6,9,10 Light alcohol consumption has been reported to be protective.14

This study was designed to determine the incidence of ischemic stroke in young women in the UK and to identify factors associated with an increased risk using the UK General Practice Research Database (GPRD).

Materials and Methods
The GPRD contains the anonymous general practice (GP) medical and prescribing records for just >8.5 million people registered with general practitioners throughout the UK. It has been used extensively for studies of disease epidemiology and drug safety.15 Automated data recorded by GPs as part of routine clinical practice are collected and made anonymous for the purposes of research. The owners of the database (the UK Medicines and Healthcare Products Regulatory Agency) monitor the quality of the data. Medical diagnoses are entered using standardized Oxford Medical Information System (OXMIS)17 and Read Codes.18 Prescribing data are product-specific and contain information on the dose and duration prescribed. Other data include height, weight, smoking, and alcohol consumption.

Study Population
The study period was from 1992 to 1998. The study population consisted of all women aged 15 to 49 who contributed data to the GPRD for at least 6 months at any time during the study period.

Case Definition
The study population was searched for women with a first diagnosis of cerebrovascular accident, cerebral infarction, cerebral thrombosis, cerebellar infarction, or cerebellar thrombosis and who had at least 6 months of research standard data accrued before diagnosis. In addition, all cases fulfilled 1 or more of the following criteria: (1) death caused by ischemic stroke, including sudden deaths referred to the coroner; (2) record of hospital attendance for ischemic stroke and evidence of residual damage (eg, hemiparesis, speech or visual dysfunction, leg or facial weakness, rehabilitation or new long-term therapy with warfarin or aspirin); (3) confirmation of stroke by...
computed tomography scan or magnetic resonance imaging; (4) evidence of new long-term residual hemiplegia in the absence of a second event or diagnosis of multiple sclerosis; and (5) evidence of new epilepsy or seizures after the stroke and, in the absence of a second event, head injury or cerebral tumor.

Cases were excluded if there was evidence of previous stroke, a diagnosis of cerebral tumor before or within 4 months after the event date, major trauma, or surgery within 6 weeks before the event date. We also excluded women with a diagnosis of systemic lupus erythematosus before or within 6 months after the event date or a diagnosis of multiple sclerosis at any time in their medical record because it was difficult to distinguish symptoms of cerebral systemic lupus erythematosus and multiple sclerosis from symptoms of stroke.

Cohort Study

We conducted a cohort study to calculate the incidence rates. We calculated the number of years of observation by summing the days during each woman’s valid period of observation and dividing by 365.25. The start of each woman’s period of observation was the latter of: January 1, 1992, or January 1 in the year she reached age 15 years, or 6 months after the date on which she registered with her GP, or her GP started to contribute data to the GPRD. The end of a woman’s period of data was the earlier of: January 1, 1992, or January 1 in the year she reached age 49 years, or the last date the GP contributed data to the GPRD.

Nested Case–Control Study

Up to 6 controls were randomly selected from the study population matched to each case by GP and age. The controls were subject to the same exclusion criteria as the cases. Data were abstracted from the medical records before the index date for history of: venous thromboembolism (VTE) (deep vein thrombosis or pulmonary embolism), heart disease (myocardial infarction, heart failure, valvular heart disease, or congenital heart disease), cancer, transient ischemic attacks, hyperlipidemia, treated diabetes, epilepsy (diagnosis or prescription of a specific antiepileptic medication in the absence of any other explanatory diagnosis between 2 years and 6 months before the index date), or hypertension (diagnosis or diastolic blood pressure ≥95 mm Hg before index date). A woman was deemed to have a history of migraine if she had a prescription for a specific antimigraine therapy or diagnosis of migraine and a prescription for a potential antimigraine medication or analgesics in the absence of any other explanatory diagnosis within 1 year before the index date. It was not possible to differentiate between classical and simple migraine.

We classified each study subject according to her reproductive status on the index date. Women were classified as pregnant if they were pregnant on the index date or had delivered within 6 weeks before the index date. COC use patterns had been previously mapped precisely for each woman on the GPRD, and it was possible to determine whether women had a prescription for a COC covering the index date and to record the formulation to which they were exposed. Women were classified as menopausal if they had a prescription for hormone replacement therapy, record of menopause, climacteric symptoms, or bilateral oophorectomy before the index date. Women who were not pregnant, menopausal, or using COCs were classified as being premenopausal and not using COCs.

When possible, body mass index (BMI) (kg/m²) was calculated for each woman from the height and weight records nearest to her index date. Smoking status before the index date was extracted, as was alcohol consumption. Alcohol consumption was partitioned into the following categories: light (1 to 9 U/week), moderate (10 to 21 U/week), heavy (≥22 U/week or a record of alcohol abuse), and unknown. There are 8 g of alcohol per “unit” in the UK.

Statistical Analysis

All analyses were conducted using STATA (Stata Statistical Software version 7.0). We offered variables to forwards and backwards stepwise logistic regression models set to accept or reject variables at a significance of $P=0.05$. Variables that were extracted from the medical records but not offered into the model were history of transient ischemic attack and hyperlipidemia because there were insufficient data available. We conducted $\chi^2$ tests on each variable offered to the stepwise logistic regression models and calculated the unadjusted ORs using univariate conditional logistic regression. Adjusted ORs were calculated for all variables accepted by the stepwise logistic regression models using multivariate conditional logistic regression.

Results

Cohort Study

We identified 190 incident cases of ischemic stroke during the study period. There were 5 336 721 years of observation on the GPRD yielding a crude incidence rate of 3.56/100 000 per year (CI95, 3.05 to 4.07). Incidence increased with age (Table 1). To compare the results of this study with those reported for Oxford, UK, we calculated the incidence rate for women aged 20 to 44 and excluded all women who were pregnant or 6 weeks postpartum on their index date and those with a history of myocardial infarction or VTE. The resulting crude incidence rate in these women (n=85) was 1.59/100 000 per year (CI95, 1.25 to 1.93). There were 29 fatalities, yielding a case fatality rate of 15.3%.

Nested Case–Control Study

Controls were matched to all cases with a mean of 5.8 controls per case (range, 1 to 6). The $\chi^2$ test results and unadjusted and adjusted ORs calculated using univariate and multivariate conditional logistic regressions are shown in Table 2. Cancer, BMI, and epilepsy were not accepted by the stepwise logistic regression models and were not included in the multivariate conditional logistic regression model. BMI was strongly correlated with hypertension in our population. Within the multivariate model, hypertension was the variable explaining most of the variance; therefore, hypertension was in the final model rather than BMI. We found a significantly increased risk associated with heavy alcohol consumption, current smoking, heart disease, history of VTE, current COC use, hypertension, migraine, and treated diabetes. The risk among women with moderate alcohol consumption was not different from that of nondrinkers, and light alcohol consumption was associated with a decreased risk. The adjusted OR for ischemic stroke among COC users was 2.30 (CI95,
1.15 to 4.59); there were too few women to allow the assessment of the effects of different formulations. We found no significant interaction between variables in the multivariate conditional logistic regression model.

Discussion

We estimated that the incidence of ischemic stroke in women aged 15 to 49 years in the GPRD population was 3.56/100 000 per year; incidence increased with increasing age. The crude and age-specific incidence rates were lower than those reported for young women in the United States (4.3 to 5.4/100 000 per year)\(^{1,3,8}\) and in continental Europe (8.9/100 000 per year).\(^{2}\) They were higher than the rate reported by the World Health Organization (WHO) study of cardiovascular disease and steroid hormone contraception for the Oxford region in the UK (0.9/100 000 per year).\(^{5}\) It is possible that women in the UK have a lower incidence of ischemic stroke than those in the USA and in continental Europe. Our method of case identification may have resulted in an underestimate of the true incidence. All cases had strong evidence to support the diagnosis and it is unlikely that false-positive cases were included. However, cases of mild stroke without a record of residual damage or rehabilitation, or when hospital admission or the results of scans were not recorded, would not have fulfilled our inclusion criteria. Such mild cases will have been missed, leading to an underestimate of the incidence rate.

The results of the case-control study are consistent with previous studies. We confirmed the association with risk factors such as hypertension, diabetes, heart disease,\(^{4,7,8}\) smoking, and heavy alcohol consumption. Light alcohol consumption was protective. Current COC use and history of migraine were associated with an increased risk. A history of VTE was associated with an increased risk; this has not been reported previously and might be a proxy indicator for abnormalities in coagulation.

The prevalence of diabetes among the controls was 0.9%; this is comparable with other estimates in the UK.\(^{21}\) We found an increased risk of ischemic stroke associated with diabetes. The CI was wide because of the small number of cases and

<table>
<thead>
<tr>
<th>Table 2. Variables Associated With Ischemic Stroke</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient Characteristic</strong></td>
</tr>
<tr>
<td>Preexisting heart disease†</td>
</tr>
<tr>
<td>Treated diabetes†</td>
</tr>
<tr>
<td>Hypertension†</td>
</tr>
<tr>
<td>Previous VTE†</td>
</tr>
<tr>
<td>History of migraine†</td>
</tr>
<tr>
<td>Alcohol consumption</td>
</tr>
<tr>
<td>None</td>
</tr>
<tr>
<td>Light (1–9 units/week)</td>
</tr>
<tr>
<td>Moderate (10–21 units/week)</td>
</tr>
<tr>
<td>Heavy (22+ units/week)</td>
</tr>
<tr>
<td>Unknown</td>
</tr>
<tr>
<td>Smoking status</td>
</tr>
<tr>
<td>Nonsmoker</td>
</tr>
<tr>
<td>Current smoker</td>
</tr>
<tr>
<td>Unknown</td>
</tr>
<tr>
<td>Reproductive status</td>
</tr>
<tr>
<td>Premenopausal (no COCs)</td>
</tr>
<tr>
<td>Pregnant</td>
</tr>
<tr>
<td>Current COC user</td>
</tr>
<tr>
<td>Menopausal</td>
</tr>
<tr>
<td>History of cancer†</td>
</tr>
<tr>
<td>History of epilepsy</td>
</tr>
<tr>
<td>Body Mass Index (kg/m(^2))</td>
</tr>
<tr>
<td>&lt;20–24.9</td>
</tr>
<tr>
<td>25–29.9</td>
</tr>
<tr>
<td>30+</td>
</tr>
<tr>
<td>Unknown</td>
</tr>
</tbody>
</table>

\(^{df}\), degrees of freedom.
\(^{†}\)Reference group: cases and controls without the specified disease.
\(^{‡}\)1 unit \(\sim 8\) g of alcohol.
controls with diabetes. Our findings are consistent with those of Lidegaard and Kreiner, the WHO, and Petitti et al who reported ORs of 5.25 (CI95, 1.71 to 16.11), 2.60 (CI95, 0.74 to 9.12), and 7.15 (CI95, 3.17 to 16.13), respectively. The increased risk associated with hypertension is also consistent with studies by Lidegaard and Kreiner, the WHO, and Petitti et al who reported ORs of 3.67 (CI95, 1.90 to 7.11), 3.38 (CI95, 1.88 to 6.08), and 7.79 (CI95, 3.15 to 17.31), respectively.

Our study is consistent with other studies of ischemic stroke in young women that have reported that the risk in current smokers is ~2- to 3-times that in nonsmokers. Classification of alcohol consumption varies widely from study to study, as does the reported risk associated with it. We found an increased risk associated with heavy alcohol consumption of >21 U/week (~168 g of alcohol/week) and a decreased risk associated with light alcohol consumption of 1 to 9 U/week. The WHO reported that alcohol consumption was not a significant factor in the risk of ischemic stroke. Similarly, Petitti et al reported no significant association between alcohol consumption and ischemic stroke; however, the heaviest consumption category was 1 to 3 drinks per week. Malarcher et al found no increased risk of ischemic stroke associated with the consumption of >24 g per day. Haapaniemi et al investigated the effect of acute heavy alcohol consumption in the 24 hours before stroke and reported an increased risk associated with >40 g with an OR of 5.68 (CI95, 1.75 to 18.48). You et al reported that heavy alcohol consumption of 60 g or more of alcohol per day was associated with an increased risk of stroke (OR, 15.3; CI95, 1.0 to 232.0). We could not classify the study subjects for smoking status and 35% for alcohol consumption. However, the “unknown” categories for both variables were not different from the reference categories in the regression models, suggesting that this has not led to distorted risk estimates.

Current users of COCs were found to have an increased risk of ischemic stroke (OR, 2.50; CI95, 1.15 to 4.59). This finding is consistent with that of the Transnational Study (OR, 2.89; CI95, 2.02 to 4.04), the WHO Study (European centers OR, 2.99; CI95, 1.65 to 4.50), and the study by Haapaniemi et al (OR, 4.19; CI95, 1.74 to 10.11). Three other studies have reported no significant association between current use of COCs and ischemic stroke. We found no increased risk associated with surgical or natural menopause at age younger than 49 or with pregnancy, and we have not identified any reports of such an association.

The increased risk associated with migraine (OR, 2.33; CI95, 1.04 to 5.21) is consistent with previously reported ORs of 3.25 (CI95, 1.47 to 3.78) to 3.54 (CI95, 1.30 to 9.61). It was not possible to differentiate between simple and classical migraine in our analysis. Using our definition, the prevalence of migraine among the cases was 8.4%, and among the controls it was 3.9%. Studies in which questionnaires or interviews based on the International Headache Society’s criteria for migraine have been used to ascertain the presence or absence of migraine have reported the prevalence to be between 29.9% and 60% in the cases and 11.8% to 30% in the controls. The use of structured interviews would inevitably lead to the diagnosis of migraine in women who had not consulted their doctor for migraine or who had not had migraine previously diagnosed. We would not have identified such women. It is possible that there is residual confounding from under-identification of migraine among the study subjects.

Many epidemiological studies rely on information regarding medical history and use of medications such as COCs. Database studies are not subject to recall bias because data are entered at the time of consultation, but they are subject to information bias. Although the GPRD is a research database, data are not collected to investigate a specific hypothesis; they are medical records from a clinical setting and deemed to be of importance to the management of the patient. We found a low prevalence of migraine compared with other studies designed to prospectively identify women with migraine, and this could be an example of such information bias. In addition, we were unable to control for potential risk factors such as genetic disposition, diet, and coagulation abnormalities (e.g., antiphospholipid antibodies). Therefore, there may be some residual confounding.

Acknowledgments and Funding

This study was supported by an unconditional grant from NV Organon and Schering AG. The sponsors had no influence on the design of the study, the analysis, or the preparation of this paper.

References

12. Tzourio C, Tehindranarivelo A, Iglesias S, Alperovitch A, Chedru F, d’Anglejan-Chatillon J, Bousser MG. Case-control study of
Ischemic Stroke in Young Women: A Nested Case–Control Study Using the UK General Practice Research Database
Alison L. Nightingale and Richard D.T. Farmer

*Stroke*. 2004;35:1574-1578; originally published online May 13, 2004;
doi: 10.1161/01.STR.0000129789.58837.e4

*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2004 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/35/7/1574