GRECOS project. The use of genetics to predict the vascular recurrence after stroke

Israel Fernández-Cadenas, PhD[1,2,*], **Maite Mendióroz, MD, PhD**[2,3,*], **Dolors Giralt, MsC**[2], **Cristina Nafria, MsC**[2], **Elena Garcia, MsC**[1], **Caty Carrera, MD**[2], **Cristina Gallego-Fabrega, PhD**[1], **Sophie Domingues-Montanari, PhD**[2], **Pilar Delgado, MD, PhD**[2], **Marc Ribó, MD, PhD**[4], **Mar Castellanos, MD, PhD**[5], **Sergi Martínez, MD, PhD**[6], **Mari Mar Freijo, MD, PhD**[7], **Jordi Jiménez-Conde, MD, PhD**[8], **Marta Rubiera, MD, PhD**[4], **José Alvarez-Sabín, MD, PhD**[4], **Carlos A. Molina, MD, PhD**[4], **Maria Angels Font, MD**[9], **Marta Grau Olivares, PhD**[10], **Ernest Palomeras, MD, PhD**[11], **Natalia Perez de la Ossa, MD, PhD**[12], **Maite Martinez-Zabaleta, MD, PhD**[13], **Jaime Masjuan, MD, PhD**[14], **Francisco Moniche, MD, PhD**[15], **David Canovas, MD, PhD**[16], **Carlos Piñana, MD, PhD**[17], **Francisco Purroy, MD, PhD**[18], **Dolores Cocho, MD, PhD**[19], **Inma Navas, MD, PhD**[20], **Carlos Tejero, MD, PhD**[21], **Nuria Aymerich, MD, PhD**[22], **Natalia Cullell, MsC**[1], **Elena Muiño, MD**[1], **Joaquín Serena, MD, PhD**[5], **Francisco Rubio, MD, PhD**[9], **Antoni Davalos, MD, PhD**[12], **Jaume Roquer, MD, PhD**[8], **Juan Francisco Arenillas, MD, PhD**[23], **Joan Martí-Fábregas, MD, PhD**[6], **Keith Keene, PhD**[24], **Wei-Min Chen, PhD**[25], **Bradford Worrall, MD, MsC**[26], **Michele Sale, PhD**[27], **Adrià Arboix, MD, PhD**[10], **Jerzy Krupinski, MD, PhD**[20], and **Joan Montaner, MD, PhD**[2,*] on behalf of the GRECOS study group

[1]Stroke Pharmacogenomics and Genetics, Fundacio Docència i Recerca MutuaTerrassa, Terrassa, Spain

[2]Neurovascular Research Laboratory and Neurovascular Unit. Neurology and Medicine Departments-Universitat Autònoma de Barcelona. Vall d'Hebrón Hospital, Barcelona, Spain

[3]Neuroepigenetics Research Group. Navarrabiomed, Pamplona, Spain

[4]Neurology Department, Vall d'Hebrón Hospital, Barcelona, Spain

[5]Department of Neurology, Hospital Universitari Dr Josep Trueta, Institut d'Investigació Biomèdica de Girona, Girona, Spain

[6]Department of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

[7]Department of Neurology, Hospital de Basurto, Bilbao, Spain

[8]Department of Neurology, Hospital del Mar and Centre de Regulació Genòmica, Universitat Pompeu Fabra, Barcelona, Spain

[9]Department of Neurology, Hospital de Bellvitge, Hospitalet del Llobregat, Spain

---

[±]Address correspondence: Israel Fernandez: Stroke Pharmacogenomics and Genetics, Fundació Docència i Recerca MutuaTerrassa, Sant Antoni 19, 08221, Terrassa, Spain. israelcadenas@yahoo.es, phone number: +34651184139.
[*]These authors contributed equally.

[10]Department of Neurology. Hospital Universitari Sagrat Cor, Barcelona, Spain

[11]Department of Neurology. Hospital de Mataró. Mataró, Spain

[12]Department of Neurology. Hospital Universitari Germans Trias i Pujol. Badalona, Spain

[13]Department of Neurology. Hospital Donostia. San Sebastián, Spain

[14]Department of Neurology. Hospital Ramón y Cajal. Madrid, Spain

[15]Department of Neurology. Hospital Universitario Virgen del Rocío. Sevilla, Spain

[16]Department of Neurology. Corporació Sanitària Parc Taulí. Sabadell, Spain

[17]Department of Neurology. Hospital Universitario Virgen de las Nieves. Granada, Spain

[18]Department of Neurology. Hospital Universitario Arnau De Vilanova. Lleida, Spain

[19]Department of Neurology. Hospital de Granollers. Granollers, Spain

[20]Department of Neurology. Hospital Mútua de Terrassa. Terrassa, Spain

[21]Department of Neurology. Hospital Clinico Universitario Lozano Blesa. Zaragoza, Spain

[22]Department of Neurology. Hospital De Navarra. Pamplona, Spain

[23]Department of Neurology. Hospital Clínico Universitario de Valladolid. Valladolid, Spain

[24]Department of Biology; Center for Health Disparities, East Carolina University, Greenville, NC, USA

[25]Department of Public Health Sciences; Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

[26]Department of Neurology; Department of Public Health Sciences, University of Virginia, Charlottesville, VA, USA

[27]Department of Public Health Sciences; Department of Biochemistry & Molecular Genetics; Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

Abstract

Background and Purpose—Vascular recurrence occurs in 11% of patients during the first year after ischemic stroke (IS) or transient ischemic attack (TIA). Clinical scores do not predict the whole vascular recurrence risk, therefore we aimed to find genetic variants associated with recurrence that might improve the clinical predictive models in IS.

Methods—We analyzed 256 polymorphisms from 115 candidate genes in three patient cohorts comprising 4,482 IS or TIA patients. The discovery cohort was prospectively recruited and included 1,494 patients, 6.2% of them developed a new IS during the first year of follow-up. Replication analysis was performed in 2,988 patients using SNPlex or HumanOmni1-Quad technology. We generated a predictive model using Cox regression (GRECOS score), and generated risk groups using a classification tree method.

Results—The analyses revealed that rs1800801 in the *MGP* gene (HR: 1.33, $p= 9\times10^{-03}$), a gene related to artery calcification, was associated with new IS during the first year of follow-up. This

polymorphism was replicated in a Spanish cohort (n=1.305), however it was not significantly associated in a North American cohort (n=1.683). The GRECOS score predicted new IS ($p=$ 3.2×10^{-09}) and could classify patients, from low risk of stroke recurrence (1.9%) to high risk (12.6%). Moreover, the addition of genetic risk factors to the GRECOS score improves the prediction compared to previous SPI-II score ($p=0.03$).

Conclusions—The use of genetics could be useful to estimate vascular recurrence risk after IS. Genetic variability in the *MGP* gene was associated with vascular recurrence in the Spanish population.

Keywords

stroke recurrence; genetics

1-Introduction

Stroke remains a major cause of death and disability. In the European Union estimates showed 2,700,000 recurrent stroke cases, and 536,000 incident stroke cases per year (1, 2). The risk of stroke recurrence after a first stroke is high, especially at the early stages, being around 6–12% within the first year of the initial stroke. Moreover, stroke patients also confront a high risk of developing other vascular diseases such as acute myocardial infarction and vascular death. Data suggest that within 10 years of ischemic stroke or Transient Ischemic Attack (TIA), around 60% of patients will die and 54% will experience a new vascular event. In the European Union countries, the total annual direct and indirect costs of stroke are estimated at €27 billion (3, 4). Unfortunately, we cannot clearly predict individualized stroke recurrence risk. Therefore, identification of patients at highest risk of recurrent stroke, including through the development of specific measures for early detection and prevention, remains a high priority. Several modifiable, lifestyle-associated risk factors such as physical inactivity, smoking, and alcohol consumption are implicated in vascular recurrence (5). Moreover, age, hypertension, atrial fibrillation, hyperlipidemia, homocysteine levels and diabetes also increase risk of stroke recurrence (6). Two main scores including clinical variables have been generated to predict vascular recurrence after ischemic stroke, such as the Essen Stroke Risk Score (ESRS) (7) and the Stroke Prognosis Instrument (SPI-II) (8), however the predictive ability of these models is limited and not useful for clinical practice (9).

Although genetic background explains some of the inter-individual differences in cardiovascular events recurrence, very few studies, with a sample size higher than 1000 subjects, have analyzed the role of individual genetic background in the risk of vascular recurrence among first-ever ischemic stroke patients (10–13).

We aimed to identify genetic variants associated with vascular and with ischemic stroke recurrence in order to estimate the ability of a clinical-genetic model to predict cardiovascular recurrence after a first-ever ischemic stroke.

2-Methods

2.1 Study Population

The Discovery cohort ("Cohort-A") (n = 1494) consisted of consecutive Caucasian patients with a first ischemic stroke that were admitted to the emergency room of one of the 23 participant Spanish hospitals (supplemental data) between July 2005–May 2009. Replication cohort B (n = 844) and C (n = 461) were recruited at Vall d'Hebron Hospital and 5 more Spanish hospitals as a part of GenotPA project between November 2000–May 2005 (14). Replication cohort D (n = 1683) was comprised of European-ancestry individuals recruited between 1997–2001 in North America and Scotland as a part of the Vitamin Intervention for Stroke Prevention (VISP) clinical trial (15). We used a prospective cohort study design, and considered patients with a vascular recurrence as cases and patients without a vascular recurrence as controls. We analyzed two end-points: 1-Ischemic stroke recurrence; and 2-Vascular recurrence (new ischemic stroke, myocardial infarction, peripheral vascular disease or cardiovascular death). This study was approved by local ethics committee and written informed consent was obtained from all individuals.

2.2 Clinical Protocol

Included patients from Cohort A met the following criteria: First episode of persistent focal neurological deficit of less than 1 h (TIA) or more than 1 h with Neuroimaging (CT or MRI) confirming an ischemic stroke, ruling out other pathology for both TIA and ischemic stroke. We excluded individuals who have previously been diagnosed with a stroke, individuals with mRS at discharge 4 and with a life expectancy <1 year at the time of inclusion and patients participating in a clinical trial of secondary prevention of stroke. Patients with a presence of diseases that are associated with a life expectancy of less than one year, such as cancer or AIDS in advanced stages, were excluded. On the basis of their experience and information provided by other physicians, the neurologists decided whether to include or exclude each patient from the study.

We enrolled the patients during the first 24h after stroke and performed a follow-up for at least one year to record the occurrence of ischemic stroke recurrence or global vascular recurrences. Ischemic stroke recurrence was defined by the occurrence of a new ischemic stroke following first-ever stroke. Vascular recurrence was defined by a new ischemic stroke, myocardial infarction, peripheral vascular disease or cardiovascular death. We used standard questionnaires to capture follow-up clinical and demographic data for each participant during phone calls performed every three months, and performed Morisky Green tests for treatment compliance (16).

Replication cohorts (Cohorts B–C) were recruited as a part of the Genot-PA study (14), clinical data regarding ischemic stroke were assessed in the emergency room or during hospitalization, and vascular recurrence data and follow-up data were obtained by reviewing the clinical records and by telephone interview. The VISP trial was a multi-center clinical trial that enrolled patients aged 35 with a non-disabling cerebral infarction defined as an ischemic brain infarction not due to embolism from a cardiac source. The trial was designed to determine if daily intake of a multivitamin tablet reduced ischemic stroke recurrence

events. The recurrence was detected by telephone contacts and in-clinic visits every 3 months for up to 2 years. Patients were excluded if the presence of cancer, pulmonary disease, or another illness was observed which, in the opinion of the study physician, would limit the life expectancy of the patient to less than two years.

The VISP cohort differed from the discovery Cohort A with regard to the B vitamin intervention and exclusion of patients with TIAs or cardioembolism.

2.3 Genetic Analysis

We identified 256 SNPs in 115 genes selected from the literature in May 2010 (methods in supplemental data) related to angiogenesis, stroke, hypertension and other pathways associated with stroke (Supplemental Table I).

Genotyping was performed using SNPlex™ technology (Applied Biosystems, Foster City, California). For quality control, two HapMap samples (NA10860/NA10861) were included, and their genotype concordance was verified. The 256 SNPs reached the minimal call rate.

VISP samples (Cohort-D) were genotyped at the Johns Hopkins Center for Inherited Disease Research, using the Illumina HumanOmni1-Quad_v1-0_B BeadChip (Illumina, San Diego, CA), (methodology in supplemental data).

2.4 Statistical Analysis

Sample size calculation was performed using Ene 2.0 software (http://www.e-biometria.com). One hundred and ten subjects were needed to detect SNP frequencies >0.30 in the experimental group (patients with recurrence) and <0.15 in the reference group (patients without recurrence) to achieve statistical power of 80% with p-value = 0.05.

Statistical analysis was performed using SPSS software, v.15 (IBM, Chicago, Illinois). Statistical significance for each SNP in the discovery cohort A was assessed by Kaplan-Meier curves using log-rank test. Multivariable analysis was performed using Cox regression in cohort A. Statistical significance for each SNP in the replication cohorts B, C and D were assessed by chi-square analysis or Fisher's exact test.

Continuous variables were compared by analysis of variance (ANOVA), Mann-Whitney or Kruskal-Wallis tests. We generated a predictive score based on cox regression (CR) beta coefficients using the data from cohort A, using a forwards stepwise procedure, with a p-value of 0.05 as the threshold for entry.

Receiver operating characteristics (ROC) curves were plotted, and predictive capacity was calculated by measuring the area under the curve (AUC). The different AUCs were compared using z-test (17) from MedCalc version 9.2.0.1 (Mariakerke, Belgium).

The decision tree algorithm was generated with SPSS classification tree test (SPSS software, IBM, Chicago, Illinois). The association of the nodes with ischemic stroke recurrence was tested with the chi-square test using SPSS (SPSS software, IBM, Chicago, Illinois).

The clinical variables included in the GRECOS score were those statistically (p<0.05) associated with stroke recurrence after COX regression and the SNPs included in the GRECOS score were those statistically associated with stroke recurrence after replication analysis (p<0.05).

3-Results

In discovery cohort A, new ischemic strokes were observed in 6.2% of patients during the first year of follow-up and global vascular recurrence was observed in 10.2%. Clinical and demographic data are available based on their availability in at least the 80% of the cohort (Supplemental Table II). Eleven variables were nominally associated with recurrence of new ischemic stroke events (p<0.1) and were included in the Cox regression analysis (Table 1).

3.1 Genotyping data

Ischemic Stroke Recurrence Association Studies—In the discovery stage (Cohort A, n=1494), 17 SNPs were associated with the occurrence of a new ischemic stroke under a dominant-recessive model (p<0.05). However, none of these SNPs remained significant after replication stages (Supplemental Table III).

Under an additive model, 60 SNPs were associated (p<0.05) with the occurrence of a new ischemic stroke in the discovery stage (Cohort A) (Supplemental Table IV). After replication stages (in Cohort B, n= 844 and Cohort C, n= 461; 476 cases and 829 controls) only rs1800801 in *MGP* gene was associated with the risk of new ischemic stroke following a genetic additive model: Discovery cohort: $p= 9\times10^{-03}$; Replication cohort (cohort B and C): 1.48 OR (95% CI, 1.04–1.81); $p=0.025$. Replication analysis within VISP cohort (cohort D) (n= 1683) revealed that rs1800801 was not associated with stroke recurrence in a non-Spanish cohort (p>0.05). However the Kaplan-Meier curves performed with prospective cohort A showed a clear cumulative risk of recurrent ischemic strokes during the time of follow-up (Figure 1). The rs1800801 SNP was only imputed in the cohort D using MACH version 1.0.16.

Vascular Recurrence Association Study—Under a dominant-recessive model and under the additive model, 13 and 53 SNPs, respectively were associated with vascular recurrence. None of them remained significant after replication stages (Supplemental Table III).

3.2 Ischemic Stroke Recurrence Clinical Predictors

Univariate analysis in Cohort A identified eleven clinical variables (Table 1) nominally associated with the recurrence of new ischemic stroke events (p<0.1) from the 50 analyzed (Supplemental Table II). We included these 11 variables in a Cox regression analysis, along with rs1800801, to determine whether rs1800801 was independently associated with ischemic stroke recurrence and to generate the clinical-genetics score. We selected the median age (71 years) as cut-off for the analysis.

After Cox regression only TIA at enrollment, median age above 71 years and previous acute myocardial infarction or angina prior to enrollment were associated with ischemic stroke

recurrence during the first year following the first ischemic stroke or TIA. When we introduced the rs1800801 SNP in the Cox regression, this SNP was a predictor of ischemic stroke recurrence (rs1800801, HR: 2.26 [1.462–3.488]; $p = 1.27\ 10^{-4}$).

3.3 Ischemic stroke recurrence score generation

A pilot predictive score from the Cox regression model was generated to determine the usefulness of the combination of clinical and genetic variables to predict ischemic stroke recurrence. We included age, TIA at enrollment, previous acute myocardial infarction or angina and rs1800801 (G allele) in the following formula based on the independent variables obtained by Cox regression: Score = $(1.85 \times$ Age>71 years) + $(2 \times$ inclusion TIA) + $(3.6 \times$ AMI/ANGINA) + $(2.26 \times$ rs1800801 (G allele)). The coefficient for each variable was the Hazard Ratio value from the COX regression. This combined genetic and clinical model (the GRECOS score) predicts ischemic stroke recurrence ($p = 3.2 \times 10^{-09}$) with an AUC of 0.684 [0.569–0.675].

The GRECOS score had higher predictive capacity than the previously reported SPI-II score or ESRS score, although the difference with ESRS was not significant (AUC-GRECOS score: 0.684 vs. AUC-SPI-II: 0.601, $p = 0.03$ and AUC-GRECOS score: 0.684 vs. AUC-ESRS: 0.622, $p = 0.11$). The 3 clinical variables included in the GRECOS score alone were slightly but not significantly different from the classical scores (AUC-clinical variables alone: 0.622 vs. AUC-SPI-II: 0.601, $p = 0.9$); AUC-clinical variables alone:0.622 vs. AUC-ESRS: 0.622 $p=0.9$). As shown by AUC values, neither genetic data nor clinical data alone significantly increased the predictive capacity compared to the SPI-II or ESRS scores, only the combination of clinical and genetic variables improves the capacity of prediction.

3.4 Classification tree

We included the three clinical variables and the rs1800801 SNP in the classification tree algorithm implemented in the SPSS software (Figure 2). We observed that 12.6% of patients harboring allele G of rs180081 (risk allele) and with previous angina or acute myocardial infarction (Node 7) suffered a new ischemic stroke during follow-up (Figure 2). Interestingly, only 1.9% of patients without previous angina/acute myocardial infarction, without TIA at enrollment and harboring allele A of rs180081 (Node 8) suffered a new ischemic stroke.

The association of all nodes with ischemic stroke recurrence was very significant (p value=5.58×10^{-5}). In addition, when we analyzed the node with the highest risk (node 7) versus the others, the results were significant (p value=0.023), as were the results when we compared the node with the lowest risk (node 8) versus the others (p value= 2.8×10^{-4}).

4-Discussion

The clinical-genetic score generated with three clinical variables (age >71 years, TIA at the enrollment and previous AMI/Angina) and rs1800801 (GRECOS score) improved AUC results versus previously reported clinical scores SPI-II or ESRS, with a statistically significant improvement in relation to the SPI-II score. However when we used only the clinical variables, the GRECOS score did not show statistically significant differences versus

the SPI-II score indicating that the inclusion of genetic risk factors in clinical scores can improve the prediction of stroke recurrence. The scores generated to predict vascular recurrence in stroke patients have low predictive values, questioning the usefulness of these risk scores in daily clinical practice (18), afterwards the increase of the prediction using genetic biomarkers could be important for a clinical implementation.

Furthermore, the GRECOS classification tree indicates a major patient risk group (Node 7) harboring rs180081 allele G (risk allele) and with previous angina/acute myocardial infarction that showed a 12% risk of ischemic stroke recurrence, which is double the mean of all ischemic stroke patients (6%). In contrast, we observed a minor patient risk group with a recurrence risk of only 1.9%, one third of the risk for all the patients. If these results are replicated in an independent study, it could be very useful from a clinical point of view. A neurologist could modify the treatment of this group of high risk patients, for example by increasing the anti-aggregation doses or using double anti-aggregation (aspirin and clopidogrel).

An international replication did not find any association between rs1800801 and new recurrent ischemic strokes, suggesting that this SNP may be associated with stroke recurrence only in the Spanish population. However, the VISP project (cohort D) was a clinical trial in which the patients were treated with vitamin interventions, introducing a potential confounding factor that could influence the genetic results. Other differences among VISP and the Spanish cohorts are the higher frequency of genetically influenced clinical risk factors, including hypertension and diabetes mellitus (Supplemental Table V), compared to the GRECOS cohort. Moreover, VISP enrolled participants between 1997 and 2001, prior to the majority of individuals enrolled in the Spanish populations and during a time of transition in secondary stroke prevention. Another problem might be the sample size needed for the replication and the fact that in the VISP project patients with TIA or cardioembolism were not enrolled. Additional studies will be needed to determine whether the association of rs1800801 SNP with ischemic stroke recurrence is observed in other populations. The rs1800801 variant is located in the promoter region of the *MGP* gene. This gene encodes a matrix Gla protein which is an inhibitor of bone formation. The *MGP* −7A>G (rs1800801) polymorphism has been associated with higher risk of coronary artery calcification (19, 20). The functional relevance of this SNP is controversial: it is located at a possible transcription start site (19), but its effects on matrix Gla protein expression and serum levels are unclear (19–21). Several authors hypothesized that *MGP* −7A>G is not a functional variant and that the amino acid change Thr83Ala, in linkage disequilibrium with *MGP* −7A>G (D′ = 0,97), might be responsible for MGP changes (19). The Thr83Ala substitution changes protein polarity and could affect its capacity to bind to calcium, leading to calcium deposition in the arterial wall (19), however the effect of rs1800801 has not yet been established.

Matrix Gla protein may play a protective role in atherosclerosis progression (21). It has been found in plaque deposits and its expression is increased in atherosclerotic plaques. Moreover, a knock-out mouse model showed intense arterial calcification and death within the first two months (22). *MGP* is a mineralization-associated gene, as it is *OPN*, which encodes osteopontin, a protein previously reported to be associated with poor prognosis after

stroke by our group (23). Both *MGP* and *OPN* genes are up-regulated through the ERK1/2 pathway (24) and are expressed at calcification sites within atherosclerotic lesions (25).

Previous studies have found SNPs associated with vascular recurrence in stroke (10–13). Two studies performed with Asian population (10, 11) observed associations of *ANRIL* and *NINJ2* polymorphisms with vascular recurrence. SNP rs10757278 of the *ANRIL* gene was not associated with stroke recurrence in our cohort. However, we did not check *NINJ2* SNPs because previously unpublished results did not find any association between this locus and ischemic stroke in Spanish population. These two studies analyzed less than five SNPs and were performed in Asian populations, thus it is possible that these polymorphisms are only genetic risk factors for Asian populations. Two other studies have been performed in Caucasian population (12, 13). One of these studies recently found an association of three *CRP* polymorphisms with recurrent stroke. In our study we did not analyze these three polymorphisms due to the selection of the genes and polymorphism for our project was performed previously to the publication of the mentioned study (12).

Only in one paper (13) the potential role of genetics to predict recurrent strokes has been calculated. In this paper the authors did not find any improvement in the use of the genetics to predict the occurrence of new recurrent ischemic strokes compared to the use of clinical factors. However the authors only analyzed SNPs associated with non-recurrent ischemic stroke from previous studies. These results are in concordance with our results indicating that genetic risk factors associated with recurrent strokes may be different than genetic risk factors associated with non-recurrent strokes. In contrast, in our study with different SNPs selection criteria we have observed a prediction improvement when the genetic risk factors are included in a score. The differences observed in the two studies can be explained regarding the strategy followed by the two groups. In our study we have found the SNPs associated with stroke recurrence, as a first stage, and secondly we have performed the score with the significant genetics variables, however the study of Achterberg et al. performed the genetics score with SNPs associated previously with ischemic stroke without checking the association with recurrent strokes. Based on our results, we think that genetics could be useful for stroke recurrence prediction when more studies will be performed in the field, including Genome Wide Association Studies (GWAS).

Limitations

Replication cohorts were not recruited prospectively as it has been performed with the discovery cohort. Prevention strategies in stroke have changed during the last ten years but these data are difficult to quantify, but we make the assumption that cases with recurrences and cases without recurrences recruited in the same time period received comparable treatment and prevention strategies. Sub-analyses depending on the subtype of ischemic stroke were not performed due to the limited sample size. It would be very interesting to test the GRECOS score in the other replication cohorts. However, one of the main clinical variables associated with stroke recurrence is previous myocardial infarction (before first stroke) and this data was not included in the replication cohorts. The MGP polymorphism was not significant after Bonferroni correction. However, we think that Bonferroni penalizes gene candidate studies too harshly, and the SNP was replicated in independent cohorts.

Bonferroni is a useful multivariable test correction in several situations; however, the Bonferroni method is conservative and has low power (26, 27). It has been suggested that Bonferroni is a useful method in situations in which several subgroups of analysis have been tested, such as multi-stratification of the cohort (selection of women only, younger subjects only, etc.) or in "omics studies" with no selection of genes, proteins, or clinical factors with a biologically plausible association with the end-point of the study (26). We think replication is the most important phase of the study to validate the results obtained in gene candidate analysis (28).

We did not count the number of SNPs discarded during the first search of SNPs. The data could be valuable for future studies, but unfortunately we do not have this information. Economic studies could estimate the usefulness of including this SNP among the clinical tests

Summary/Conclusions

The generation of a score using clinical and genetic variables to predict ischemic stroke recurrence has revealed its potential to be useful in the future as a diagnostic tool during the follow-up of ischemic stroke patients.

Accurate estimates of vascular recurrence risk would allow clinicians to individualize secondary preventive treatment and to perform an intense follow-up of the major risk patients. This study is a potential first step in the development of personalized medicine tools for vascular recurrence after a first ischemic stroke

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Di Carlo A, Launer LJ, Breteler MMB, Fratiglioni L, Lobo A, Martinez-Lage J, et al. Frequency of stroke in Europe: a collaborative study of population-based cohorts. ILSA Working Group and the

Neurologic Diseases in the Elderly Research Group. Italian Longitudinal Study on Aging. Neurology. 2000; 54(Suppl 5):S28–S33. [PubMed: 10854359]

2. Launer LJ, Hofman A. Frequency and impact of neurologic diseases in the elderly of Europe: a collaborative study of population-based cohorts. Neurology. 2000; 54(Suppl 5):S1–S8.

3. Rosamond W, Flegal K, Furie K, Go A, Greenlund K, Haase N, et al. Heart disease and stroke statistics—2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation. 2008; 117:e25–146. [PubMed: 18086926]

4. Di Carlo A. Human and economic burden of stroke. Age Ageing. 2009; 38:4–5. [PubMed: 19141505]

5. Burn J, Dennis M, Bamford J, Sandercock P, Wade D, Warlow C. Long-term risk of recurrent stroke after a first-ever stroke. Stroke. 1994; 25:333–337. [PubMed: 8303740]

6. Hankey GJ, Jamrozik K, Broadhurst RJ, Forbes S, Burvill PW, Anderson CS, et al. Long-term risk of first recurrent stroke in the Perth Community Stroke Study. Stroke. 1998; 29:2491–2500. [PubMed: 9836757]

7. Weimar C, Goertler M, Röther J, Ringelstein EB, Darius H, Nabavi DG, et al. SCALA Study Group. Predictive value of the Essen Stroke Risk Score and Ankle Brachial Index in acute ischaemic stroke patients from 85 German stroke units. J Neurol Neurosurg Psychiatry. 2008; 79:1339–1343. [PubMed: 18586863]

8. Kernan WN, Viscoli CM, Brass LM, Makuch RW, Sarrel PM, Roberts RS, et al. The stroke prognosis instrument II (SPI-II): A clinical prediction instrument for patients with transient ischemia and nondisabling ischemic stroke. Stroke. 2000; 31:456–462. [PubMed: 10657422]

9. Thompson DD, Murray GD, Dennis M, Sudlow CL, Whiteley WN. Formal and informal prediction of recurrent stroke and myocardial infarction after stroke: a systematic review and evaluation of clinical prediction models in a new cohort. BMC Med. 2014:12–58. [PubMed: 24460622]

10. Hsieh YC, Seshadri S, Chung WT, Hsieh FI, Hsu YH, Lin HJ, et al. Formosa Stroke Genetic Consortium (FSGC). Association between genetic variant on chromosome 12p13 and stroke survival and recurrence: a one year prospective study in Taiwan. J Biomed Sci. 2012:19–21. [PubMed: 22316086]

11. Zhang W, Chen Y, Liu P, Chen J, Song L, Tang Y, et al. Variants on chromosome 9p21. 3 correlated with ANRIL expression contribute to stroke risk and recurrence in a large prospective stroke population. Stroke. 2012; 43:14–21. [PubMed: 22034006]

12. Williams SR, Hsu FC, Keene KL, Chen WM, Nelson S, Southerland AM. METASTROKE The Genomics and Randomized Trials Network GARNET Collaborative Research Group. Shared genetic susceptibility of vascular-related biomarkers with ischemic and recurrent stroke. Neurology. 2016; 86:351–359. [PubMed: 26718567]

13. Achterberg S, Kappelle LJ, de Bakker PI, Traylor M, Algra A. SMART Study Group and the METASTROKE Consortium. No Additional Prognostic Value of Genetic Information in the Prediction of Vascular Events after Cerebral Ischemia of Arterial Origin: The PROMISe Study. PLoS One. 2015; 10:e0119203. [PubMed: 25906364]

14. del Río-Espínola A, Fernández-Cadenas I, Giralt D, Quiroga A, Gutiérrez-Agulló M, Quintana M, et al. GRECOS Investigators. A predictive clinical-genetic model of tissue plasminogen activator response in acute ischemic stroke. Ann Neurol. 2012; 72:716–729. [PubMed: 23280790]

15. Spence JD, Howard VJ, Chambless LE, Malinow MR, Pettigrew LC, Stampfer M, et al. Vitamin Intervention for Stroke Prevention (VISP) trial: rationale and design. Neuroepidemiology. 2001; 20:16–25. [PubMed: 11174041]

16. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care. 1986; 24:67–74. [PubMed: 3945130]

17. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology. 1983; 148:839–843. [PubMed: 6878708]

18. Lemmens R, Smet S, Thijs VN. Clinical scores for predicting recurrence after transient ischemic attack or stroke: how good are they? Stroke. 2013; 44:1198–1203. [PubMed: 23482596]

19. Herrmann SM, Whatling C, Brand E, Nicaud V, Gariepy J, Simon A, et al. Polymorphisms of the human matrix gla protein (MGP) gene, vascular calcification, and myocardial infarction. Arterioscler Thromb Vasc Biol. 2000; 20:2386–2393. [PubMed: 11073842]

20. Brancaccio D, Biondi ML, Gallieni M, Turri O, Galassi A, Cecchini F, et al. Matrix GLA protein gene polymorphisms: clinical correlates and cardiovascular mortality in chronic kidney disease patients. Am J Nephrol. 2005; 25:548–552. [PubMed: 16210837]

21. Farzaneh-Far A, Davies JD, Braam LA, Spronk HM, Proudfoot D, Chan SW, et al. A polymorphism of the human matrix gamma-carboxyglutamic acid protein promoter alters binding of an activating protein-1 complex and is associated with altered transcription and serum levels. J Biol Chem. 2001; 276:32466–32473. [PubMed: 11425864]

22. Luo G, Ducy P, McKee MD, Pinero GJ, Loyer E, Behringer RR, et al. Spontaneous calcification of arteries and cartilage in mice lacking matrix GLA protein. Nature. 1997; 386:78–81. [PubMed: 9052783]

23. Mendioroz M, Fernández-Cadenas I, Rosell A, Delgado P, Domingues-Montanari S, Ribó M, et al. Osteopontin predicts long-term functional outcome among ischemic stroke patients. J Neurol. 2011; 258:486–493. [PubMed: 20967551]

24. Khoshniat S1, Bourgine A, Julien M, Petit M, Pilet P, Rouillon T, et al. Phosphate-dependent stimulation of MGP and OPN expression in osteoblasts via the ERK1/2 pathway is modulated by calcium. Bone. 2011; 48:894–902. [PubMed: 21147284]

25. Canfield AE, Farrington C, Dziobon MD, Boot-Handford RP, Heagerty AM, Kumar SN, et al. The involvement of matrix glycoproteins in vascular calcification and fibrosis: an immunohistochemical study. J Pathol. 2002; 196:228–234. [PubMed: 11793375]

26. Perneger TV. What's wrong with Bonferroni adjustments. BMJ. 1998; 316:1236–8. [PubMed: 9553006]

27. Lin YT, Lee WC. Importance of presenting the variability of the false discovery rate control. BMC Genet. 2015:16–97. [PubMed: 25887754]

28. Dichgans M, Markus HS. Genetic association studies in stroke: methodological issues and proposed standard criteria. Stroke. 2005; 36:2027–31. [PubMed: 16051898]
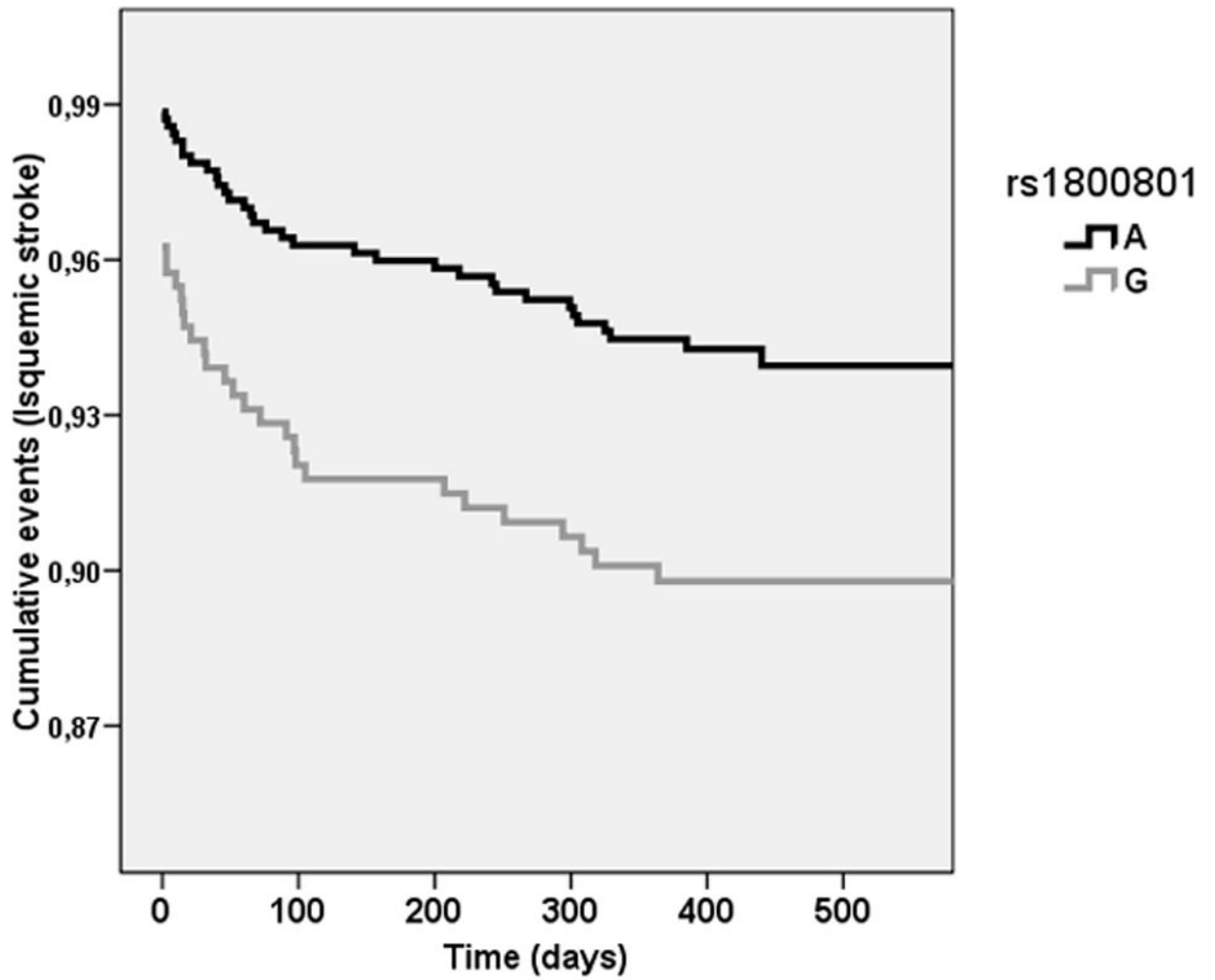
Figure 1. Kaplan-Meier curves of rs1800801 and time of recurrent ischemic stroke event
X axis: time (days) from first stroke or TIA to recurrent ischemic stroke, p value = 9×10^{-3}.
Y axis: cumulative events (ischemic strokes). Data obtained from Discovery cohort A, N = 1,494.
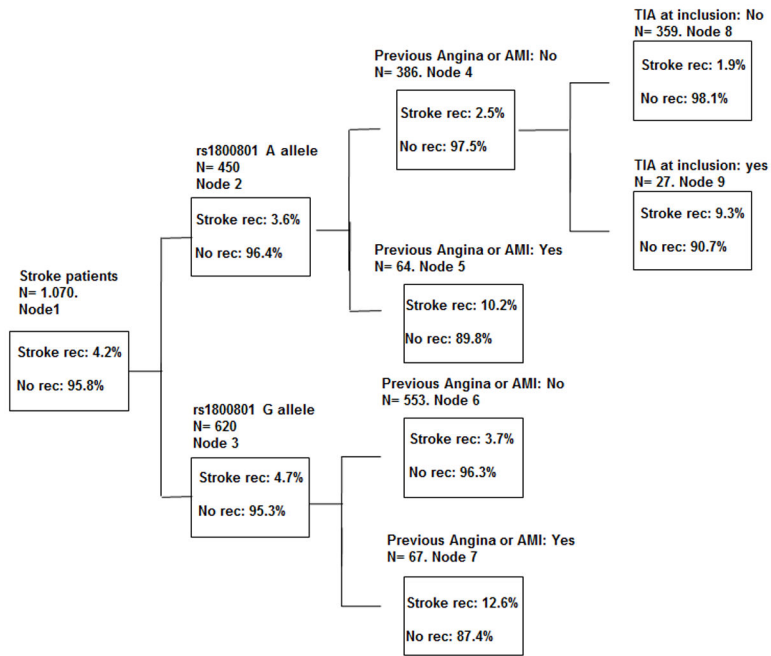
Figure 2. Classification tree

Depending on the clinical and genetic risk factors, the patient will be incorporated in a node. Each node has a different ischemic stroke recurrence risk. Node 7 corresponds to the maximum risk of suffering a recurrent ischemic stroke; 12.6% of patients in this node will suffer an ischemic stroke during follow-up. Node 8 corresponds to the minimum risk of suffering a recurrent ischemic stroke. Only 1.9% of patients in Node 8 will suffer a new ischemic stroke during follow-up. %: percentage of patients, AMI: acute myocardial infarction, Rec: stroke recurrence during the first year, N: number of patients. Data obtained from Discovery cohort A. Only patients with all the data available were included in the classification tree analysis (N= 1,070).

Table 1

Clinical variables associated with ischemic stroke recurrence during the first year. Discovery cohort (Cohort A)

The *p* values were obtained after Kaplan-Meier analysis, and *p* values of Cox regression were obtained after Cox regression. MI: Myocardial infarction, na: Not assessed, ns: Non-significant.

Clinical Variable	*p* value (Kaplan-Meier)	Hazard Ratio	*p* value (Cox regression)
Median Age (71 years)	0.001	1.85	**0.006**
Dyslipidemia	0.046	na	ns
Diabetes	0.005	na	ns
Angina or MI	0.003	3.65	**$9.5\ 10^{-09}$**
Diabetic nephropathy	**1.9×10^{-05}**	na	ns
TIA	0.006	2	**0.028**
TOAST	0.013	na	ns
Clopidogrel treatment	0.048	na	ns
Diuretics	0.020	na	ns
B-blockers	0.037	na	ns
Oral antidiabetic drugs	0.017	na	ns