Study protocol

The Ischemic Stroke Genetics Study (ISGS) Protocol

James F Meschia*[1], Thomas G Brott[1], Robert D Brown Jr[1], Richard JP Crook[2], Michael Frankel[3], John Hardy[4], José G Merino[5], Stephen S Rich[6], Scott Silliman[5] and Bradford Burke Worrall[7]

Address: [1]Department of Neurology, Mayo Clinic, Jacksonville, Florida, U.S.A, [2]Department of Neuroscience, Mayo Clinic, Jacksonville, Florida, U.S.A, [3]Emory University School of Medicine, Atlanta, Georgia, U.S.A, [4]National Institute on Aging, Bethesda, Maryland, U.S.A, [5]University of Florida/Shands Hospital, Jacksonville, Florida, U.S.A, [6]Department of Public Health Sciences and Neurology, Wake Forest University School of Medicine, Winston-Salem, North Carolina, U.S.A and [7]University of Virginia Health System, Charlottesville, Virginia, U.S.A

Email: James F Meschia* - Meschia.James@mayo.edu; Thomas G Brott - Brott.Thomas@mayo.edu; Robert D Brown - Brown@mayo.edu; Richard JP Crook - Crook.Richard@mayo.edu; Michael Frankel - mfranke@emory.edu; John Hardy - hardyj@mail.nih.gov; José G Merino - jose.merino@jax.ufl.edu; Stephen S Rich - srich@wfubmc.edu; Scott Silliman - scott.silliman@jax.ufl.edu; Bradford Burke Worrall - bbw9r@virginia.edu

* Corresponding author

Abstract

Background: The molecular basis for the genetic risk of ischemic stroke is likely to be multigenic and influenced by environmental factors. Several small case-control studies have suggested associations between ischemic stroke and polymorphisms of genes that code for coagulation cascade proteins and platelet receptors. Our aim is to investigate potential associations between hemostatic gene polymorphisms and ischemic stroke, with particular emphasis on detailed characterization of the phenotype.

Methods/Design: The Ischemic Stroke Genetic Study is a prospective, multicenter genetic association study in adults with recent first-ever ischemic stroke confirmed with computed tomography or magnetic resonance imaging. Patients are evaluated at academic medical centers in the United States and compared with sex- and age-matched controls. Stroke subtypes are determined by central blinded adjudication using standardized, validated mechanistic and syndromic classification systems. The panel of genes to be tested for polymorphisms includes β-fibrinogen and platelet glycoprotein Ia, Iba, and IIb/IIIa. Immortalized cell lines are created to allow for time- and cost-efficient testing of additional candidate genes in the future.

Discussion: The study is designed to minimize survival bias and to allow for exploring associations between specific polymorphisms and individual subtypes of ischemic stroke. The data set will also permit the study of genetic determinants of stroke outcome. Having cell lines will permit testing of future candidate risk factor genes.

Background

Cross-sectional, longitudinal, and twin studies strongly support an inherited component to stroke risk, but except for rare mendelian and mitochondrial stroke syndromes, the molecular basis for inherited ischemic stroke risk remains ill defined. The ability to identify high-risk patients through genetic testing could make screening for treatable intermediate phenotypes more cost-effective. For

example, identification of patients with a high genetic risk of cervical carotid atherosclerosis might enable the efficient use of endarterectomy for primary stroke prevention [1]. In addition, a clear and comprehensive understanding of genetic risk may promote advances in gene therapy and in the development of novel pharmaceutical agents.

Herein we describe the protocol of an ongoing prospective, multicenter study, the Ischemic Stroke Genetics Study (ISGS). This study uses a candidate gene approach in which rates of variant polymorphisms of the candidate genes are compared between patients with ischemic stroke and stroke-free control subjects.

Selection of Candidate Genes for ISGS
The candidate genes for this study include the gene encoding β-fibrinogen and the genes encoding platelet glycoprotein (GP) receptors Ia/IIa, Ib/IX/V, and IIb/IIIa. In selecting candidate genes for an association study, effects of polymorphisms on structure, function, or expression of a gene product should be considered. Failure to consider the underlying pathophysiologic mechanism when searching for polymorphisms associated with stroke might result in mistaking association for causation [2]. We studied genes related to thrombosis because the importance of thrombosis in acute ischemic stroke has been established conclusively in numerous clinical trials of treatment and prevention [3–8]. We focused on genetic variations in the fibrinogen gene cluster because of the efficacy of fibrinolytic agents in the acute treatment of ischemic stroke. In addition, we studied genes encoding for platelet receptors because of the efficacy of platelet anti-aggregant therapy in preventing first-time and recurrent ischemic strokes.

To restrict the choice of polymorphisms worthy of further study, we constructed an evidence table from reports appearing in MEDLINE-indexed English language journals describing cross-sectional or longitudinal studies of at least one thrombosis gene polymorphism in at least 100 patients with stroke. Results were classified as positive or negative according to whether a significant association ($P < 0.05$) was found between stroke (or carotid atherosclerosis) and a polymorphism. Because it is biologically plausible that a prothrombotic polymorphism may exert a differential effect across different ages, sexes, and ethnic groups, we classified studies as having positive results even if they had only one positive subgroup. We considered a polymorphism worthy of further study if it was not already a clearly established stroke risk factor and if at least one association study was positive.

Regarding a possible relationship to stroke risk, most studies of hemostasis genes have been inconclusive at best and unconvincing at worst. On the basis of the evidence, we concluded that the polymorphisms of factor VII R353Q, factor XIII Val34Leu, plasminogen activator inhibitor-1 4G/5G, and prothrombin G20210A were not worthy of further investigation because large studies had consistently yielded negative results (Table 1). For similar reasons, we decided not to study factor V R506Q (G1691A; i.e., the factor V Leiden mutation), despite its apparent association with cerebral vein thrombosis [9]. Although unknown point mutations in the coding regions of these genes may relate to stroke and relevant variations in gene expression elements may exist, we decided to focus on more immediately high-yield candidate genes.

The results of three large European studies listed in Table 1 led us to conclude that the β-fibrinogen gene might be a promising candidate. Fibrinogen is a 340,000-Da GP consisting of three polypeptide chains: α, β, and γ. The genes that encode these polypeptides reside on chromosome 4q in a cluster. In a study of the β-fibrinogen G455A polymorphism, Kessler et al. [10] did not find an overall association between genotype and stroke, but heterozygosity for the A allele was associated with large-vessel ischemic stroke ($P = 0.045$). Schmidt et al. [11] observed an association between carotid atherosclerosis and the C148T polymorphism in a population-based cross-sectional study of persons with normal neurologic status. Carotid atherosclerosis was seen in 53.6% of persons with the C/C genotype, 54.1% of those with the C/T genotype, and 88% of those with the T/T genotype ($P = 0.003$). Abnormal results on carotid ultrasonography were significantly more common in the T/T genotype group (OR, 6.29; 95% CI, 1.91 to 20.71). Data from the study by Carter et al. [12] on the G448A polymorphism of the β-fibrinogen gene suggested that mechanisms linking fibrinogen and the development of cerebrovascular disease may be different in men and women.

Several studies listed in Table 1 suggested that polymorphisms of genes controlling the three platelet glycoprotein receptors Ia/IIa, Ib/IX/V, and IIb/IIIa, which play a role in adhesion, might also be promising candidate risk factors for stroke. GP Ia/IIa (integrin $\alpha_2\beta_1$) is involved in collagen-induced platelet aggregation. It does not bind collagen monomers, but it does bind collagen fibrils and immobilized collagen. Binding of GPIa/IIa to collagen induces a conformational change in receptor structure that enhances affinity. Thus, one platelet GP of interest is GPIa. Carlsson et al. [13] compared the GPIa (α_2) C807T genotype distribution in patients with ischemic stroke or transient ischemic attacks with that in hospitalized patients without cerebrovascular disease and in healthy blood donors. An association between the polymorphism and stroke was not seen overall. However, there was an overrepresentation of the C807T polymorphism in patients with stroke age 50 years or younger ($n = 45$)

Table 1: Association Studies of Thrombosis Genes in Ischemic and Carotid Atherosclerosis

Gene	Reference	Polymorphism	Cases, *n*; controls, *n*	Results
Factor V	Catto et al., 1995 [56]; Catto et al., 1996 [57]	G1691A	348; 247	Negative
	Kontula et al., 1995 [58]	G1691A	236; 137	Negative
	Ridker et al., 1995 [15]	G1691A	209; 209	Negative
	Martinelli et al., 1997 [59]	G1691A	155; 155	Negative
	Press et al., 1996 [60]	G1691A	116; 54/161/287	Negative
	van der Bom et al., 1996 [61]	G1691A	112; 222	Negative
	Longstreth et al., 1998 [62]	G1691A	106; 391	Negative
	Halbmayer et al., 1997 [63]	G1691A	229; 71	Negative
	Markus et al., 1996 [64]	G1691A	180; 80	Negative
Factor VII	Nishiuma et al., 1997 [65]	R353Q	137; 83/87	Negative
	Heywood et al., 1997 [66]	R353Q	286; 198	Negative
	Corral et al., 1998 [67]	R353Q/323A2	104; 104	Negative
Factor XIII	Catto et al., 1998 [68]	Val34Leu	529; 437	Negative
β-Fibrinogen	Schmidt et al., 1998 [11]	C148T	399	**Positive**
	Carter et al., 1997 [12]	G448A	305; 197	**Positive**
	Kessler et al., 1997 [10]	G455A	227; 225	**Positive**
GPIa/IIa	Carlsson et al., 1999 [13]	C807T	227; 170	**Positive**
	Carlsson et al., 1997 [69]	HPA-5	218; 165/321	Negative
GPIbα	Gonzalez-Conejero et al., 1998 [14]	VNTR	104; 104	**Positive**
		HPA-2	104; 104	**Positive**
GPIIb/IIIa	Carter et al., 1998 [16]	PlA2	505; 402	**Positive**
	Ridker et al., 1997 [70]	PlA2	209; 209	Negative
	Carlsson et al., 1997 [69]	HPA-1	218; 165/321	Negative
		HPA-3	218; 165/321	Negative
PAI 1	Catto et al., 1997 [71]	4G/5G	421; 172	Negative
Prothrombin	Longstreth et al., 1998 [62]	G20210A	106; 391	Negative
	Poort et al., 1996 [72]	G20210A	104; 104	Negative
	Martinelli et al., 1997 [59]	G20210A	155; 155	Negative
	Ridker et al., 1999 [73]	G20210A	259; 1,744	Negative

GP, glycoprotein; PAI, plasminogen activator inhibitor.

versus age-matched controls (OR, 3.02; 95% CI, 1.20 to 7.61). No such overrepresentation was detected in older patients.

The second platelet GP of interest is GPIbα, a transmembranous platelet GP (molecular weight, 143,000) that forms noncovalent complexes with GPIbβ, GPIX, and GPV to form the GPIb/IX/V receptor, which is involved in shear stress-induced platelet activation by binding to von Willebrand factor (vWF). This receptor may be particularly relevant in large-vessel atherosclerotic ischemic stroke because high shear stresses like those seen in atherosclerotic arteries increase ligand-receptor affinity. The receptor may also have a role in so-called aspirin failure, in which patients suffer stroke despite taking daily aspirin prophylaxis. Cyclooxygenase inhibition by aspirin has little effect on initial aggregation in response to shear forces. One GPIbα polymorphism is referred to as "VNTR" because it consists of a variable number of tandem repeats of 39 base pairs, each repeat leading to a 13-amino acid addition that pushes a vWF-binding domain further away from the platelet membrane surface. Another is human

platelet antigen-2 (HPA-2), a mutation that codes for either a thr (HPA-2a) or met (HPA-2b) at position 145. The HPA-2 site resides next to the vWF and high-affinity thrombin binding sites.

In a case-control study of these polymorphisms, Gonzalez-Conejero et al. [14] found that cerebrovascular disease was associated with both the C/B genotype of the VNTR polymorphism (OR, 2.83; 95% CI, 1.16 to 7.07; *P* = 0.0114) and the β allele of the HPA-2 polymorphism. Of the 104 patients with cerebrovascular disease, 22.11% carried at least one β allele compared with 10.58% of controls (OR, 2.40; 95% CI, 1.04 to 5.63; *P* = 0.0244). Neither polymorphism showed significant differences related to age, sex, or type of cerebrovascular disease. Both polymorphisms also correlated with coronary artery disease, but neither correlated with deep vein thrombosis. This is the converse of what Ridker et al. [15] and others found for factor V Leiden. Taken together, the studies suggest that polymorphisms predisposing to arterial thrombosis may differ from polymorphisms predisposing to deep vein thrombosis. This hypothesis supports the

rationale for a hemostasis candidate-gene association study such as ISGS, which investigates ischemic stroke specifically and does not regard all acute thrombotic events, whether arterial or venous, as a single clinical entity.

The third candidate platelet GP gene controls GPIIb/IIIa (integrin $\alpha_{IIb}\beta_3$), a transmembranous heterodimer with several ligands, including fibrinogen, fibrin, fibronectin, and vWF. Many receptors are involved in platelet adhesion and many agonists stimulate platelet aggregation, but platelet aggregation requires GPIIb/IIIa. When platelets aggregate, GPIIb/IIIa binds to fibrinogen and vWF. Binding to vWF gains importance under conditions of high shear stress. Carter et al. [16] found no overall association between the P1A2 polymorphism of the GPIIb/IIIa gene and cerebral infarction confirmed by computed tomography (CT). However, a subgroup analysis showed significant genotype distribution differences in nonsmokers. The risk of stroke was greater in nonsmokers heterozygous for the P1A2 allele than in those homozygous for P1A2 (OR, 2.37; 95% CI, 1.19 to 4.74; $P = 0.01$). Information on young stroke patients was limited ($n = 37$), but in a logistic regression model that included P1A genotype status, smoking, hypertension, and diabetes, the OR for stroke in those possessing the A2 allele was 1.68 (95% CI, 1.00 to 2.82; $P = 0.05$). This study highlights the need for further studies of the interaction between genes and environmental factors, in this case smoking, in attempts to elucidate inherited stroke risk.

Aims

The primary aim of ISGS is to test the association between ischemic stroke and the following putative risk factor polymorphisms: β-fibrinogen C148T, G448A, and G455A; GPIa C807T; GPIbα HPA-2 and VNTR; and GPIIb/IIIa P1A. Exploratory aims are to investigate whether any association found between ischemic stroke and the panel of tested polymorphisms is contingent on sex, age, ethnic origin, smoking status, or stroke subtype and to investigate whether hemostatic gene sequence variations are associated with 90-day functional outcome after acute ischemic stroke.

Methods

Design and Overview

The ISGS is a prospective multicenter study using a case-control design (Fig. 1). Patients and control subjects are screened at one of five clinical centers (Appendix 1 – Additional file: 1), stroke status is verified, the index stroke for each patient is subtyped, and baseline clinical and demographic data are collected. Blood samples are collected from all enrolled patients and control subjects by means of a one-time venipuncture. The samples are shipped to a central DNA bank for processing and storage, and the
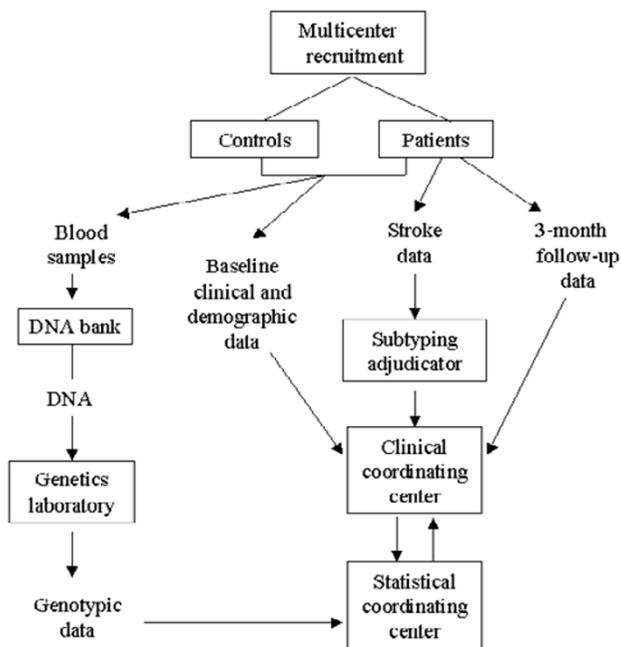

Figure 1
Overview of flow of samples and data from study subjects to genetic analyses.

processed DNA samples are sent to a central genetics laboratory for genotyping. The genotype data are then merged with the clinical, stroke, and follow-up data and analyzed to ascertain potential associations between stroke risk and genes for β-fibrinogen and platelet GPIa, GPIbα, or GPIIb/III.

Study Population

Patients With Stroke (Cases)

Each patient with suspected stroke admitted to a participating center is evaluated by a study neurologist according to current standards for care [17,18]. The evaluation includes patient history, physical examination, CT or magnetic resonance imaging (MR) of the head, and laboratory testing. Where clinically indicated, the evaluation may also include carotid ultrasonography; MR, CT, or digital subtraction angiography; transthoracic or transesophageal echocardiography; resting and ambulatory electrocardiography; intracranial arterial imaging; and additional blood testing.

Adult men and women who meet the following criteria are entered into the study: 1) diagnosis of first-ever ischemic stroke confirmed by the study neurologist on the basis of history, physical examination, and head imaging by CT or MR; 2) enrollment within 30 days after onset of

stroke symptoms; 3) attained 18th birthday by the time of enrollment; 4) complete blood cell count, casual or fasting blood glucose, prothrombin time, and activated partial thromboplastin time available; and 5) written informed consent from the patient or surrogate.

Stroke is defined according to the World Health Organization criteria [19] as rapidly developing signs of a focal or global disturbance of cerebral function with symptoms lasting 24 hours or longer or leading to death, with no apparent cause other than vascular origin. A diagnosis of ischemic stroke is made only if the patient has a clinical diagnosis of stroke and if a CT scan or MR of the brain done after onset of symptoms either is normal or shows the relevant infarct. Patients with hemorrhagic transformation of an infarct remain eligible.

Time of stroke onset is defined as the time when the subject was last noted to be at baseline neurologic status. If the patient awoke with stroke symptoms, the time of onset is taken as the last time the patient was known to be awake and without any symptoms of stroke. We restrict enrollment to within 30 days after onset of symptoms from the first-ever stroke to avoid potential survival bias [20–23]. The date of enrollment is the date of obtaining signed informed consent.

Exclusion criteria parallel those of the Siblings With Ischemic Stroke Study (SWISS) [24]. Patients who are already enrolled in SWISS are not eligible for participation in ISGS.

To be able to assess the extent to which enrolled patients represent all potential subjects, clinical study coordinators at each site keep logs of every eligible stroke patient who is offered participation in the study, whether or not they are enrolled. The logs will contain initials and date of birth of the eligible stroke patients, date of screening, sex, and race/ethnicity.

Controls
Controls are adult men and women who have attained their 18th birthday at the time of enrollment, have not had a stroke, are unrelated by blood to patients enrolled in the study, and who give written informed consent to participate in the study. We confirm that controls have not had a prior stroke by means of the Questionnaire for Verifying Stroke-free Status (QVSS), a structured interview that was validated in an adult population (age, > 60 years) using systematic review of electronic medical records as the benchmark [25]. The QVSS was further validated in an independent population using history and physical examination by a study neurologist as the benchmark [26]. Interviewers administering the QVSS may exclude a subject they judge to be an unreliable historian on the basis

of a global impression of moderate or severe impairment of speech, language, hearing, or memory. Hospitalized patients being treated for coronary or peripheral vascular disease are not eligible for enrollment as controls, but nonhospitalized subjects with a history of these conditions are eligible.

Cases and controls are matched one-to-one. Matching criteria are sex and age (within 3 years for patients who are younger than 30 years and within 5 years for patients who are 30 years or older). We recruit controls mainly from among spouses and unrelated friends of the patients. Each center has a backup plan for recruiting community volunteers should there be a lag in recruitment of properly matched controls [27].

Data Collection
The schedule for data collection is shown in Table 2.

Structured Interview
A structured interview is conducted by the study coordinator with each patient (case) or surrogate and each control subject to explain the study, obtain informed consent, and obtain standardized information on baseline medication and demographic, medical, social, and behavioral variables. Information regarding race and ethnicity is recorded according to self-report. A proband-derived family history is taken for all living or deceased full siblings, all biological children, and both biological parents [28]. Investigators do not independently verify stroke status of family members as part of this protocol. Self-reported cerebrovascular histories are obtained for all patients and control subjects by administering the QVSS during the baseline interview [25].

Medical Records
Study coordinators review the medical records of the initial evaluation of stroke cases to complete case report forms for documenting eligibility and baseline data and to construct the abstracted medical record used for stroke subtyping. The following information is recorded on the case report forms: patient history, physical examination, CT or MR of the head, white blood cell count, platelet count, and hemoglobin concentration, casual or fasting blood glucose, prothrombin time, and activated partial thromboplastin time, vital signs (height, weight, blood pressure, and temperature), international normalized ratio, lipid profile, plasma homocysteine concentration, and size and location of the symptomatic cerebral infarct as seen on head imaging.

The clinical coordinator also constructs the abstracted medical record used for subtyping of the index stroke by copying admission notes, physician progress notes, discharge summaries, radiology reports, electrocardiograms,

Table 2: Events Table for the Ischemic Stroke Genetics Study

| Variables | Enrollment | | 90 ± 14 days after stroke |
	Cases	Controls	Cases
Structured interview			
Demographics	X	X	
Medications	X	X	
Family history	X	X	
Past medical history	X	X	
Social and behavioral history	X	X	
QVSS	X	X	
Blood sample	X	X	
Medical records			
Vital signs	X		
Laboratory tests	X		
Head imaging	X		
Vascular imaging	X		
Treatments	X		
Stroke status			
Subtyping	X		
Time of stroke onset	X		
NIH Stroke Scale	X		
Barthel Index	X		T
Oxford Handicap Scale	X		T
Glasgow Outcome Scale	X		T

NIH, National Institutes of Health; QVSS, Questionnaire for Verifying Stroke-free Status; T, assessment by telephone; X, assessment by face-to-face interview or examination.

echocardiography reports, laboratory reports, and rehabilitation notes.

Investigators and coordinators use standardized definitions for major medical and surgical comorbid conditions (Appendix 2 – see Additional file: 2). Because blood pressure typically falls during the first few days after an acute ischemic stroke, we have adopted a modification of the Northern Manhattan Stroke Study (NOMASS) technique [29], in which we use the systolic and diastolic blood pressure values measured after admission to the hospital floor (or intensive care unit) rather than measurements taken by the emergency medical service or in the emergency department. Blood pressure is measured from the left brachial artery (if attainable) with a sphygmomanometer, with the patient sitting upright. If the patient is bedbound, the measurement is made with the head of the bed elevated to at least a 45° angle.

Characterization of Ischemic Stroke
An on-site, study-appointed neurologist confirms the diagnosis of ischemic stroke and the time of stroke onset by interviewing patients or any available observers present when the stroke was first noticed [30]. The examiner seeks corroborating evidence (such as ambulance reports) and carefully screens for the possibility of onset during sleep.

The severity of the neurologic deficit is assessed within 48 hours of the patient's enrollment in the study by means of the National Institutes of Health Stroke Scale (NIHSS) [31] administered by a certified examiner. The first CT or MR obtained is used to measure infarct size by means of standardized criteria (Appendix 3 – see Additional file: 3).

Prestroke functional status is assessed retrospectively by the study coordinator with the Oxford Handicap Scale [32]. Acute poststroke functional status is assessed with the Oxford Handicap Scale [32], Barthel Index [33,34], and Glasgow Outcome Scale [35] within 48 hours of enrollment.

Subtyping of the index ischemic stroke is done centrally on the basis of the abstracted medical record by a neurologist adjudicator blinded to genotype data and personal indentifiers. Because final subtype diagnosis has been shown to vary from initial diagnosis in approximately one-third of cases [36], the adjudicator uses all available and relevant information obtained after completion of the stroke work-up. The Trial of ORG10172 in Acute Stroke Treatment (TOAST) [37], Oxfordshire Community Stroke Project (OCSP) [38], and Baltimore-Washington Young Stroke Study (BWYSS) [39] classification systems are used for subtyping.

Follow-up of Stroke Patients
Patients with stroke are followed up by the study coordinator at the center from which they were recruited. The coordinator reassesses patients by using the Oxford Handicap Scale, Barthel Index, and Glasgow Outcome Scale by telephone interview 90 ± 14 days after onset of stroke symptoms. Coordinators preferentially interview the patients themselves. However, if acquired deficits of speech, language, or cognition prevent the patient from participating in the telephone outcomes assessment, a surrogate history is taken from a caregiver or live-in relative. Coordinators will also record mortality and history of cause of death from collateral sources.

Genetic Data
The study coordinator at the local center obtains two tubes of peripheral blood from each patient and each control subject, collected in a 10-mL (8.5-mL draw) acid-citrate-dextrose solution A (ACD) tube. The blood samples are shipped to a central DNA bank by overnight courier and assigned a unique repository identifier. Lymphocytes are isolated, and 0.5 mL of blood is retained in the original tube as a quality control specimen for identity testing. Isolated lymphocytes are cryopreserved using controlled-rate freezing and stored at the liquid phase of nitrogen.

The DNA bank prepares high-quality, high-molecular-weight DNA from the cell pellet using a modification of the salting-out procedure of Miller et al. [40]. Quality control studies on DNA consist of estimation of quantity by OD260/OD280 ratio, estimation of integrity by gel electrophoresis and restriction digestion, and verification of identity by microsatellite analysis and sex determination.

Transformation of lymphocytes is done using Epstein-Barr virus and phytohemagglutinin. Lymphocyte cultures are expanded to produce sufficient stock for 10 to 12 ampules. All cell cultures are done in the absence of antibiotics. Cryopreservation is again done by controlled-rate freezing. Samples are stored at the liquid phase of nitrogen. In-house and remote fail-safe stocks are generated. The following routine quality control studies for cell culture are performed: recovery of frozen stock and determination of viability, sterility testing for bacterial and fungal contamination, testing for mycoplasma contamination by polymerase chain reaction (PCR), and confirmation of the identity of the culture by comparing the DNA fingerprint of the culture with that of the quality control specimen. If the first attempt fails, a second aliquot of cryopreserved lymphocytes can be transformed.

Genotyping is done at a central genetics laboratory. Currently, the following sequencing methods are used, but the laboratory will be responsive to technological advances in the field.

PCR is carried out in a 75-μL volume on 40-ng genomic DNA by use of plain primers under standard conditions with a 57° to 52°C (0.5°C/cycle) "touch-down" annealing temperature. To remove excess unincorporated primers that would compete as sequencing primers in the cycle sequencing reaction, the amplified product is filtered with MultiScreen PCR filters (Millipore) and resuspended in 50 μL. The sequencing reaction is carried out using the BigDye Terminator cycle sequencing kit (Applied Biosystems) as per the manufacturer's conditions. To remove excess dye terminators, the sequencing product is purified by ethanol precipitation and resuspended in 10 μL of HiDi formamide. The samples are then denatured and electrophoresed on an ABI 3100 capillary analyzer. Data analysis is carried out with a software suite (ABI) consisting of Sequencing Analysis (base calling), Factura (heterozygous base detection), and Sequence Navigator (sequence comparison).

Adverse Events
All adverse events and serious adverse events will be recorded by the study coordinators and forwarded to the statistical center. Potential physical risks are minimal and relate to the one-time phlebotomy. An independent medical safety monitor will review summary reports of adverse events and serious adverse events periodically and will forward assessments to the study Principal Investigator.

Outcome Measures
The main end point of the study is whether any of the polymorphisms (the β-fibrinogen polymorphisms C148T, G448A, and G455A and the platelet GP polymorphisms GPIa C807T, GPIbα HPA-2 and VNTR, and GPIIb/IIIa P1A) are associated with ischemic stroke. Thus, patients with stroke will be compared with controls as to frequency and distribution of these polymorphisms. Other end points include potential associations among these polymorphisms and individual subtypes of ischemic stroke, ethnic origin, or 90-day poststroke functional outcome and mortality. Additional analyses will include testing for interactions between inherited susceptibility (genotype) and environmental exposures (e.g., smoking).

Data Analyses
Several types of analyses will be performed to assess the relationships between outcome (binary and continuous) and risk factors (both genetic and environmental). The statistical techniques we use to analyze the data depend on the distribution of the independent (predictor) and dependent (outcome) variables. When the outcome variables are categorical (i.e., stroke [yes/no]) we will use chi-square tests and logistic regression techniques; when the outcome variables are continuous (with or without inclusion of collected longitudinal data), we will use

repeated measures analysis of covariance (ANCOVA) techniques.

Cases will be compared with controls with respect to predictor variables of interest and with respect to genotypes that may provide increased risk for stroke. Where needed, adjustments for potential covariates (age, sex, hypertension, etc.) will also be included in these analyses. To perform the above comparisons, we will use logistic regression, a statistical technique for modeling the relationship of a binary outcome and a set of independent (or predictor) variables [41].

For our case-control comparisons, a second series of analyses relates to the distribution of polymorphisms within the selected candidate genes in the cases (stroke positive) and controls (stroke negative). Each subject will be characterized by the polymorphism (or mutation) found at each candidate gene. Different analytical strategies need to be employed for the different candidate loci under study. For the fibrinogen gene cluster, it is proposed that the genotype be considered as a binary variable, that is, as a single nucleotide polymorphism (SNP) with two possible alleles, since the outcome of polymorphisms in genes of this cluster has resulted in apparent increased risk for ischemic heart disease through increased circulating fibrinogen levels. Thus, initial analyses using members of the fibrinogen cluster (or with other candidate genes such as GPIbα) will focus on each case/control being classified by the presence of a risk allele (yes/no), with comparison of SNP allele frequencies between groups (stroke/control).

For other candidate genes (GPIbα, etc.), distributions of genotypes or strata determined by alleles or haplotypes can be established for comparison. For example, GPIbα has a series of polymorphisms that may define an allele of interest. Extending SNPs to haplotypes has the advantage of increased information, yet decreased power in small samples due to the increased number of possible haplotypes to be tested. Restriction of haplotypes by "binning" may provide some increase in power, yet it assumes previous knowledge of risk haplotypes. These few haplotypes can be used to establish strata in the analyses proposed. In separate analyses, the SNP/polymorphism status can be considered to be an exposure variable. Thus, risk factor distributions can be compared between "exposed" subjects (with a polymorphism in a candidate gene) and those "not exposed" (without a polymorphism). For levels of continuous risk factors, this analysis will compare means by analysis of variance methods. For dichotomous risk factors, contingency table methods can be used.

In the case-control analyses, each candidate locus can be analyzed individually as a potential modifier of disease risk (stroke) and as a determinant of other intermediate (continuous trait) end points. To evaluate the relationships between gene interaction (gene-gene) and interaction between inherited susceptibility (genotype) and environmental exposures (e.g., hypertension or smoking), several strategies can be employed. One approach is stratification by genotype at each candidate locus, with analysis of "case" status with the second genetic (or environmental) exposure. For dichotomous exposures, this reduces to a comparison of contingency tables within genotype strata.

A second approach uses logistic regression. Logistic regression can be used for continuous risk factors within genotype strata. Similarly, multivariate logistic regression can be used to predict group membership (stroke vs. control) based on age, sex, environmental risk factors that appear significant in univariate analyses, and genotype(s), with first-order interaction terms of genotype with environment and gene 1 with gene 2.

Sample Size and Study Power

We intend to enroll a total of 900 participants, including 450 patients with ischemic stroke (cases) and 450 stroke-free volunteers (controls), at five academic medical centers in the United States over 3 years. Estimates of potential for recruitment were based on the assumption that the rates of hospitalized stroke cases at the five participating hospitals will remain at 1999 levels throughout the patient recruitment phase of the study.

We anticipate that the study population will be approximately two-thirds white and one-third African American, giving sample sizes of approximately 300 each for white patients with stroke and control subjects and 100 each for African American patients with stroke and control subjects. Thus, for the above analyses, we are concerned primarily with dichotomous outcomes with fixed sample sizes (100, 300, or 450 stroke cases and 100, 300, or 450 controls). Power for any analyses involving continuous measures can be approximated using an independent *t* test comparison between the incident cases and controls. In the case-control study, the power to address specific issues relating to the genetic associations is dependent on the sample size available, the frequency of the polymorphisms (in the fibrinogen cluster and other candidate genes), and the size of the effect to be detected.

The power is determined on the basis of the frequency of the SNP in the control group and the detectable difference in SNP allele frequency in the control group. For example, if an SNP has a frequency of 0.40 in the total group of controls (n = 450), we can detect (with 75% power) a 20% difference in cases [(0.20)(0.40) = 0.08] or frequencies greater than 0.48 or less than 0.32 in cases. This is obvi-

ously a small difference (or a relatively weak genetic relative risk). Using this same example, we would have 46% power in whites and 22% power in African Americans. For the African American group, we generally have power for relatively high-frequency polymorphisms in controls (approximately 45%-50%) and detectable differences of 40% in cases (0.20 difference). Thus, we could detect an SNP allele frequency of 0.70 in cases and 0.45 in controls with 80% power.

In general, the power to detect gene-environment interaction effects is less than that for main effects and is dependent on the type of interaction. For example, using methods discussed by Goodman [42], the power to detect interactions for multiple scenarios is possible. For these scenarios, 2 × 2 tables can be used to describe the overall main effects of genetic and environmental risks (i.e., both the genetic and the environmental risk factors display a main effect size [odds ratio] of about 4).

For large interactions we have > 99% statistical power. Only in the smallest subgroup would power be appreciably affected. For example, for the African American group (100 cases and 100 controls), with the same distribution of individuals as above, the power to detect the large interactions (OR of 0.33 in controls, OR of 5.33 in cases) would still be ~99%, while the power to detect the modest interaction (OR of 0.80 in controls, OR of 1.50 in cases) would be reduced to 13%.

Ethical Considerations

Local institutional review boards (IRBs) governing each clinical center have approved the study protocol. Written informed consent is obtained for every study subject. To the extent permitted by local IRBs, surrogate consent is permitted for patients rendered incompetent by stroke to avoid biasing the study toward discovery of risk factors for mild or moderate stroke but not severe stroke.

Genetic information has the potential to adversely affect insurability and employability [43,44]. An individual's genetic information could also lead to stigmatization within a family and a community [45]. Because of the highly sensitive nature of genetic information, we have developed the following plan to prevent intentional or unintentional misuse of genetic data, which is in keeping with the Privacy Workshop Planning Subcommittee Guidelines of the National Action Plan on Breast Cancer [46].

Every ISGS investigator who obtains or has access to genotypic data is blinded to individual personal identifiers, and every investigator who obtains or has access to individual personal identifiers is blinded to genotypic data. Personal identifiers are defined as individual names,

addresses, phone numbers, fax numbers, and e-mail addresses. Personal identifiers and linkage codes are kept only at the clinical center where the study subject was enrolled and will not be recorded on case report forms or stored in the study electronic database. Experimental research data are not placed in a participant's medical record. There have been various opinions regarding whether family members should be considered human subjects in pedigree research and when it is permissible to waive consent [28,47–49]. In ISGS, no personal identifiers are collected on family members.

The protocol of ISGS calls for centrally banking DNA and creating immortalized cell lines. Unlike the situation in clinical trial research, no broad consensus on research ethics exists among investigators and eligible participants in the field of human molecular genetics research. Of recent contention are procedures for ensuring ethical future use of stored genetic material [50,51]. The potential for future use of DNA should be anticipated, even if the specific studies cannot be known. A future-use agreement should respect genetic privacy rights and autonomy of human subjects while encouraging scientific inquiry. Establishing a transparent and ethically sound future-use agreement for research in this field facilitates multicenter collaborations and future research. After reviewing available policy statements from genetics societies and governmental agencies, we developed the following list of key principles, which are incorporated into a future-use agreement governing DNA banked for this study: 1) Subjects must provide informed consent to the original study at the time of DNA collection. 2) The original study must have a rigorous procedure in place to protect privacy of study subjects prior to DNA collection. 3) Patients should have the option to consent to, or refrain from, participation in future research at the time of the initial consent process. 4) Levels of consent must be clear, explicit, and exclusive (e.g., original study only, any stroke study, any study). 5) Applications for future use need to be reviewed formally. 6) Study investigators can release DNA for future use only after determining whether the new study conforms to the type of research permitted by the donor. 7) All specimens are stripped of direct personal identifiers before future use. 8) Anonymous data sets with genetic information from different studies may be merged for hypothesis-generating analyses. We believe that this carefully developed, publicly scrutinized future-use agreement is a unique strength of this study and hope that this proactive approach will avoid some of the rancorous misunderstandings that other investigative groups have encountered in genetic and pedigree research [51].

Discussion

Defining the molecular basis for the inherited component to ischemic stroke risk will require converging lines of evi-

dence from various methodologies, including genomewide screens [24] and candidate gene association studies. Collection of DNA samples from a large cohort of ischemic stroke pedigrees is logistically challenging. Collection of DNA samples from patients with stroke and unrelated stroke-free controls is more feasible.

Some stroke genetics research has been done in the context of epidemiological studies; for example, a study by Ridker and colleagues [15] tested for an association between factor V Leiden and stroke. One limitation of this approach is that there may be relatively few stroke end points in an epidemiological study, particularly because stroke, myocardial infarction, and vascular death are often combined as the primary end point. Furthermore, for various reasons, epidemiological studies may include a selected population, for example, only one sex or a limited socioeconomic stratum. The use of such highly selected samples may compromise the ability to generalize the findings of a genetic association study.

Genetic association can be studied within the context of a clinical trial of an intervention for primary prevention. In a prevention study, DNA would be available both from patients with stroke and from stroke-free controls, but the intervention may alter the outcome (stroke) sufficiently to confound interpretation of the results of a genetic association study.

A genetic study in the context of a randomized trial of treatment of stroke patients is a robust study design when the goal is to discover genetic determinants of stroke outcome or response to therapy (pharmacogenomics). However, such a study design may not achieve the goal of discovering genetic risk factors for stroke because no genetic material from stroke-free controls outside of the study would be available. Furthermore, highly restrictive criteria for eligibility into a clinical trial can drastically limit the representativeness of a study population. For example, in one study of intra-arterial thrombolysis for treatment of acute ischemic stroke, 12,323 patients with stroke were screened, and only 180 patients were selected [4]. Other stroke studies include only patients with non-disabling strokes [52], or include only patients with moderate to severe strokes [4]. Because ISGS is a dedicated stroke genetic association study, it has broad eligibility criteria relative to clinical trials of drugs or devices. We expect this to enhance the external validity of the results of the study.

Additionally, we have given careful attention to appropriate selection of controls [27]. The controls are concurrently enrolled at the same centers as are the patients, and controls are also screened for a medical history of stroke or transient ischemic attack and for the presence of symp-

toms of stroke or transient ischemic attack which may have occurred in the absence of a corresponding medical history.

A unique strength of ISGS is that the protocol was specifically developed to permit valid future pooled analyses with the ongoing affected sibling pair linkage study SWISS [24,53], which uses genome-wide screening within a cohort of pedigrees to identify chromosomal regions – rather than specific polymorphisms – that are linked to stroke. Conducting a candidate gene association study in cases and controls has advantages compared with the genome-wide linkage approach in siblings. Isolated cases are more likely to be available than sibling pairs, increasing the potential for statistical power. Furthermore, determination of phenotype is often retrospective in genome-wide linkage studies because it is rarely possible to enroll all affected members of a pedigree shortly after onset of stroke. In contrast, the case-control approach allows prospective phenotyping of all subjects at the time of onset of stroke, which may be more reliable than retrospective phenotyping. This is an advantage because ischemic stroke has a heterogeneous phenotype, and distinctions between specific clinical subtypes of ischemic stroke may be relevant [54,55]. To address the heterogeneity of the ischemic stroke phenotype, ischemic stroke is subtyped by a genotype-blinded adjudicator in both ISGS and SWISS.

ISGS and SWISS also use the same definitions for ischemic stroke and comorbidity and the same criteria for classifying stroke mechanism. Both studies use the same key exclusion criteria; for example, both studies exclude iatrogenic and vasospastic ischemic stroke patients and exclude the same mendelian and mitochondrial disorders. In addition, the stroke-free status of controls in ISGS and of discordant siblings in SWISS is verified using the same structured interview instrument (i.e., the QVSS [25]). We anticipate that the two studies will complement each other in contributing to the broad, long-term objective of defining the molecular basis for inherited ischemic stroke risk.

List of abbreviations
BWYSS, Baltimore-Washington Young Stroke Study

CT, Computed tomography

GP, Glycoprotein

ISGS, Ischemic Stroke Genetics Study

MR, Magnetic resonance imaging

NIHSS, National Institutes of Health Stroke Scale

OCSP, Oxfordshire Community Stroke Project

OR, Odds ratio

PCR, Polymerase chain reaction

QVSS, Questionnaire for Verifying Stroke Status

SWISS, Siblings With Ischemic Stroke Study

TOAST, Trial of ORG10172 in Acute Stroke Treatment

vWF, von Willebrand factor

Competing interests
None declared.

Authors' contributions
Dr. Meschia is the principal investigator in ISGS, developing and writing the protocol from its inception to the final version. Drs. Brott, Brown, Frankel, Merino, Silliman and Worrall are site investigators who contributed to developing and writing the final protocol. Dr. Hardy and Mr. Crook contributed to writing the portions of the manuscript that refer specifically to DNA sequencing and analysis. Dr. Rich contributed to writing the portion of the manuscript that refers to the statistical analytical plan.

Additional material

> ### Additional file 1
>
> Click here for file
> [http://www.biomedcentral.com/content/supplementary/1471-2377-3-4-S1.doc]
>
> ### Additional file 2
>
> Click here for file
> [http://www.biomedcentral.com/content/supplementary/1471-2377-3-4-S2.doc]
>
> ### Additional file 3
>
> Click here for file
> [http://www.biomedcentral.com/content/supplementary/1471-2377-3-4-S3.doc]

References
1. Executive Committee for the Asymptomatic Carotid Atherosclerosis Study: **Endarterectomy for asymptomatic carotid artery stenosis** *JAMA* 1995, **273:**1421-1428.
2. Gambaro G, Anglani F and D'Angelo A: **Association studies of genetic polymorphisms and complex disease** *Lancet* 2000, **355:**308-311.
3. del Zoppo GJ, Higashida RT, Furlan AJ, Pessin MS, Rowley HA and Gent M: **PROACT: A phase II randomized trial of recombinant pro-urokinase by direct arterial delivery in acute middle cerebral artery stroke. PROACT Investigators. Prolyse in Acute Cerebral Thromboembolism** *Stroke* 1998, **29:**4-11.
4. Furlan A, Higashida R, Wechsler L, Gent M, Rowley H, Kase C, Pessin M, Ahya A, Callahan F and Clark WM *et al.*: **Intra-arterial prourokinase for acute ischemic stroke. The PROACT II study: a randomized controlled trial. Prolyse in Acute Cerebral Thromboembolism** *JAMA* 1999, **282:**2003-2011.
5. Lewandowski CA, Frankel M, Tomsick TA, Broderick J, Frey J, Clark W, Starkman S, Grotta J, Spilker J and Khoury J *et al.*: **Combined intravenous and intra-arterial r-TPA versus intra-arterial therapy of acute ischemic stroke: Emergency Management of Stroke (EMS) Bridging Trial** *Stroke* 1999, **30:**2598-2605.
6. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group: **Tissue plasminogen activator for acute ischemic stroke** *N Engl J Med* 1995, **333:**1581-1587.
7. Antiplatelet Trialists' Collaboration: **Collaborative overview of randomised trials of antiplatelet therapy – I: prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients** *BMJ* 1994, **308:**81-106.
8. Hart RG, Benavente O, McBride R and Pearce LA: **Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis** *Ann Intern Med* 1999, **131:**492-501.
9. Zuber M, Toulon P, Marnet L and Mas JL: **Factor V Leiden mutation in cerebral venous thrombosis** *Stroke* 1996, **27:**1721-1723.
10. Kessler C, Spitzer C, Stauske D, Mende S, Stadlmuller J, Walther R and Retlig R: **The apolipoprotein E and beta-fibrinogen G/A-455 gene polymorphisms are associated with ischemic stroke involving large-vessel disease** *Arterioscler Thromb Vasc Biol* 1997, **17:**2880-2884.
11. Schmidt H, Schmidt R, Niederkorn K, Horner S, Becsagh P, Reinhart B, Schumacher M, Weinrauch V and Kostner GM: **Beta-fibrinogen gene polymorphism (C148→T) is associated with carotid**

atherosclerosis: results of the Austrian Stroke Prevention Study** *Arterioscler Thromb Vasc Biol* 1998, **18:**487-492.

12. Carter AM, Catto AJ, Bamford JM and Grant PJ: **Gender-specific associations of the fibrinogen B beta 448 polymorphism, fibrinogen levels, and acute cerebrovascular disease** *Arterioscler Thromb Vasc Biol* 1997, **17:**589-594.

13. Carlsson LE, Santoso S, Spitzer C, Kessler C and Greinacher A: **The alpha2 gene coding sequence T807/A873 of the platelet collagen receptor integrin alpha2beta1 might be a genetic risk factor for the development of stroke in younger patients** *Blood* 1999, **93:**3583-3586.

14. Gonzalez-Conejero R, Lazano ML, Rivera J, Corral J, Iniesta JA, Moraleda JM and Vicente V: **Polymorphisms of platelet membrane glycoprotein Ib associated with arterial thrombotic disease** *Blood* 1998, **92:**2771-2776.

15. Ridker PM, Hennekens CH, Lindpaintner K, Stampfer MJ, Eisenberg PR and Miletich JP: **Mutation in the gene coding for coagulation factor V and the risk of myocardial infarction, stroke, and venous thrombosis in apparently healthy men** *N Engl J Med* 1995, **332:**912-917.

16. Carter AM, Catto AJ, Bamford JM and Grant PJ: **Platelet GP IIIa P1A and GP Ib variable number tandem repeat polymorphisms and markers of platelet activation in acute stroke** *Arterioscler Thromb Vasc Biol* 1998, **18:**1124-1131.

17. Adams HP Jr, Brott TG, Crowell RM, Furlan AJ, Gomez CR, Grotta J, Helgason CM, Marler JR, Woolson RF and Zivin JA *et al.*: **Guidelines for the management of patients with acute ischemic stroke: a statement for healthcare professionals from a special writing group of the Stroke Council, American Heart Association** *Stroke* 1994, **25:**1901-1914.

18. Adams HP Jr, Brott TG, Furlan AJ, Gomez CR, Grotta J, Helgason CM, Kwiatkowski T, Lyden PD, Marler JR and Torner J *et al.*: **Guidelines for Thrombolytic Therapy for Acute Stroke: a Supplement to the Guidelines for the Management of Patients with Acute Ischemic Stroke. A statement for healthcare professionals from a Special Writing Group of the Stroke Council, American Heart Association** *Stroke* 1996, **27:**1711-1718.

19. WHO MONICA Project Principal Investigators: **The World Health Organization MONICA Project (monitoring trends and determinants in cardiovascular disease): a major international collaboration** *J Clin Epidemiol* 1988, **41:**105-114.

20. McCarron MO, Muir KW, Weir CJ, Dyker AG, Bone I, Nicoll JA and Lees KR: **The apolipoprotein E epsilon4 allele and outcome in cerebrovascular disease** *Stroke* 1998, **29:**1882-1887.

21. Catto AJ, McCormack LJ, Mansfield MW, Carter AM, Bamford JM, Robinson P and Grant PJ: **Apolipoprotein E polymorphism in cerebrovascular disease** *Acta Neurol Scand* 2000, **101:**399-404.

22. Carter AM, Catto AJ, Bamford JM and Grant PJ: **Association of the platelet glycoprotein IIb HPA-3 polymorphism with survival after acute ischemic stroke** *Stroke* 1999, **30:**2606-2611.

23. Carter AM, Catto AJ and Grant PJ: **Association of the alpha-fibrinogen Thr312A1a polymorphism with poststroke mortality in subjects with atrial fibrillation** *Circulation* 1999, **99:**2423-2426.

24. Meschia JF, Brown RD Jr, Brott TG, Chukwudelunzu FE, Hardy J and Rich SS: **The Siblings With Ischemic Stroke Study (SWISS) Protocol** *BMC Med Genet* 2002, **3:**1.

25. Meschia JF, Brott TG, Chukwudelunzu FE, Hardy J, Brown RD Jr, Meissner I, Hall LJ, Atkinson EJ and O'Brien PC: **Verifying the stroke-free phenotype by structured telephone interview** *Stroke* 2000, **31:**1076-1080.

26. Jones WJ, Williams LS, Redmon G and Meschia JF: **Validating the questionnaire to verify stroke-free status by patient history and physical examination [abstract]** *Stroke* 2001, **32:**362.

27. Worrall BB, Brown DL, Brott TG, Brown RD, Silliman SL and Meschia JF: **Spouses and unrelated friends of probands as controls for stroke genetics studies** *Neuroepidemiology* 2003, **22:**239-244.

28. Worrall BB, Chen DT and Meschia JF: **Ethical and methodological issues in pedigree stroke research** *Stroke* 2001, **32:**1242-1249.

29. Sacco RL, Kargman DE and Zamanillo MC: **Race-ethnic differences in stroke risk factors among hospitalized patients with cerebral infarction: the Northern Manhattan Stroke Study** *Neurology* 1995, **45:**659-663.

30. Marler JR, Tilley BC, Lu M, Brott TG, Lyden PC, Grotta JC, Broderick JP, Levine SR, Frankel MP and Horowitz SH *et al.*: **Early stroke treatment associated with better outcome: the NINDS rt-PA stroke study** *Neurology* 2000, **55:**1649-1655.

31. Lyden P, Lu M, Jackson C, Marler J, Kothari R, Brott T and Zivin J: **Underlying structure of the National Institutes of Health Stroke Scale: results of a factor analysis. NINDS tPA Stroke Trial Investigators** *Stroke* 1999, **30:**2347-2354.

32. Bamford JM, Sandercock PAG, Warlow CP and Slatterly J: **Interobserver agreement for the assessment of handicap in stroke patients (letter)** *Stroke* 1989, **20:**828.

33. Collin C, Wade DT, Davies S and Horne V: **The Barthel ADL Index: a reliability study** *Int Disabil Stud* 1988, **10:**61-63.

34. Shinar D, Gross CR, Bronstein KS, Licata-Gehr EE, Eden DT, Cabrera AR, Fishman IG, Roth AA, Barwick JA and Kunitz SC: **Reliability of the activities of daily living scale and its use in telephone interview** *Arch Phys Med Rehabil* 1987, **68:**723-728.

35. Jennett B and Bond M: **Assessment of outcome after severe brain damage** *Lancet* 1975, **1:**480-484.

36. Madden KP, Karanjia PN, Adams HP Jr and Clarke WR: **Accuracy of initial stroke subtype diagnosis in the TOAST study: Trial of ORG 10172 in Acute Stroke Treatment** *Neurology* 1995, **45:**1975-1979.

37. Adams HP Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL and Marsh EE III: **Classification of subtype of acute ischemic stroke: definitions for use in a multicenter clinical trial. TOAST: Trial of Org 10172 in Acute Stroke Treatment** *Stroke* 1993, **24:**35-41.

38. Bamford J, Sandercock P, Dennis M, Burn J and Warlow C: **Classification and natural history of clinically identifiable subtypes of cerebral infarction** *Lancet* 1991, **337:**1521-1526.

39. Johnson CJ, Kittner SJ, McCarter RJ, Sloan MA, Stern BJ, Buchholz D and Price TR: **Interrater reliability of an etiologic classification of ischemic stroke** *Stroke* 1995, **26:**46-51.

40. Miller SA, Dykes DD and Polesky HF: **A simple salting out procedure for extracting DNA from human nucleated cells** *Nucleic Acids Res* 1988, **16:**1215.

41. Hosmer DW Jr and Lemeshow S: **Applied Logistic Regression** *New York, Wiley* 1989.

42. Goodman LA: **Simple methods for analyzing three-factor interaction in contingency tables** *J Am Stat Assoc* 1964, **59:**319-352.

43. Roberts R and Ryan TJ: **Task force 3: clinical research in a molecular era and the need to expand its ethical imperatives** *J Am Coll Cardiol* 1998, **31:**942-949.

44. The Ad Hoc Committee on Genetic Testing/Insurance Issues: **Genetic testing and insurance** *Am J Hum Genet* 1995, **56:**327-331.

45. Foster MW and Freeman WL: **Naming names in human genetic variation research (editorial)** *Genome Res* 1998, **8:**755-757.

46. Fuller BP, Kahn MJ, Barr PA, Biesecker L, Crowley E, Garber J, Mansoura MK, Murphy P, Murray J and Phillips J *et al.*: **Privacy in genetics research** *Science* 1999, **285:**1359-1361.

47. Botkin J: **Protecting the privacy of family members in survey and pedigree research** *JAMA* 2001, **285:**207-211.

48. Meschia JF: **The problem of ignoring interconnectedness in genetic research (letter)** *J Med Ethics* 2000, **26:**477.

49. Sommerville A and English V: **Genetic privacy: orthodoxy or oxymoron?** *J Med Ethics* 1999, **25:**144-150.

50. Gulcher JR and Stefansson K: **The Icelandic Healthcare Database and informed consent** *N Engl J Med* 2000, **342:**1827-1830.

51. Annas GJ: **Rules for research on human genetic variation – lessons from Iceland** *N Engl J Med* 2000, **342:**1830-1833.

52. North American Symptomatic Carotid Endarterectomy Trial Collaborators: **Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis** *N Engl J Med* 1991, **325:**445-453.

53. Meschia JF, Brott TG, Brown RD Jr, Kissela PM and Rich SS: **Planning genetic studies and human stroke: sample size estimates based on family history data** *Neurology* 2002, **59:**1292.

54. Meschia JF: **Subtyping in ischemic stroke genetic research** *Semin Cerebrovasc Dis Stroke* 2002, **2:**2-13.

55. Meschia JF: **Addressing heterogeneity of the ischemic stroke phenotype in human genetics research** *Stroke* 2002, **33:**2770-2774.

56. Catto A, Carter A, Ireland H, Bayston TA, Philippou H, Barrett J, Lane DA and Grant PJ: **Factor V Leiden gene mutation and thrombin generation in relation to the development of acute stroke** *Arterioscler Thromb Vasc Biol* 1995, **15:**783-785.

57. Catto A, Carter A and Grant PJ: **Factor V Leiden mutation and completed stroke (letter)** *Stroke* 1996, **27:**573.

58. Kontula K, Ylikorkala A, Miettinen H, Vuorio A, Kauppinen-Makelin R, Hamalainen L, Palomaki H and Kaste M: **Arg506Gln factor V mutation (factor V Leiden) in patients with ischaemic cerebrovascular disease and survivors of myocardial infarction** *Thromb Haemost* 1995, **73:**558-560.

59. Martinelli I, Franchi F, Akwan S, Bettini P, Merati G and Mannucci PM: **The transition G to A at position 20210 in the 3'-untranslated region of the prothrombin gene is not associated with cerebral ischemia (letter)** *Blood* 1997, **90:**3806.

60. Press RD, Liu XY, Beamer N and Coull PM: **Ischemic stroke in the elderly: role of the common factor V mutation causing resistance to activated protein C** *Stroke* 1996, **27:**44-48.

61. van der Bom JG, Bots ML, Haverkate F, Slagboom PE, Meijer P, de Jong PT, Hofman A, Grobbee DE and Kluft C: **Reduced response to activated protein C is associated with increased risk for cerebrovascular disease** *Ann Intern Med* 1996, **125:**265-269.

62. Longstreth WT Jr, Rosendaal FR, Siscovick DS, Vos HL, Schwartz SM, Psaty BM, Raghunathan TE, Koepsell TD and Reitsma PH: **Risk of stroke in young women and two prothrombotic mutations: factor V Leiden and prothrombin gene variant (G20210A)** *Stroke* 1998, **29:**577-580.

63. Halbmayer WM, Haushofer A, Angerer V, Finsterr J and Fischer M: **APC resistance and factor V Leiden (FV:Q506) mutation in patients with ischemic cerebral events: Vienna Thrombophilia in Stroke Study Group (VITISS)** *Blood Coagul Fibrinolysis* 1997, **8:**361-364.

64. Markus HS, Zhang Y and Jeffrey S: **Screening for the Factor-V Arg 506 Gln mutation in patients with TIA and stroke** *Cerebrovasc Dis* 1996, **6:**360-362.

65. Nishiuma S, Kario K, Nakanishi K, Yakushijin K, Kageyama G, Matsunaka T, Matsuo T, Kanai N, Ikeda U, Shimada K and Matsuo M: **Factor VII R353Q polymorphism and lacunar stroke in Japanese hypertensive patients and normotensive controls** *Blood Coagul Fibrinolysis* 1997, **8:**525-530.

66. Heywood DM, Carter AM, Catto AJ, Bamford JM and Grant PJ: **Polymorphisms of the factor VII gene and circulating FVII:C levels in relation to acute cerebrovascular disease and poststroke mortality** *Stroke* 1997, **28:**816-821.

67. Corral J, Gonzalez-Conejero R, Lozano ML, Rivera J and Vicente V: **Genetic polymorphisms of factor VII are not associated with arterial thrombosis** *Blood Coagul Fibrinolysis* 1998, **9:**267-272.

68. Catto AJ, Kohler HP, Bannan S, Stickland M, Carter A and Grant PJ: **Factor XIII Val 34 Leu: a novel association with primary intracerebral hemorrhage** *Stroke* 1998, **29:**813-816.

69. Carlsson LE, Greinacher A, Spitzer C, Walther R and Kessler C: **Polymorphisms of the human platelet antigens HPA-1, HPA-2, HPA-3, and HPA-5 on the platelet receptors for fibrinogen (GPIIb/IIIa), von Willebrand factor (GPIb/IX), and collagen (GPIa/IIa) are not correlated with an increased risk for stroke** *Stroke* 1997, **28:**1392-1395.

70. Ridker PM, Hennekens CH, Schmitz C, Stampfer MJ and Lindpaintner K: **PIA1/A2 polymorphism of platelet glycoprotein IIIa and risks of myocardial infarction, stroke, and venous thrombosis** *Lancet* 1997, **349:**385-388.

71. Catto AJ, Carter AM, Stickland M, Bamford JM, Davies JA and Grant PJ: **Plasminogen activator inhibitor-1 (PAI-1) 4G/5G promoter polymorphism and levels in subjects with cerebrovascular disease** *Thromb Haemost* 1997, **77:**730-734.

72. Poort SR, Rosendaal FR, Reitsma PH and Bertina RM: **A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis** *Blood* 1996, **88:**3698-3703.

73. Ridker PM, Hennekens CH and Miletich JP: **G20210A mutation in prothrombin gene and risk of myocardial infarction, stroke, and venous thrombosis in a large cohort of US men** *Circulation* 1999, **99:**999-1004.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2377/3/4/prepub