Polygenic Disease: A Study of Genetic Risk in an Australian Stroke Population

The Adelaide Genetic Stroke Study

Jim Jannes

Department of Medicine
The University of Adelaide
South Australia

A thesis submitted in fulfilment of the requirements
for the degree of PhD in Medicine
February 2004

Table of Contents

Chapter 1: Introduction

Chapter 2: Genetics and Ischemic Stroke

Chapter 3: Thesis Aims and Rationale

Chapter 4: Research Methods

Chapter 5: Results

Chapter 6: Discussion

Chapter 7: Future Directions

Thesis Abstract

Twin, family and animal studies support this thesis that ischemic stroke is a polygenic disease. The magnitude of this predisposition varies according to stroke subtype, with the greatest risk associated with lacunar and atherothromboembolic stroke. To date, the precise genetic determinants remain unknown.

The primary aim of this thesis was to determine the risk of ischemic stroke associated with eight single nucleotide polymorphisms (SNPs) that were selected using a candidate gene approach: Paraoxonase (PON1) –107T/C and M54L, Glycoprotein 1b 145Thr/Met, Glycoprotein IIb/IIIa PlA1/A2, β fibrinogen –148 C/T, Prothrombin 20210 G/A, Tissue Plasminogen Activator (TPA) –7,351 C/T and Plasminogen Activator Inhibitor (PAI-1) 5G/4G. This thesis also aimed to determine the relevance of each SNP to ischemic stroke subtypes and to determine the effect of interaction between each SNP and known cerebrovascular risk factors.

The objectives were met using a case-control study that recruited hospital inpatients with a diagnosis of acute ischemic stroke. Patients were evaluated for known cerebrovascular risk factors and classified for stroke subtype. A cerebrovascular risk factor profile was also determined in a randomly selected, age and gender matched control group. The SNP genotypes were determined using a polymerase chain reaction (PCR) method. Logistic regression was used to determine the risk of ischemic stroke associated with each SNP.

During a 26-month period, 182 patients and 301 non-hospitalised controls consented to participate. In a multivariate model that adjusted for important confounders, a 1.9-fold (95%CI 1.01-3.6) increased risk of ischemic stroke was associated with the TPA –7,351 TT genotype. This association, however, was not significant in a multivariate model that incorporated all potential confounders (OR 1.8, 95%CI 0.9-3.4). In a subgroup analysis, a statistically significant 2.6 and 2.4-fold increased risk of lacunar

stroke was associated with the TPA –7,351 TT and PON1 -107 CC genotypes respectively. No other association or effect of interaction was observed.

The findings suggest that TPA –7,351 C/T and PON1 -107 T/C SNP's may play a role in the pathogenesis of lacunar stroke. Confirmation by a larger study of greater statistical power is required, which may then provide a better means to predict the risk of lacunar stroke.

Acknowledgements

This thesis would not be possible without the input from many people. I am sincerely grateful to my main supervisor, Dr Simon Koblar, for his belief in my ability, guidance and mentorship throughout this PhD. Dr Koblar's influence has extended beyond enhancing my capacity to perform scientific research, for which I am truly grateful. I also thank my other supervisors, Associate Professors Brian Smith and Louis Pilotto for their advice, particularly during the planning phase of this thesis.

My sincere gratitude also extends to Dr Anne Hamilton-Bruce for her tireless effort towards many aspects of this thesis. I am particularly grateful for her assistance in the recruitment of community controls and the numerous hours spent reviewing manuscripts and thesis chapters.

I also thank Mr Anthony Condina, Mr Steven Davis, Ms Robyn Attewell, Ms Bernadette Kenny, Dr Charles Mulligan, Dr Warren Flood and The Australian Red Cross (South Australia) tissue-typing laboratory for their technical assistance.

Finally, I sincerely thank my wife Mary and my two daughters, Kathryn and Nicola, for their patience and loving support. Gratitude also extends to my mother and father-in-law and my parents for providing the opportunity to pursue my career aspirations. I hope that their sacrifice is at least in part rewarded by the completion of this thesis.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference is made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Signed:..... ..... Date: 24/2/04.....

Conference Presentations

Poster Presentation:
"Polygenic Disease: A Study of Genetic Risk in an Australian Stroke Population" Australasian Association of Neurologists Annual Scientific Meeting, Adelaide, May 2000.

Poster Presentation:
"Polygenic Disease: A Study of Genetic Risk in an Australian Stroke Population" Australian Neuroscience Society, 23rd Annual Scientific Meeting, Adelaide, January 2003.

Platform Presentation:
"Tissue Plasminogen Activator –7,351 Enhancer Polymorphism is a Risk Factor for Lacunar Stroke". The Stroke Society of Australasia Annual Scientific Meeting, Sydney, September 2003.
Recipient of the "New Investigator Award for Scientific Research"

Platform Presentation:
"Tissue Plasminogen Activator –7,351 Enhancer Polymorphism is a Risk Factor for Lacunar Stroke". The Queen Elizabeth Hospital Research Day, Adelaide, October 2003.
Recipient of the "Best Higher Degree Research Award" and the "Allan Kerr Grant Most Outstanding Postgraduate Learner Award".

Platform Presentation:
"Tissue Plasminogen Activator –7,351 Enhancer Polymorphism is a Risk Factor for Lacunar Stroke". ASMR 42nd National Scientific Conference, Adelaide, November 2003.

List of Figures

List of Tables

Abbreviations

Adenosine	A
Adenosine Diphosphate	ADP
Computerised Tomography	CT
Cytosine	C
Deoxyribonucleic Acid	DNA
Diastolic Blood Pressure	DBP
Disability Adjusted Life Year	DALY
Glycoprotein	Gp
Guanine	G
High Density Lipoprotein	HDL
Human Platelet Alloantigen	HPA
Lacunar Syndrome	LS
Leucine	L
Low Density Lipoprotein	LDL
Magnetic Resonance Imaging	MRI
Messenger Ribodeoxynucleic Acid	MRNA
Metaloproteases	MMP
Methionine	M
National Heart Foundation	NHF
North East Melbourne Stroke Incidence Study	NEMESIS
Oxfordshire Community Stroke Project	OCSP
Paraoxonase	PON1
Partial Anterior Circulation Syndrome	PACS
Patent Foramen Ovale	PFO
Perth Community Stroke Study	PCSS
Plasminogen Activator Inhibitor	PAI
Polymerase Chain Reaction	PCR
Population Research and Outcome Studies	PROS
Posterior Circulation Syndrome	PCS

Ribonucleic Acid	RNA
Sequence Specific Primer Polymerase Chain Reaction	SSP-PCR
Sibling Transmission Disequilibrium Test	S-TDT
Single Nucleotide Polymorphism	SNP
Spontaneously Hypertensive Rat	SHR
Stroke Prone Spontaneously Hypertensive Rat	SP-SHR
Systolic Blood Pressure	SBP
The Trial of ORG 10172 in Acute Stroke Treatment	TOAST
Threonine	Thr
Thymidine	T
Tissue Plasminogen Activator	TPA
Total Anterior Circulation Syndrome	TACS
Transcription factor IID	TFIID
Transient Ischemic Attack	TIA
Transmission Disequilibrium Test	TDT
World Health Organization	WHO

Chapter 1

Introduction

1.1 The Global Burden of Stroke

Amongst all the adult neurological conditions known, cerebrovascular disease remains the most prevalent and poses the greatest disease burden on the developed world. In 2001, an estimated 5.4 million people died from cerebrovascular disease worldwide, placing it second to ischaemic heart disease (WHO, 2002). To put this into perspective, death from cerebrovascular disease was almost double that caused by human immunodeficiency virus infection (2.8 million deaths), quadruple that of lung cancer (1.2 million deaths) and approximately six-times higher than death caused by renal disease (825,000 deaths) or diabetes (895,000 deaths). Furthermore, it has been estimated that an additional 15 million non-fatal strokes occur annually, and that of the 50 million stroke survivors currently alive, one in five will suffer a recurrent stroke within five years (Chalmers and Chapman, 2001).

In 1997, the results of the Global Burden of Disease Study provided new insight into global impact of cerebrovascular disease (Murray and Lopez, 1997b). This study was completed in collaboration with the World Health Organization, and estimated worldwide and regional cause-of-death patterns in 1990, for 14 age-gender groups in eight geographical regions. The authors also pioneered a new summary measure for disease burden, the 'DALY' (disability-adjusted life year), which has been widely adopted subsequently in population health statistics. The 'DALY' incorporates the impact of both disease morbidity and mortality, by combining years of life lost due to premature mortality, with years lived with disability. When measured in disability-adjusted life years, cerebrovascular disease was the sixth greatest contributor to global disease burden, following lower respiratory tract infections, diarrhoeal diseases, perinatal conditions, unipolar major depression and ischemic heart disease (in ranked

order). It projected that by 2020, if immediate and effective strategies were not undertaken, cerebrovascular disease would increase to become the second greatest cause of death and disability in developed countries, ranking fourth worldwide (Murray and Lopez, 1997a).

1.2 Stroke: an Australian Perspective

In Australia, cerebrovascular disease has mirrored trends seen in other developed countries and remains a major public health challenge. It is estimated that 116,500 Australians, or 0.6% of the population, have suffered a stroke, with the burden compounded by 40,000 new strokes each year (NHF, 2001). Over the last twelve years, mortality in Australia from cerebrovascular disease has declined by 34.9% for males, and 34.7% for females; this is attributable to a decrease in disease incidence rather than improved outcome (Jamrozik *et al.*, 1999). Despite this downward trend, in 2001 cerebrovascular disease remained Australia's greatest killer second to ischemic heart disease with 12,146 deaths recorded (ABS, 2001).

In the last decade, The Perth Community Stroke Study (PCSS) and The North East Melbourne Stroke Incidence Study (NEMESIS) have provided further insight into the true incidence and natural history of stroke in Australia. Both were well-designed, population-based cohort studies that focused on accurate clinical diagnosis and ascertainment of all stroke patients within geographically defined regions, including those patients who did not present to hospital (Thrift *et al.*, 2001, Jamrozik *et al.*, 1999). Comparison using census data for the remainder of their respective cities showed that the recruited cohorts were of similar socio-demographic composition and generally representative of the entire urban population. The PCSS registered 536 acute cerebrovascular events in 492 patients (370 patients with first-ever stroke), from the inner metropolitan region of Perth, Western Australia, for an 18-month period during 1989 to 1990. The investigators used the same methodology to ascertain a further 290 stroke events in 281 patients (213 patients with first-ever stroke) for a 13-month period during 1995 to 1996. This provided the first longitudinal, population-based data for various stroke indices in Australia. The study showed that the overall incidence for all first-ever strokes standardized for age and gender to the world population fell from 104

per 100,00 persons in 1989-1990 to 76 per 100,000 persons in 1995-1996. As the 28-day case mortality did not change significantly between 1989-1990 and 1995-1996 (23% vs 24%), the authors concluded that the fall in the Australian stroke mortality rate was due primarily to a significant decline in stroke incidence, rather than to an improvement in stroke outcome (Jamrozik *et al.*, 1999).

The NEMESIS was designed to determine the true incidence of stroke in Melbourne, Australia. Preliminary data had suggested that the stroke incidence overall in this region was significantly higher than that reported in the PCSS. The former study identified 381 stroke events amongst 353 patients in a well-defined area of inner northeast Melbourne, for a 12-month period during 1996-1997 (Thrift *et al.*, 2000). The overall annual first-ever stroke incidence rate standardized for age and gender to the world population was 100 per 100,000 persons, a finding that was considerably higher than that reported in the PCSS (76 per 100,000 persons in 1995-1996), but similar to that found in other developed countries (Thrift *et al.*, 2000). A subgroup analysis in the NEMESIS showed that the observed difference in stroke incidence was mainly due to a greater incidence of ischemic stroke amongst men, and intracerebral haemorrhage amongst women (Thrift *et al.*, 2001). There was however, insufficient raw data from the PCSS to determine if this difference achieved statistical significance. It is possible that the difference in stroke incidence between Australian states reflects differences in recruitment and stroke classification methodology. Alternatively, it may be explained by the ethnic, demographic and lifestyle diversity between regions.

Cerebrovascular disease is Australia's leading cause of long-term disability with an estimated 63,530 stroke survivors affected. Of these, three quarters require assistance with mobility, self-care or communication (NHF, 2001). In the Australian Burden of Disease Study, cerebrovascular disease was ranked the second largest contributor to the disease mortality burden in Australia, representing 8.3% of the total years of life lost from all conditions in 1996 (Mathers *et al.*, 2000). When measured in disability-adjusted life years, cerebrovascular disease accounted for 5.4% of Australia's total burden of disease and injury, second only to ischaemic heart disease. The long-term disability data after first-ever stroke in Australia has also been reported recently

(Hankey *et al.*, 2002). This study presented the 5-year outcome from the original cohort of 462 stroke patients recruited in the PCSS between 1989-1990. The study found that of the 277 cases (76.5%) that had survived beyond 28 days, more than half were alive at five years (55%). Of these long-term survivors, 1 in 7 (14.4%) were placed in permanent institutionalised care subsequently, and approximately one third were left with a major disability (36%). The cumulative risk of death or disability at 5 years amongst 30-day stroke survivors was 68%.

The enormous health burden that cerebrovascular disease places on the Australian community comes at a great financial cost. The total health cost attributable to stroke was $630.5 million in 1994, representing approximately 17% of total health system cost for all circulatory system disorders (NHF, 2001). Of this, approximately a third was spent on public hospital admissions, a third on providing nursing home beds, with the remainder distributed amongst other medical, private and allied health services. Interestingly, less than 1% of the total stroke budget ($5.9 million) was allocated to cerebrovascular disease research (NHF, 2001). By extrapolating data obtained from the NEMESIS study, Dewey *et al.* developed a comprehensive incidence–based cost-of-illness model that was used to estimate the total cost of stroke in Australia in 1997 (Dewey *et al.*, 2001b). The study concluded that in 1997, the total cost burden for first-ever strokes to Australia was $555 million, of which two thirds contributed to the cost of acute hospitalisation, inpatient rehabilitation and nursing home care. The average cost per case in the first year was estimated at $18,956, with an estimated lifetime cost of $44,428.

1.3 Classification of Stroke

1.3.1 Stroke Subtypes

The World Health Organization defines stroke as syndrome of "rapidly developing clinical signs of focal (or global) disturbance of cerebral function lasting more than 24 hours (unless interrupted by surgery or death) with no apparent cause other than that of vascular origin" (Hatono, 1976). The definition incorporates a wide range of vascular pathologies including atherosclerosis, vasculitis, aneurysmal dilatation, arterial dissection and developmental malformations. These cause two main types of

secondary brain parenchymal change: cerebral infarction or haemorrhage (both intracerebral and sub-arachnoid haemorrhage). In Australia, the NEMESIS found cerebral infarction to be the most common subtype of stroke, responsible for nearly three quarters (72.5%) of all first–ever strokes. This equates to a crude annual incidence rate of ischaemic stroke of 149 per 100,000 persons, a finding comparable to that of other developed countries (Thrift *et al.*, 2001). Intracerebral haemorrhage was the second most common stroke subtype (14.5%), followed by sub-arachnoid haemorrhage (4.3%). Insufficient clinical information prevented classification in the remaining 8.7% (Thrift *et al.*, 2001). Similar findings were reported by Hankey *et al.* in the PCSS (Hankey *et al.*, 2002). Furthermore, the balance of stroke subtype in the PCSS did not change significantly between cohorts recruited in 1989-1990 and 1995-1996.

1.3.2 Pathophysiological Classification of Ischemic Stroke

Cerebral infarction is often classified on the basis of arterial pathology, with most cases falling into one of three categories: large-vessel disease, small-vessel disease and cardio-embolic infarction. In 30-40% of patients, accurate categorization is not possible, and these cases are often labelled as infarction of unknown type (Whisnant, 1990). Atherosclerosis underlies the pathogenesis of large-vessel occlusion. This was first described in 1967 by Constantinides, who performed a post-mortem microscopic examination of the major branches of the circle of Willis from ten consecutive patients with large-vessel ischemic stroke (Constantinides, 1967). In all cases, vascular occlusion was caused by acute thrombosis that was associated with fissures within adjacent atherosclerotic lesions. As haemorrhage was observed within the atherosclerotic plaque and collagen fragments were found deep within the thrombotic core, arterial wall fissuring could not be explained by post-mortem artefact. The authors concluded that acute thrombotic occlusion occurred following rupture of atherosclerotic lesions and exposure of blood to subendothelial tissues or expelled intra-luminal collagen.

Since this discovery was made, there has been a considerable advance in the understanding of atherosclerotic plaque vulnerability. It is now known that the risk of

atherosclerotic plaque rupture is determined by interplay of cellular, structural, molecular and mechanical factors. The accumulation of plaque cholesterol ester, reduction of plaque collagen density, influx of inflammatory cells, reduced density of smooth muscle cells and secretion of matrix-digesting enzymes have all been implicated in the pathogenesis of plaque rupture (Kullo *et al.*, 1998). Atherosclerotic lesions typically affect specific sites along the cerebral arterial vasculature, including the carotid bifurcation, carotid siphon, proximal segment of the middle cerebral artery (M1 segment), segments 1 and 4 of the vertebral arteries and the proximal basilar artery (Mohr *et al.*, 1997). These sites are regions of persistent high mechanical stress that create an uneven distribution of tension on the luminal surface of vulnerable atherosclerotic plaque. As the plaque is stretched, the capacity to withstand a mechanical load lessens, with eventual plaque rupture (Kullo *et al.*, 1998). Apart from causing in-situ thrombotic occlusion, atherosclerotic lesions may also cause cerebral infarction by severely compromising distal flow of oxygenated blood during times of decreased global cerebral perfusion (e.g. cardiac failure, peri-operatively), or due to plaque fragmentation and distal embolism (Whisnant, 1990).

Atrial fibrillation, anterior myocardial infarction, mitral valve stenosis, and prosthetic heart valves are all known to predispose to intra-mural cardiac thrombus formation and are established causes of cardio-embolic stroke. Less frequently, the embolic source may arise from cardiac tumours, endocarditis, or due to a patent foramen ovale associated with an atrial septal aneurysm allowing for paradoxical venous embolism. As emboli pursue a path of least resistance, cardio-embolic strokes typically cause cortical infarction in the distribution of the proximal middle cerebral artery or its distal branches (Whisnant, 1990).

Lacunar infarcts are small rounded lesions (less than 1.5cm in diameter) located commonly within the internal capsule, thalamus and striatum. They arise from the occlusion of deep penetrating arteries which branch at right angles from their main intracerebral arteries (e.g. lenticulostriate arteries). Diagnosis of lacunar syndrome is made via neuro-imaging or by determining if the clinical presentation fits a "classic" lacunar syndrome: pure motor hemiparesis; sensorimotor stroke; pure sensory stroke;

dysarthria-clumsy-hand syndrome and ataxic hemiparesis. To date, much of the understanding of the pathophysiology of lacunar infarction is derived from the early work by Charles Miller Fisher (Fisher, 1979). Through meticulous post-mortem microscopic examination of brain tissue in a series of patients who had suffered infarction of the internal capsule, two pathological processes were identified. In most cases microatheroma was identified as the pathology underlying lacunar stroke. This was typically located at the proximal portion of the small vessel or at its origin. Vascular occlusion occurred either due to superimposed thrombosis, implying a similar mechanism to that seen in large vessel disease, or due to luminal obliteration causing severe hemodynamic compromise. The second mechanism, termed lipohyalinosis or complex small vessel disease, involved an acute phase of fibrinoid necrosis followed by the deposition of fibrin and collagen that effectively obliterated the vessel lumen. In a minority of cases, where no vascular lesions were identified, cerebral embolism was presumed to have occurred (Fisher, 1979).

The pathophysiological classification of cerebral infarction has become an essential component in the management of ischemic stroke, allowing the treating physician to implement case-specific acute and secondary prevention therapies. Classification of cerebral infarction in this manner is also important for research purposes, particularly in studies investigating the aetiology of ischemic stroke where disease-causing factors are likely to differ between pathological subtypes. In these studies, a classification system created by The Trial of ORG 10172 in Acute Stroke Treatment (TOAST) investigators, is used frequently (Kolominsky-Rabas *et al.*, 2001). This categorizes cerebral infarction into four subtypes: (a) large-artery atherosclerosis (includes large artery thrombosis and artery-to-artery embolism); (b) cardioembolism; (c) small artery occlusion; and (d) stroke of undetermined cause.

Although robust, the TOAST classification system relies greatly on sophisticated imaging technologies that may not be readily available in some centres. Its use in larger epidemiological and clinical studies may thus be impracticable. The investigators from the Oxfordshire Community Stroke Project (OCSP) overcame this problem with a simple clinical classification system for cerebral infarction that can be

performed by clinicians and nursing staff at the patient's bedside. This method uses the findings of the clinical examination to predict the anatomical site of pathology, and categorizes cerebral infarction into four subtypes: total anterior circulation syndrome (TACS); partial anterior circulation syndrome (PACS); posterior circulation syndrome (POCS); and lacunar syndrome (Appendix I) (Bamford *et al.*, 1991). The OCSP showed significant differences in the natural history between cerebral infarction groups. Patients presenting with TACS had a negligible chance of good functional recovery, with high mortality, whilst stroke reoccurrence was highest in those with PACS (17% at 1 year) (Bamford *et al.*, 1991). The OCSP stroke classification has been subsequently validated by Mead *et.al.,* who showed that there was high correlation between stroke subtype and intra-cranial artery involvement and that it could accurately predict the anatomical site of cerebral infarction as seen on computerized tomography scanning (Mead *et al.*, 2000b, Mead *et al.*, 2000a). Furthermore, Dewey *et al.*, found a moderate agreement between neurologists, and concluded that the OCSP stroke classification system was a reliable tool that could be applied in community-based epidemiological studies of cerebral infarction (Dewey *et al.*, 2001a).

1.4 Thesis Overview

This thesis will address the relevance of 8 common SNP's to ischemic stroke. Chapter 2 will address the evidence for a genetic predisposition of ischemic stroke and conclude with a review of the literature for each of the candidate SNP's. Chapter 3 will discuss the choice of study design and the criteria for candidate gene selection. Chapters 4 and 5 will present the study methods and results respectively. A detailed discussion of the findings from both a pathogenetic and statistical perspective is provided in Chapter 6. The final chapter outlines the future implications of the study findings and concludes with recommendations for a model genetic association study.

Chapter 2

Genetics and Ischemic Stroke

There is strong evidence indicating that genetic factors play an important role in the aetiology of ischemic stroke. In the last decade, a number of single-gene disorders associated with ischemic stroke have been characterised (Table 2.1). These disorders are often considered in the differential diagnosis underlying ischemic stroke in young patients, particularly those free of established cerebrovascular risk factors, or those with a family history of early-onset stroke.

Table 2.1
Monogenic Disorders Causing Ischemic Stroke

Pathophysiology	Disorder	Inheritance
Prothrombotic states	Protein S deficiency	AD
	Antithrombin III deficiency	AD
	Protein C deficiency	AD
	Familial Sneddon Syndrome	AD
Haemaglobinopathies	Sickle cell disease	AR
Vasculopathies	Fibromuscular dysplasia	AD
	Familial moyamoya disease	AD
Connective tissue disorders	Neurofibromatosis type1	AD
	Ehlers-Danlos syndrome type 4	AD
	Marfan's syndrome	AD
Metabolic disorders	Fabry's disease	XR
	Homocyteinuria	AD
Small vessel disorders	CADASIL*	AD
Mitochondrial disorders	MELAS**	AD
Channelopathies	Familial hemiplegic migraine	AD

AD = autosomal dominant; AR = autosomal recessive; XR = X-linked recessive
** Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leucoencephalopathy*
*** Mitochondrial Encephalomyelopathy, Lactic Acidosis and Stroke-like episodes*
(Hademenos et al., 2001)

The mutations associated with these disorders are highly penetrant, present with ischemic stroke in early adulthood and follow classic Mendelian inheritance patterns (Hassan and Markus, 2000). On a population level, single gene disorders are rare and account for less than 1% of ischemic stroke. Most patients are affected by the sporadic form, which presents later in life and does not follow classic inheritance patterns. The aetiology in these cases is postulated to be multifactorial. Although there is often an association with well-established risk factors such as hypertension and smoking, there is strong evidence from epidemiological and animal studies implicating genetic factors.

2.1 The Attributable Risk of Known Risk Factors

Two large population-based epidemiological studies performed in Rochester (Davis *et al.*, 1987) and Perth (Jamrozik *et al.*, 1999), have shown that a significant proportion of ischemic stroke patients cannot be accounted for by established risk factors. The Rochester, Minnesota study was a population-based, nested case-control study that used a highly sophisticated medical record linkage system to identify 1,444 incidence cases of ischemic stroke between 1960 and 1984 in Rochester, United States of America. Registered ischemic stroke cases were compared with a similar number of stroke-free individuals from the same geographical region, matched for age, gender, date of stroke and duration of medical record at the time of stroke diagnosis. Using a conditional logistic-regression model that adjusted for confounders, the study was one of the first to provide a detailed profile of modifiable risk factors that could be applied on an individual basis to determine the probability of first ischemic stroke (Table 2.2) (Whisnant *et al.*, 1996).

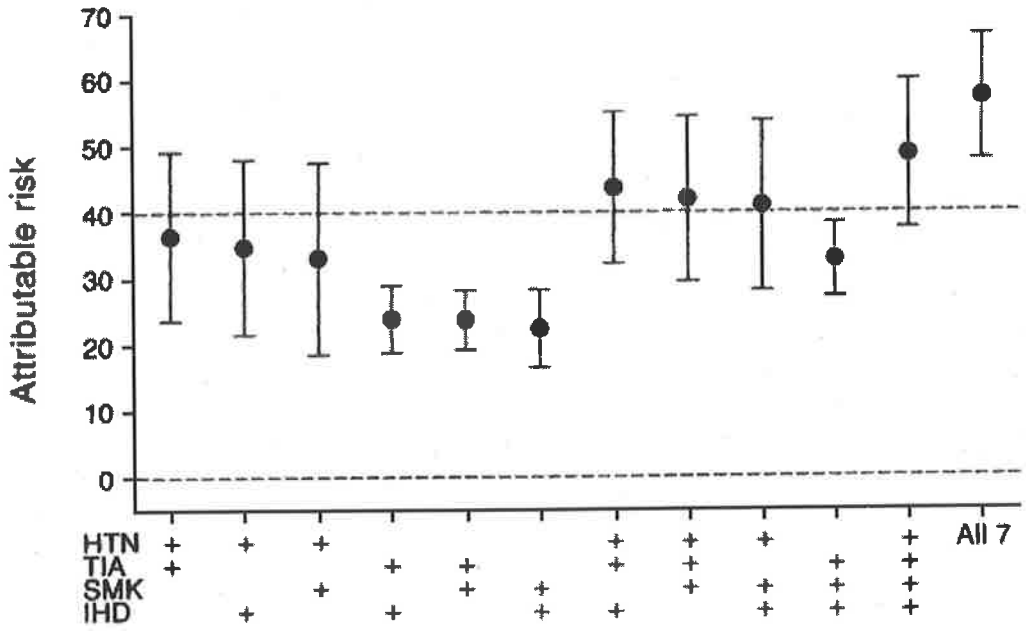
The population attributable risk is defined as the proportion of ischemic stroke within a population that can be attributed to a risk factor. It is a function of the intrinsic relative risk of a risk factor and its prevalence, and is a measure of the total risk-burden placed on a population. When applied to the Rochester data, the population attributable risk of all seven major risk factors was 57% (Whisnant, 1997). Similar findings were reported in the PCSS, with up to 68% of the population attributable risk of stroke being unaccounted for (Jamrozik *et al.*, 1994).

Table 2.2
Odds Ratios, Confidence Intervals and Probability Values for Risk Factors for First Ischemic Stroke: The Rochester, Minnesota Study

Risk Factor	Odds Ratio	95% CI	P value
Transient Ischemic Attack	5.6	3.7 - 8.5	.0001
Atrial Fibrillation	2.0	1.5 - 2.8	.0001
Hypertension	2.0	1.6 - 2.5	.0001
Left Ventricular Hypertrophy	1.8	1.4 - 2.4	.0001
Diabetes Mellitus	2.0	1.5 - 2.8	.0001
Cigarette Smoking	2.0	1.5 - 2.7	.0001
Congestive Cardiac Failure	2.1	1.5 - 3.0	.0001
Myocardial Infarction	1.8	1.3 - 2.5	.0002
Angina Pectoris	2.0	1.5 - 2.7	.0001
Mitral Valve Disease	2.4	1.5 - 4.1	.0008
Cardiac Wall Abnormality	2.7	1.3 - 5.7	.0086

(Whisnant et al., 1996)

Figure 2.1
Attributable Risks of a Combination of Risk Factors for First Ischemic Stroke: The Rochester, Minnesota Study



HTN = hypertension; TIA = transient ischemic attack; SMK = current cigarette smoking; IHD = ischemic heart disease; All 7 = HTN + TIA+ SMK+ IHD + atrial fibrillation + diabetes mellitus + mitral valve disease. (Whisnant, 1997)

The findings from both studies suggest that established risk factors account for between 32 to 57 percent of ischemic stroke and suggest that other undetermined factors are implicated in a substantial proportion.

2.2 Evidence of a Genetic Predisposition to Stroke

2.2.1 Twin Studies

Twin studies provide good evidence that genetic factors play an important role in the aetiology of ischemic stroke. These studies estimate the heritability of a trait by comparing the degree of concordance of disease expression (i.e. both twin siblings affected) between monozygotic and dizygotic pairs, with the former being greater if a genetic predisposition occurs. Using this study design, Brass *et al.* (1992) determined the concordance of stroke between 9,475 male twins born in the United States between 1917 and 1927. Through questionnaire response, the proband concordance rates were 17.7% for monozygotic twins and 3.6% for dizygotic twins, giving a relative risk of 4.94 ($p<0.05$) (Brass *et al.*, 1992). The study was unable to estimate the magnitude of heritability, as the absolute number of concordant twin pairs was small (monozygotic pairs = 1, dizygotic pairs = 7). Furthermore, as stroke was identified by history alone, no conclusions could be made on the genetic predisposition of different stroke subtypes.

In another study, a twin-based approach was used to investigate the relevance of genetic factors on a marker of sub-clinical ischemic cerebrovascular disease, i.e. white matter hyper-intensity on brain magnetic resonance imaging (Carmelli *et al.*, 1998). This study involved 74 monozygotic and 74 dizygotic US male twins aged 69 to 71 years. The concordance rates for white matter hyperintensity were 61 percent for monozygotic pairs and 38 percent for dizygotic pairs. As the prevalence of stroke in the entire sample was 15 percent, the risk of stroke in a monozygotic co-twin of an affected twin was 4-fold, whereas the risk for a dizygotic co-twin of an affected twin was 2.5-fold. Furthermore, the study found that genetic factors accounted for 71 percent of the variation in white matter hyper-intensity volume.

In the largest study published to date, the concordance of stroke amongst twins has been determined (Bak *et al.*, 2002). The study used a population-based Danish twin register comprising of 11,564 same-sex twin pairs. Of these, death from stroke was recorded in one of the co-twins in 351 monozygotic and 639 dizygotic pairs. The surviving co-twin was then followed for subsequent stroke. In comparison to dizygotic twins, a 2-fold increase in concordance for stroke death was found among monozygotic twins (relative risk 2.06, 95% CI 1.27-2.33). An important aspect of this study however, was its large sample size that allowed for an estimation of heritability. The study concluded that inherited genetic factors accounted for 32 percent of the total liability of stroke death between twins.

2.2.2 Family-based Studies

Until recently, the hypothesis of an aggregation of ischemic stroke among first-degree relatives remained controversial, with early studies reporting conflicting findings. Eight studies reported an overall doubling of stroke risk associated with a history of stroke amongst a first-degree relative (Kiely *et al.*, 1993, Jousilahti *et al.*, 1997, Welin *et al.*, 1987, Liao *et al.*, 1997, Diaz *et al.*, 1986, Graffagnino *et al.*, 1994, Khaw and Barrett-Connor, 1986, Carrieri *et al.*, 1994), whilst others have shown no statistical association (Herman *et al.*, 1983, Brass and Shaker, 1991, Boysen *et al.*, 1988). This contradiction may be explained by differences in study design, ascertainment of information and sample size. Most studies also failed to validate ischemic stroke in either the proband or the first-degree relative and relied solely on historical information.

The importance of the clinical validation of stroke in the design of family-based studies is highlighted in a study of the Framingham cohort. The Framingham study is an ongoing population-based, prospective cohort study that commenced in 1948. This study recruited a cohort of 5,209 persons (2,336 men and 2,873 women) who were free of cardiovascular and cerebrovascular disease and resided in Framingham, Massachusetts, in a four-year period between 1948 and 1952. The morbidity and mortality of this cohort have been closely followed by means of twice yearly interviews and physical examinations. The Framingham study was instrumental in

determining many of the now well-established risk factors for cerebrovascular disease including increasing age, male gender, hypertension, cardiovascular disease, diabetes, cigarette smoking and atrial fibrillation (D'Agostino *et al.*, 1994).

In 1971, the Framingham investigators commenced evaluation of the Framingham offspring cohort, which consisted of 5,124 descendents or spouses of descendents from the original Framingham cohort. Information from both cohorts was then used to determine the familial aggregation of stroke amongst first-degree relatives (Kiely *et al.*, 1994). The study found no significant association between stroke among members of the original Framingham cohort and a reported parental history of death from stroke (RR = 0.99, 95% CI 0.82-1.19). Among the offspring cohort, where validated clinical information was available for both proband and first-degree relatives, the relative risk of ischemic stroke for paternal stroke was 2.4 (95% CI 0.96-6.03) and 1.4 (CI 0.6-3.25) for maternal stroke. Although there is a trend supporting a familial aggregation of stroke, there remains some doubt as to the significance of these findings, as the confidence intervals are wide and the number of incidence stroke cases in the young offspring cohort is small (n=33). The small sample size also precluded estimates of heritability. The authors hope to strengthen their findings with a follow-up study in an older offspring cohort (and therefore a greater number of incident stroke cases) in the near future.

The Family Heart Study, published in 1997, provides the strongest evidence supporting a familial aggregation of stroke (Liao *et al.*, 1997). It is the largest family-based study of stroke to date, consisting of 3,168 probands and 29,325 first-degree relatives pooled from three ongoing parent cohort studies: the Framingham Offspring Study, the Utah Family Tree Study and the Atherosclerosis Risk in Communities Study. The study found a statistically significant association between stroke risk and history of stroke in any first-degree relative (OR 1.52, 95% CI 1.02-2.38). This association was even stronger for a paternal history of stroke (OR 2.01, 95% CI 1.10-3.69), and was independent of other established cerebrovascular risk factors (Table 2.3).

Table 2.3
Odds Ratios and Confidence Intervals for Proband Stroke by Familial History: The Family Heart Study

Risk Factor	Odds Ratio*	95% CI
No family history for stroke	1.00	-
History of stroke in any first-degree relative	1.56	1.02 - 2.38
History of stroke in one parent	1.85	1.18 - 2.90
History of stroke in father	2.01	1.10 - 3.69
History of stroke in mother	1.35	0.73 - 2.48
History of stroke in a sibling	0.93	0.44 - 1.97
History of stroke in a child	1.05	0.13 - 8.84
History of stroke in a spouse	2.13	0.67 - 6.76

** Adjusted for proband age, ethnicity, gender, history of elevated cholesterol, cigarette smoking status, history of coronary heart disease, hypertension, diabetes mellitus and sampling strata (Liao et al., 1997)*

2.2.3 Animal Studies

Studies of the stroke-prone spontaneously hypertensive rat (SP-SHR) have reinforced human epidemiological findings supporting a polygenic predisposition to ischemic stroke. This inbred rat line was established in 1974 following an observation of premature stroke in a colony of inbred spontaneously hypertensive rats (SHR) when given a diet high in sodium and low in protein and potassium (Okamoto *et al.*, 1974). Strokes were ischaemic and haemorrhagic, with histopathology showing similarities with vascular fibrinoid necrosis and lipohyalinosis seen in human small vessel disease. As the stroke phenotype was highly penetrant, susceptibility was thought to have arisen from a segregation of genetic factors during early inbreeding.

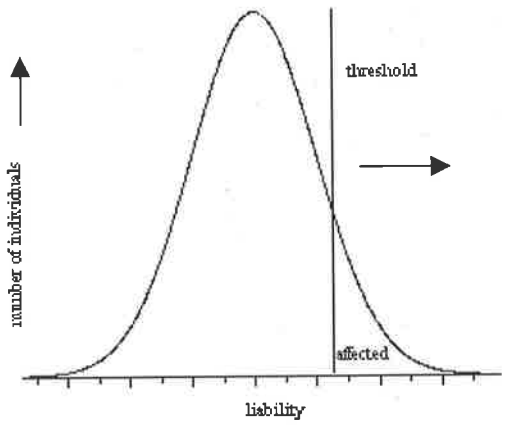
In an ingenious study, the genetic loci responsible for the SP-SHR phenotype were identified (Rubattu *et al.*, 1996). This was achieved by performing a genome-wide screen of polymorphic markers in a hybrid line of mice established from mating SP-SHR's and SHR's (F2 hybrid). The investigators hypothesized that the stroke susceptibility genes would cosegregate in the F2 hybrids, whilst alleles that contribute to hypertension would be equally shared, thus eliminating it as a potential confounder. The study showed a significant variation of stroke susceptibility in the F2 hybrids, whilst hypertension was uniformly maintained. The position of qualitative trait loci,

that is the identification of polymorphic alleles that co-segregated with the stroke phenotype, was determined. Three loci in the SP-SHR were associated with variation in stroke susceptibility. The STR1 locus on chromosome 1 was strongly associated with an increased susceptibility to ischemic stroke, accounting for 17.3 percent of the overall variance in the risk of stroke. In contrast, two other loci, the STR2 and the STR3 loci on chromosomes 4 and 5 respectively, were found to have a protective effect. The STR2 locus accounted for 9.6 percent of the variance of stroke risk and mapped close to the gene coding for atrial natriuretic factor, a peptide that regulates endothelial-mediated vasodilatation. The protective effect of STR2 was further enhanced in animals homozygous for the stroke-causing STR1 allele. The findings provide good evidence of the polygenic predisposition of ischemic stroke, with complex interactions between genes and ultimate phenotypic expression determined by the net accumulative effect of a number of genes with pathogenic properties.

2.3 The Genetic Heterogeneity of Ischemic Stroke

Both epidemiological and animal studies support a polygenic predisposition to ischemic stroke. Genetic factors may influence stroke risk in several ways. Firstly, genes may be disease causing, with a direct pathogenic effect that is independent of other cerebrovascular risk factors or environmental influences. As the pathogenesis of ischemic stroke varies between subtypes, it is likely that the genes responsible also differ according to subtype. The genes responsible for macrovascular atherosclerosis are likely to be different to those accounting for small vessel lipohyalinosis causing lacunar infarction. Furthermore, within each subtype, different genes are likely to be involved in different aspects of the underlying pathology. For example, genetic factors that predispose to atherosclerosis are likely to differ from genes responsible for plaque disruption and genes promoting intravascular thrombosis and occlusion. Genetic factors may also influence stroke risk by indirect means, such as predisposing to well-established risk factors such as hypertension, diabetes and hypercholesterolemia. Alternatively, genetic factors may influence the susceptibility to known cerebrovascular risk factors, with disease expression caused by gene-risk factor interaction. The latter may explain the phenotypic heterogeneity in those with similar risk factor profiles.

Figure 2.2
Threshold Model for the Inheritance of Common Ischemic Stroke



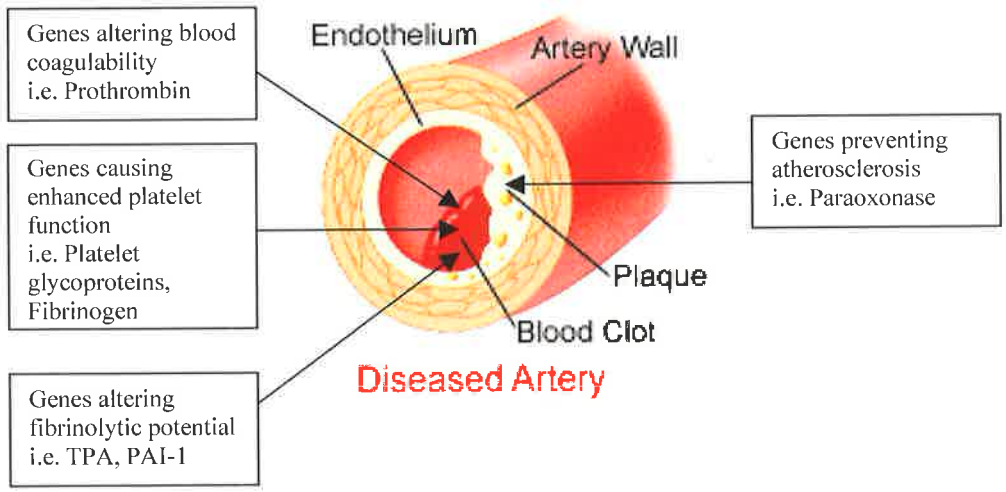
There is a liability towards ischemic stroke that consists of a combination of genetic and non-genetic factors, which is normally distributed in the population. Ischemic stroke is expressed only in individuals whose liability exceeds a threshold (Lander and Schork, 1994).

The 'threshold model of multifactorial inheritance' best describes the phenotypic expression of polygenic ischemic stroke (Lander and Schork, 1994). This model assumes a normal distribution of liability towards expression of a stroke phenotype that is composed of the sum of environmental and genetic influences (Figure 2.2). Disease expression, that is ischemic stroke, occurs when there is an accumulation of liability beyond a disease threshold. The model illustrates how individual genes may be insufficient to cause disease alone, although they remain important contributors to the total liability load. When considering this model, the genetic and environmental influences that are protective need to be taken into account, as well as any synergistic effects that occur in gene-environment or gene-gene interactions. The latter may also be disease causing or protective with the total liability predisposing to ischemic stroke being a complex function of both protective and disease causing influences and their interactions. This model implies that ischemic stroke occurs in those who are genetically predisposed by inheritance of one or more genetic factors, with the risk augmented by other established risk factors.

2.4 Candidate Genes for Ischemic Stroke

The candidate gene approach is the current mainstay of investigating the genetic predisposition of ischemic stroke. The method relies on determining an association between ischemic stroke and known genetic SNP's of functional significance. A prospective case-control study design is commonly used to compare the prevalence of a SNP between ischemic stroke cases and a normal control group. Candidate SNP's are chosen based on their influence on genes underlying the pathological process. In ischemic stroke, genes influencing atherosclerosis, platelet function, coagulation and fibrinolysis are plausible candidates for investigation based on known pathogenic models (Figure 2.3). Although biological plausibility is essential in the selection of candidate genes, other methodological issues such as sample size estimation and ethnic allelic variation also require consideration prior to candidate gene selection.

Figure 2.3
Genetic Candidates for Common Ischemic Stroke



Candidate genes are selected based on their known physiologic effect. For common ischaemic stroke, relevant pathogenic mechanisms include atherosclerosis, hypercoagulability, enhanced platelet function and impaired fibrinolysis.

In this study, seven genes were selected for investigation: a) atherosclerosis - paraoxonase; b) platelet function – glycoprotein 1b, IIIa and fibrinogen; c) coagulation – prothrombin; d) fibrinolysis – tissue plasminogen activator and plasminogen

activator inhibitor. The rationale for the choice of study design and selection of SNP's is described in detail in Chapter 3.3.

2.4.1 Paraoxonase (PON1)

Oxidation of low-density lipoprotein (LDL) plays an important role in the pathogenesis of atherosclerosis. Oxidized LDL, when trapped in the sub-endothelial space, induces the release of monocyte-specific chemoattractants, adhesion molecules and colony stimulating factors from vascular endothelium (Watson *et al.*, 1995). Further oxidation leads to arterial wall cytotoxicity and uptake of oxidized LDL particles by macrophages. This sequence of events ultimately causes atherosclerotic plaque formation. The oxidation of LDL is limited by high-density lipoprotein (HDL). This effect is mediated by PON1, an enzyme that is produced by the liver and circulates in the blood stream exclusively bound to HDL. It exerts an anti-atherogenic effect by hydrolysing sub-endothelial lipid peroxides before they accumulate in LDL particles (Watson *et al.*, 1995). In PON 1 knockout mice, HDL loses its protective effect on LDL oxidation, and mice become susceptible to diet-induced atherosclerosis (Shih *et al.*, 1998). This finding suggests that the anti-atherogenic effect of HDL is primarily mediated by PON1.

PON1 expression is largely determined by genetic factors. The PON1 gene has two SNP's, resulting in an arginine to glutamine substitution at amino acid position 191 and a methionine (M) to leucine (L) substitution at amino acid position 54 respectively (Mackness *et al.*, 2002). In addition to these, five SNP's have been identified recently in the PON1 regulatory region: -107 thymidine (T) / cytosine (C), -126 guanine (G)/C, -162 adenine (A)/G, -907 G/C and -824 G/A (Mackness *et al.*, 2002). In 2000, Leviev *et al.* showed that the T allele at position –107 was associated with significantly lower PON1 levels and enzymatic activity (Leviev and James, 2000). The study concluded that the PON1 –107C/T polymorphism was the main genetic contributor to PON1 expression, accounting for 24.7 percent of the variance in serum PON1 levels. This was followed by the PON1 M54L SNP that accounted for 4.4 percent of the variation in serum PON1. Furthermore, the –107T/C SNP is considered to have functional importance as it occurs within a binding site for an Sp1 transcription factor (Leviev

and James, 2000). The other SNP's did not have a significant functional impact on PON1 expression or activity. A similar finding was reported in another study, which showed that the -107 T/C and M54L SNP's accounted for 22.8 and 5 percent of the variance of PON1 activity respectively (Brophy *et al.*, 2001). Functional significance has also been shown for the M54L SNP, with the MM genotype providing the greatest protection from LDL oxidation (Mackness *et al.*, 1999). These findings suggest that both the PON1 54L and –107T alleles may predispose to atherosclerosis.

Few studies have examined the clinical relevance of the PON1 -107 T/C SNP. In a study of type 2 diabetics, a positive association between the PON1 –107TT genotype and coronary heart disease was found (OR 1.64, 95% CI 1.03-2.61). This was independent of other known cardiovascular risk factors (James *et al.*, 2000). A subsequent study of patients with angiographically proven coronary artery disease showed that this association was limited to those less than 60 years of age (Leviev *et al.*, 2001). To date, no studies that have addressed an association between the PON1 –107C/T SNP and ischemic stroke have been reported.

Studies of the PON1 M54L SNP have generated conflicting results. In a post-mortem analysis of the abdominal aorta, mesenteric and iliac arteries in 123 consecutive autopsies, the LL genotype was associated with a greater number and significantly thicker atherosclerotic lesions when compared to M allele carriers (Malin *et al.*, 2001). Homozygosity for the L allele has also been associated with severe internal carotid artery atherosclerosis (Fortunato *et al.*, 2003, Leus *et al.*, 2000, Schmidt *et al.*, 1998a). The largest study involved 316 randomly selected individuals, and showed that the LL genotype was associated with nearly double the risk of severe internal carotid artery atherosclerosis (adjusted OR 1.91, 95% CI 1.14 - 3.20) (Schmidt *et al.*, 1998a). In 2000, the same authors reported a strong association between the LL genotype and progression of white matter lesions on magnetic resonance imaging (adjusted OR 2.7, 95% CI 1.4 – 5.2), implicating the relevance of this SNP in the pathogenesis of small vessel disease. In contrast, two other studies of ischemic stroke did not find an association with the PON1 M54L SNP (Voetsch *et al.*, 2002, Imai *et al.*, 2000).

2.4.2 Glycoprotein 1b

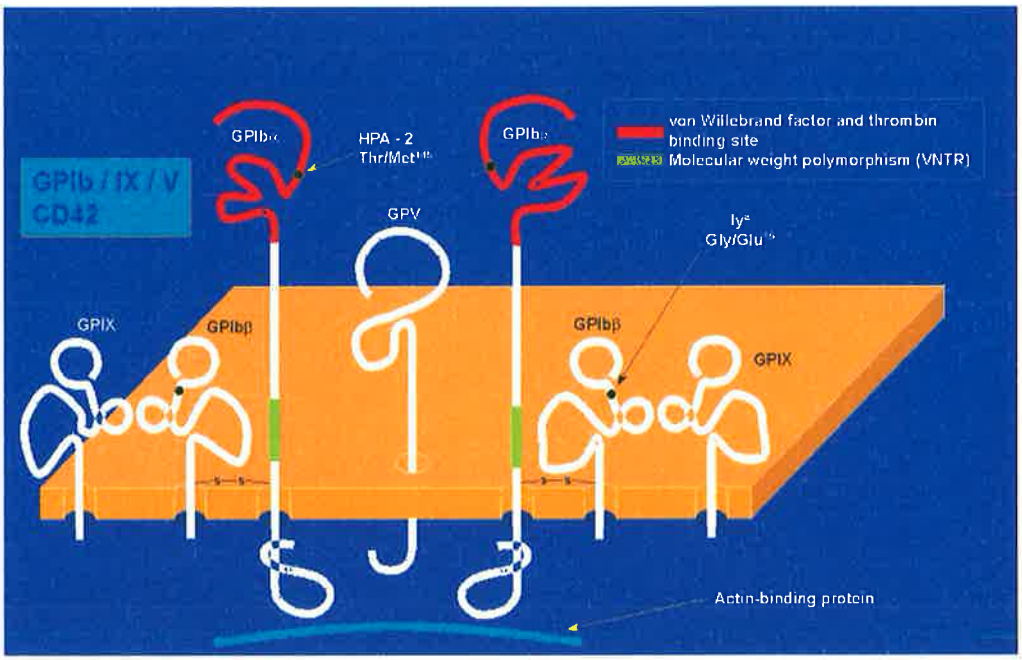
The benefit of anti-platelet therapy in preventing recurrent ischemic stroke is well established and provides strong evidence that platelets play an important role in the pathogenesis of ischemic stroke (Gubitz *et al.*, 2003). Evidence now suggests that that an exaggerated platelet response predisposes to ischemic stroke. Plasma levels of PF4 and Beta-TG, two platelet derived molecules released following cellular activation, are significantly higher in patients with ischemic stroke when compared to normal controls (Shah *et al.*, 1985, Uchiyama *et al.*, 1983). Furthermore, functional *in vitro* studies have shown that platelets from patients with ischemic stroke are excessively prone to thrombus formation, either spontaneously or following exposure to various aggregating agents or tensile stress (Konstantopoulos *et al.*, 1995, Shah *et al.*, 1985, Uchiyama *et al.*, 1983).

Haemostasis functions to limit blood loss following vascular injury. The process begins with adhesion of platelets to exposed sub-endothelial tissues. This initial step is mediated by a plasma adhesive glycoprotein, von Willebrand factor (vWF), which upon binding to exposed sub-endothelium, undergoes a conformational change revealing a cryptic ligand that is subsequently recognised by its platelet receptor, the glycoprotein 1b/IX/V complex (Berndt *et al.*, 1995). This interaction acts as a "platelet brake", slowing down circulating platelets as they move across exposed sub-endothelial tissues and permits platelet arrest via three other interactions: a) platelet integrin $\alpha 2\beta 1$ binding to sub-endothelial collagen; b) platelet integrin $\alpha 5\beta 1$ binding to sub-endothelial fibronectin; and c) platelet integrin $\alpha 2b\beta 3$ (glycoprotein IIb/IIIa) binding to sub-endothelial fibrinogen, fibrin and vWF (Savage *et al.*, 1996). Spontaneous activation of platelets can also occur under conditions of high shear stress such as those encountered over an ulcerated atherosclerotic plaque or across a high-grade vascular stenosis. The process can occur without exposure to sub-endothelial tissues and is initiated by the binding of circulating vWF to platelet glycoprotein $1b\alpha$ (Razdan *et al.*, 1994). In comparison to normal controls, platelets from ischemic stroke patients exhibit an exaggerated shear-induced platelet aggregation response. This process is not inhibited by aspirin but is totally inhibited by ticlopidine and

dipyridamole, suggesting that intra-platelet signalling does not involve the cyclo-oxygenase pathway (Uchiyama *et al.*, 1994, Konstantopoulos *et al.*, 1995).

In 1992, Kuijpers *et al.* identified a cytosine to thymidine single nucleotide SNP located at position 434 within the coding region of the glycoprotein 1b gene resulting in a Threonine (Thr) to Methionine (M) substitution at amino acid position 145 (Figure 2.4). This substitution causes a conformational change in the glycoprotein 1b molecule that alters its antigenicity and defines the human platelet alloantigen (HPA-2a/b) system that is associated with platelet transfusion refractoriness (Kuijpers *et al.*, 1992).

Figure 2.4
Structure of the Glycoprotein Ib/IX/V Complex



Schematic drawing of the Glycoprotein Ib/IX/V complex showing the position of known SNP's relative to the vWF-binding site. Note the HPA-2a/b is located within the vWF-binding domain. Figure from http://www.nibsc.ac.uk/divisions/Haem/diag_glcoprotein_hpa_antigens.html.

A cytosine allele at nucleotide position 434 corresponds to threonine at amino acid position 145 and the HPA-2a antigen subtype, whilst the thymidine allele results in a methionine substitution at amino acid position 145 and HPA2-b antigen subtype. The change in molecular shape is also thought to be functionally important as the polymorphism occurs within the leucine-rich region responsible for binding of von

Willebrand factor ligand (Kuijpers *et al.*, 1992). In a study by Douglas *et al.*, carriers of the methionine allele had increased platelet activity, evidenced by an exaggerated response to agonist-induced platelet aggregation under high shear conditions (Douglas *et al.*, 2000).

To date, the clinical studies investigating the relevance of the glycoprotein 1b 145Thr/M SNP to ischemic stroke have reported conflicting findings. Three studies have reported a positive association (Sonoda *et al.*, 2000, Reiner *et al.*, 2000, Gonzalez-Conejero *et al.*, 1998), whilst three others were negative (Meiklejohn *et al.*, 2001, Carlsson *et al.*, 1997, Baker *et al.*, 2001). In the largest study involving 200 ischemic stroke patients and 317 age and gender-matched controls, possession of the Met allele (Met/Met or Met/Thr genotype) was associated with a doubling of risk of ischemic stroke (OR 2.18, p value = 0.0005) (Sonoda *et al.*, 2000). Furthermore, in a subgroup of non-smoking women less than 60 years of age, the risk of ischemic stroke associated with the Met allele was greater than 10-fold (OR 10.6, 95% CI 2.2- 51.7). Reiner *et al.* reported a similar risk associated with homozygosity for the Met allele, although the sample size was small, giving rise to wide confidence intervals (OR 10.63, 95% CI 1.43 – 79.34) (Reiner *et al.*, 2000). In one of the negative studies, data analysis was simplistic and did not account for confounding variables such as smoking, hypertension and atrial fibrillation (Carlsson *et al.*, 1997). In another negative study involving 219 patients and 205 controls, there were no patients who were homozygous for the Met allele thus risk determination could not be performed. Although the prevalence of the heterozygous state was higher in ischemic stroke patients, the study lacked the statistical power to demonstrate a significant risk associated with this genotype (OR 1.8, 95% CI 0.94-3.4, p = .07) (Baker *et al.*, 2001).

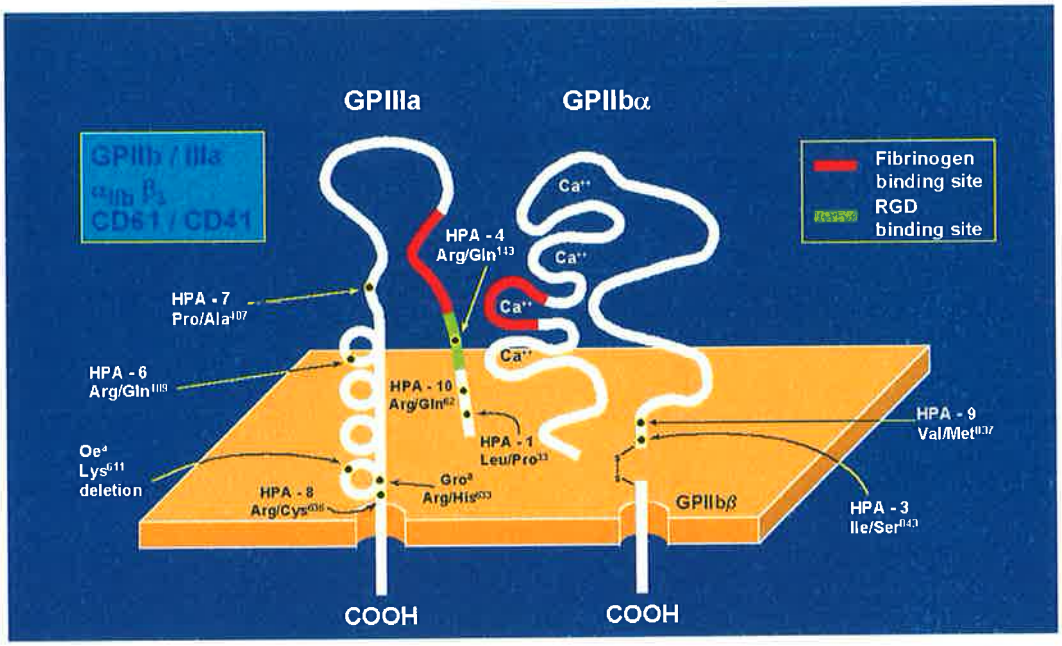
2.4.3 Glycoprotein IIb/IIIa

Glycoprotein IIb/IIIa exists in inactive state on the surface of resting platelets. After platelet exposure to sub-endothelial tissues or under high shear-stress conditions, von Willebrand factor bound to platelet glycoprotein 1b/IX/V complex activates intra-cellular mechanisms causing a conformational change within the Glycoprotein IIb/IIIa molecule. This reveals receptor sites for von Willebrand factor and fibrinogen, which

mediate cross-linking between platelets and stabilize the adhesion to exposed sub-endothelial tissues. Activated platelets also release a variety of substances that attract and activate other circulating platelets, culminating in the formation of a platelet plug (del Zoppo, 1998).

In 1989, Newman *et al.* identified a thymidine to cytosine SNP at position 1565 within exon 2 of the glycoprotein IIIa gene (Figure 2.5).

Figure 2.5
Structure of the Glycoprotein IIb/IIIa Complex



Schematic drawing of the Glycoprotein IIb/IIIa complex showing the position of known SNP's relative to the fibrinogen-binding site. Note the PIA1/A2 SNP (HPA-1) is located outside the fibrinogen-binding domain. Figure from http://www.nibsc.ac.uk/divisions/Haem/diag_glcoprotein_hpa_antigens.html

The SNP resulted in a leucine to proline substitution at amino-acid position 33 from the NH_2-terminus, and defined the PlA1 (leucine) - PlA2 (proline) platelet alloantigen system associated with neonatal alloimmune thrombocytopenia and post-transfusion purpura (Newman *et al.*, 1989). This SNP has been shown recently to be functionally significant with the degree of platelet adhesion being strongly genotype dependent. Under physiological shear-stress conditions, platelets of PlA2/A2 genotype were

associated with a 2.8-fold increase in binding to immobilized fibrinogen when compared to their wild-type counterpart (Vijayan *et al.*, 2003).

The mechanism underlying this enhanced adhesive phenotype is not well understood, as the leucine to proline substitution occurs adjacent to the fibrinogen-binding site. In one study, enhanced tyrosine phosphorylation of intracellular kinases following receptor-fibrinogen interaction in PlA2/A2 platelets resulted in a conformational change that altered intracellular responses following receptor-ligand binding that was different to the wild-type counterpart (Vijayan *et al.*, 2000).

The relevance of the platelet glycoprotein IIb/IIIa PlA1/A2 SNP to ischemic stroke remains controversial. Four studies have found a positive association (Carter *et al.*, 1998, Pongracz *et al.*, 2001, Streifler *et al.*, 2001, Wagner *et al.*, 1998). In the largest of these, involving 505 ischemic stroke patients and 402 healthy controls, a 2.4-fold increase in risk of stroke associated with heterozygosity for the PlA2 allele (OR 2.37, 95%CI 1.19 – 4.47, p value 0.01) (Carter *et al.*, 1998). Furthermore, in subgroup analysis of subjects less than 50 years of age, possession of the PlA2 allele was shown to be a significant predictor of first ischemic stroke (OR 1.68, 95%CI 1.01-2.82, p value = 0.05) (Carter *et al.*, 1998). The glycoprotein IIb/IIIa PlA2 allele was also shown recently to be an important factor determining the natural history of internal carotid artery stenosis. In this study comparing 86 symptomatic and 67 asymptomatic patients with a greater than 50% internal carotid artery stenosis, heterozygosity for the PlA2 allele conferred a 3.4-fold increased risk of ischemic stroke (OR 3.4, 95%CI 1.5-7.8). The association was independent of other known cerebrovascular risk factors (Streifler *et al.*, 2001).

In direct contrast, five studies have failed to show an increased risk of ischemic stroke associated with the platelet glycoprotein IIb/IIIa PlA1/A2 SNP (Carlsson *et al.*, 1997, Reiner *et al.*, 2000, van Goor *et al.*, 2002, Kekomaki *et al.*, 1999, Ridker *et al.*, 1997). The largest of these involved incident stroke cases from the prospective follow-up of 14,916 men participating in the Physicians' Health Study (Ridker *et al.*, 1997). Over an 8.6-year follow-up period, 209 incident strokes occurred. They were compared with

709 controls from the same cohort who were free of vascular disease. The study found no association between the glycoprotein IIb/IIIa PlA2 allele and stroke (OR 0.93, 95%CI 0.7-1.3) (Ridker *et al.*, 1997). The study has since been criticised as it failed to differentiate between ischemic and haemorrhagic stroke (Smith *et al.*, 1997). The study cohort, consisting entirely of men, was also unusually healthy with vascular event rates less than three per thousand per year and not representative of the general population at large (Bray *et al.*, 1997). There was also significant racial heterogeneity amongst participants, with approximately 6 percent being non-white. Furthermore, outcome events were based on a self-reported questionnaire, thus introducing misclassification bias.

In another negative study, a trend favouring an association between the glycoprotein IIb/IIIa PlA2 allele and was observed, but there was no adjustment for the confounding effects of established cerebrovascular risk factors (Carlsson *et al.*, 1997). Two other negative studies investigated only a small number of patients (45 and 36 ischemic stroke patients, respectively) and as such, were significantly underpowered to establish a positive association (van Goor *et al.*, 2002, Reiner *et al.*, 2000).

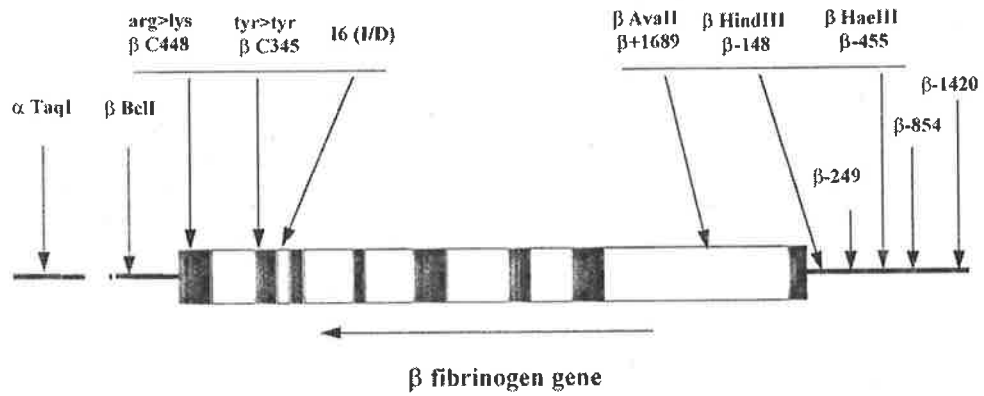
2.4.4 Fibrinogen

Hyperfibrinogenemia is an established risk factor for ischemic stroke (Kofoed *et al.*, 2003, Qizilbash *et al.*, 1991, Resch *et al.*, 1992, Wu *et al.*, 1992). In a recent population-based prospective cohort study involving 8,755 participants, a high plasma fibrinogen level was associated with almost double the risk of first ischemic stroke (OR 1.9, 95% CI 1.4-2.5) (Kofoed *et al.*, 2003). The risk was highest for men and for persons under 50 years of age (OR 5.2, 95%CI 1.1-2.6) (Kofoed *et al.*, 2003). There are several pathophysiologic mechanisms linking fibrinogen to ischemic stroke. Fibrinogen binds to platelets via glycoprotein IIb/IIIa receptors and promotes platelet aggregation. Furthermore, fibrinogen is the precursor of fibrin, a polypeptide that is incorporated into the mural thrombus and influences clot stability. Raised fibrinogen levels are associated with an accelerated platelet aggregation response and the formation of larger clots that are less amenable to fibrinolytic reversal (Koenig, 2003). Fibrinogen has also been implicated in the development of atherosclerosis. It has been

shown to bind to the endothelial cell receptors and causes a change in endothelial cell permeability allowing for deposition of sub-endothelial deposits. Fibrinogen can also directly stimulate smooth muscle cell proliferation, induce monocyte chemotaxis and provide an adsorptive surface in the sub-endothelium allowing for the accumulation of oxidized lipoproteins (Koenig, 2003).

Fibrinogen is a glycoprotein composed of pairs of three non-identical polypeptides: the alpha (α); beta (β) and gamma (γ) chains. The genes encoding for these components are clustered on the long arm of chromosome 4. As the production of fibrinogen is highly dependent on the synthesis of the β component (Roy *et al.*, 1990), most clinical research has concentrated on the β fibrinogen gene.

Figure 2.6
Structure of the β fibrinogen gene



Schematic diagram of the β fibrinogen gene and position of known SNP's. Thick black bars represent exons, wide white bars represent introns. Adjacent black lines represent flanking regions. Note the close proximity between the –148 C/T (β HindIII) and –455G/A SNP's (β HaeIII) (Behague et al., 1996).

Eleven SNP's have been identified within this gene, with most being in tight linkage disequilibrium with each other (Behague *et al.*, 1996) (Figure 2.6). Functional significance is limited to the –148 C/T SNP located in the promoter region of the β fibrinogen gene. This polymorphism is located immediately adjacent to a transcription-factor binding site that mediates the effect of interleukin-6 on fibrinogen levels (Behague *et al.*, 1996). The –148 C/T SNP is also in complete linkage disequilibrium

with the –455G/A polymorphism located in the 5' flanking region of the β fibrinogen gene (Behague *et al.*, 1996), with both the T and A alleles from the respective SNP's causing a significant increase in plasma fibrinogen level (Behague *et al.*, 1996, Liu *et al.*, 2002, Iso *et al.*, 1995, Scarabin *et al.*, 1993, Tybjaerg-Hansen *et al.*, 1997). As the –455G/A SNP is not a functional locus, its association with hyperfibrinogenemia is thought to arise through its complete linkage disequilibrium with the –148 C/T SNP.

Few studies have investigated the relevance of the β fibrinogen –148 C/T SNP to ischemic stroke. A strong association between T allele homozygotes and carotid atherosclerosis was reported in the Austrian Stroke Prevention Study (Schmidt *et al.*, 1998b). This was a prospective cross-sectional study that genotyped 399 asymptomatic individuals for the β fibrinogen –148 C/T SNP. Subjects were also examined for internal carotid artery atherosclerosis using colour-coded duplex scanning. The study found that homozygosity for the T allele was associated with a 6-fold increase in risk of internal carotid artery atherosclerosis (OR 6.17, 95%CI 1.7 - 22.4) (Schmidt *et al.*, 1998b). However, this result should be interpreted with caution, as the number of TT homozygotes was small (n=25) and an association between the –148 C/T SNP and plasma fibrinogen levels was not demonstrated. Studies of other SNP's in linkage disequilibrium with the –148 C/T SNP provide further evidence of its relevance to ischemic stroke. Four studies of the –455G/A SNP have shown a positive independent association with ischemic stroke (Nishiuma *et al.*, 1998, Martiskainen *et al.*, 2003, Kessler *et al.*, 1997, Liu *et al.*, 2002). The –455G/A SNP predisposes to stroke caused by large and small-vessel disease. In the largest study involving 227 patients and 225 controls, homozygosity for the A allele was found to be a significant predictor of ischemic stroke caused by large-vessel disease (chi-squared = 4.0, p = 0.045) (Kessler *et al.*, 1997). More recently, in a cross-sectional study of 299 ischemic stroke patients, possession of the A allele and hypertension was associated with a 4-fold increase in risk of multiple lacunar infarcts (OR 4.24, 95%CI 1.29-13.99, p = 0.02) (Martiskainen *et al.*, 2003).

In contrast to the above findings, Blake *et al.* did not find an association between the β fibrinogen –148 C/T SNP and an increased risk of ischemic stroke (OR 0.87, 95%CI

0.63-1.21) (Blake *et al.*, 2001). This study used data from the Physicians Health Study cohort and should be interpreted with caution for reasons discussed earlier (reference in 2.4.2).

2.4.5 Prothrombin

Prothrombin (or coagulation factor II) is produced by the liver and is the precursor of circulating thrombin. Upon activation of the coagulation pathway, prothrombin is converted to thrombin that has a central role in the conversion of fibrinogen to fibrin. High prothrombin levels are associated with a hypercoagulable state and predispose to venous thrombosis (Poort *et al.*, 1996). Recently, Gomez-Garcia *et al.* showed that prothrombin was also an important determinant of arterial thrombosis. In a study of 49 young ischemic stroke patients and 87 asymptomatic controls, plasma prothrombin levels were significantly higher in those with ischemic stroke (1.11 U/ml vs 0.97 U/ml, p value 0.003). This study also found that a high prothrombin level (greater than 1.1 U/ml) was associated with a 3-fold increase in the risk of ischemic stroke (RR 3.2, 95%CI 1.03-9.96) (Gomez-Garcia *et al.*, 2002).

In 1996, Poort *et al.* identified a G to A SNP at position 20210 in the 3' untranslated region of the prothrombin gene. The SNP was shown to be functionally significant with the A allele associated with high plasma prothrombin levels (Poort *et al.*, 1996). The prothrombin G20210A SNP is located in the last position of the 3'-untranslated region and is thought to influence prothrombin levels by preventing the cleavage of mRNA during protein translation. This results in higher stability of transcribed mRNA and contributes to a greater translation efficacy (Poort *et al.*, 1996).

The risk of ischemic stroke associated with the prothrombin G20210A SNP remains controversial. Homozygosity for the A allele is rare, with an estimated prevalence of 1.3 per 10,000. This genotype is strongly associated with premature death from both venous and arterial thrombotic disease, including ischemic stroke (Nguyen, 2000, Giordano *et al.*, 1999). Three studies have shown an association between the prothrombin 20210 G/A genotype and an increased risk of ischemic stroke (De Stefano *et al.*, 1998, Gomez-Garcia *et al.*, 2002, Longstreth *et al.*, 1998). De Stefano et al.

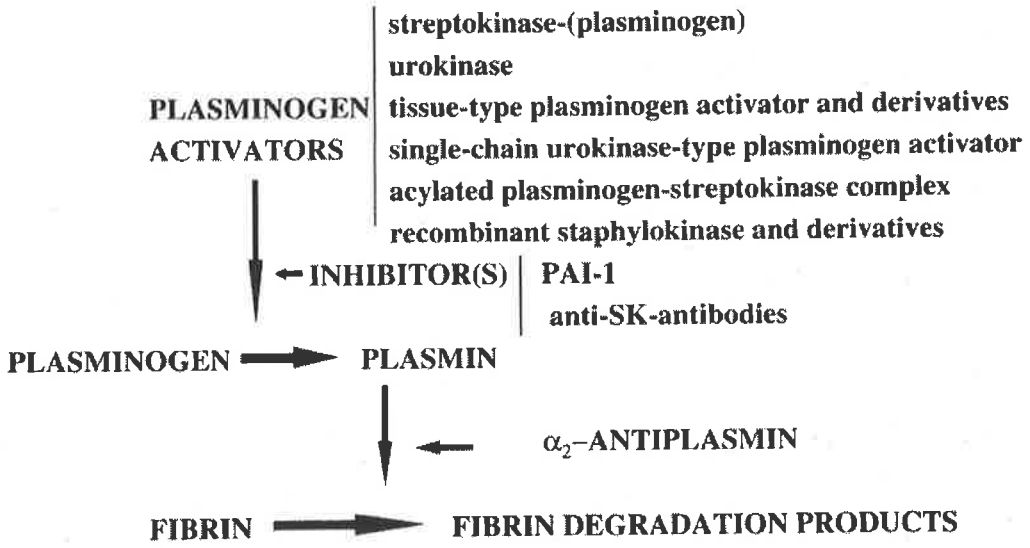
(1998) reported the strongest association, with a 3.8-fold increased risk of ischemic stroke (OR 3.8, 95%CI 1.1-13.1) (De Stefano *et al.*, 1998). The risk increased to 5-fold when the two ischemic stroke cases that were homozygous for the A allele were taken into account (Possession of A allele, OR 5.1, 95%CI 1.6-6.3). These findings, however, cannot be extrapolated to an older population with known cerebrovascular risk factors, as the study population consisted patients with cryptogenic ischemic stroke who were less than 50 years of age.

In contrast, seven studies showed no association between the prothrombin G20210A SNP and ischemic stroke (Hankey *et al.*, 2001, Corral *et al.*, 1997, Austin *et al.*, 2002, Ferraresi *et al.*, 1997, Ridker *et al.*, 1999, Szolnoki *et al.*, 2002, Lopaciuk *et al.*, 2001). In an Australian study involving 219 cases of first-ever ischemic stroke and 205 controls, heterozygosity for the A allele was associated with a trend toward an increased stroke risk, however this fell short of being statistically significant (OR 1.9, 95%CI 0.5-6.2) (Hankey *et al.*, 2001). In another large negative study, patients and controls were recruited from the US Physicians Health Study cohort and thus should be interpreted with caution for reasons described earlier (refer to 2.4.3) (Ridker *et al.*, 1999). The remaining five negative studies were limited to a young stroke population, thus the findings cannot be extrapolated beyond this group.

2.4.6 Tissue Plasminogen Activator (TPA)

Occlusive thrombus complicating an atherosclerotic plaque underlies both small and large vessel ischemic stroke (Fisher, 1979, Constantinides, 1967). Focal disruption of an atheromatous plaque exposes sub-endothelial collagen and triggers a local haemostatic response. The endogenous fibrinolytic defence system plays an important role in counteracting thrombus formation, with the degree of thrombosis governed by the balance of these opposing effects (Rosenberg and Aird, 1999) (Figure 2.7). Endothelium-derived TPA is the primary mediator of local intravascular fibrinolysis. In knockout mouse experiments, TPA-deficient mice were found to have markedly slower rates of reperfusion following thrombotic occlusion of a carotid artery, when compared with their wild-type counterparts (Matsuno *et al.*, 1999).

Figure 2.7
The Human Fibrinolytic Pathway

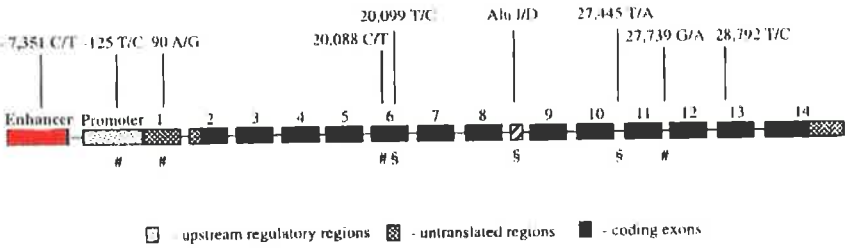


Schematic diagram showing the mediators of intravascular fibrinolysis. Upon exposure to tissue bound fibrin, TPA facilitates the activation of plasminogen to plasmin, which in turn cleaves fibrin monomers. The process is limited by circulating plasminogen activator inhibitor (PAI-1) and α2-antiplasmin, which inactivate TPA and plasmin respectively. Therefore factors that reduce TPA activity or enhance PAI-1 activity may impair fibrinolytic potential and predispose to thrombosis.

The secretion of TPA by endothelial cells is regulated by two mechanisms: (a) an acute release of local TPA that is precipitated by vascular injury and occurs within seconds to minutes and (b) a long-term basal secretion rate that is determined by the chronic activation of the fibrinolytic system and is responsible for the changes in the plasma level of TPA (Kooistra *et al.*, 1994). Indeed, a high plasma TPA antigen level has been shown to be an independent predictor of ischemic stroke in both cross-sectional (Johansson *et al.*, 2000, Kain *et al.*, 2001, Margaglione *et al.*, 1994, Lindgren *et al.*, 1996, Kristensen *et al.*, 1999, Kristensen *et al.*, 1998, Macko *et al.*, 1999) and prospective cohort studies (Ridker, 1994, Smith *et al.*, 1998). The rapid release of local TPA by the endothelium is considered to be the most important mechanism mediating the dissolution of arterial thrombus (Kooistra *et al.*, 1994). Hence a low capacity for rapid TPA release is likely to predispose to ischemic stroke.

In humans, genetic factors play an important role in the variance of endothelial TPA release (Jern *et al.*, 1999). Recently, a SNP located at position –7,351 within the enhancer region of the TPA gene was reported and shown to strongly correlate with endothelial TPA release rates (Ladenvall *et al.*, 2000). Furthermore, the SNP was found to have functional importance as it occurred within an Sp1 binding site, an important factor promoting transcription. Possession of a T allele was shown to inhibit Sp1 binding and was associated with less than half the TPA release observed in those homozygous for the C allele. The TPA –7,351C/T SNP has subsequently been shown to be clinically relevant, with a strong association of the T allele with first myocardial infarction (OR 2.68, 95%CI 1.3-5.5) (Ladenvall *et al.*, 2002). To date, no studies have reported the relevance of the TPA –7,351C/T SNP in ischemic stroke.

Figure 2.8
Structure of the TPA Gene



Schematic diagram showing the structure of the TPA gene and position of known SNP's. Note the –7,351 SNP lies within the enhancer region of the gene and influences the regulation of gene expression. (Ladenvall et al., 2000)

2.4.7 Plasminogen Activator Inhibitor – 1 (PAI-1)

PAI-1 is an important regulator of intravascular fibrinolysis. Upon release by vascular endothelial cells, PAI-1 binds and inactivates circulating TPA, which results in less plasmin degradation of fibrin. Thus an elevated PAI-1 level may result in excessive suppression of fibrinolysis and predispose to an increased tendency for arterial thrombosis. Indeed, high plasma PAI-1 levels have been associated with both first myocardial infarction and ischemic stroke (Brockmann *et al.*, 1991, Zhang *et al.*, 2001, Kain *et al.*, 2002, Eriksson *et al.*, 1995).

The regulation of PAI-1 expression is largely genetically determined. In a recent study of normotensive twins, genetic factors accounted for 87% of the variance of plasma PAI-1 (Cesari *et al.*, 1999). In 1993, Dawson *et al.* identified a single base pair insertion/deletion polymorphism (PAI-1 5G/4G) in the promoter region of the PAI-1 gene (Dawson *et al.*, 1993). This polymorphism was shown to be functionally significant with the deletion allele (4G) associated with a 6-fold increase in PAI-1 mRNA expression (Dawson *et al.*, 1993) and a 25% increase in plasma PAI-1 levels (Eriksson *et al.*, 1995). Both alleles are located within a transcription-factor binding site, with the insertion allele (5G) also binding a repressor protein that inhibits DNA transcription (Eriksson *et al.*, 1995).

Studies that have investigated the risk of ischemic stroke associated with the PAI-1 5G/4G polymorphism have produced conflicting results. Two studies have reported a significant association between the 4G allele and ischemic stroke (Zhang *et al.*, 2001, Bang *et al.*, 2001). The larger study, which involved 60 ischemic stroke patients and 100 asymptomatic controls, found a 3-fold increase in risk of ischemic stroke after adjustment for known cerebrovascular disease risk factors (OR 3.11, 95%CI 1.18-8.15) (Bang *et al.*, 2001).

In direct contrast, two studies have found a protective effect of the 4G allele on the risk of ischemic stroke. In a study that investigated patients with minor stroke or transient ischemic attack, the protective effect of the 4G allele was limited to those less than 60 years of age (OR 0.4, 95%CI 0.2-0.9) (Endler *et al.*, 2000). This study, however, did not take account of or correct for confounding variables such smoking – this being unequally distributed among cases and controls (cases 74%, controls 25%). In another population-based prospective cohort study involving 12,239 Dutch women followed for up to 18 years, homozygosity for the 4G allele was associated with a lower risk of death from cerebrovascular disease (OR 0.4, 95%CI 0.2-0.7) (Roest *et al.*, 2000). The authors speculated that the protective mechanism did not involve fibrinolysis, but occurred due to a series of reactions promoting atherosclerotic plaque destabilization. However, this study should be interpreted with caution, as it failed to differentiate between ischemic and haemorrhagic stroke.

Two additional studies have reported no association between the PAI-1 5G/4G polymorphism and the risk of ischemic stroke (Catto *et al.*, 1997, Hindorff *et al.*, 2002). In the first, a study of 558 stroke patients, although no difference in the prevalence of the PAI-1 5G/4G genotypes was observed between patients an controls, PAI -1 activity was significantly increased in stroke patients ($p = 0.0001$) and was independently associated with all-cause mortality at one and three months ($p = 0.02$ and 0.03 respectively) (Catto *et al.*, 1997). These findings implied that other influences mediate expression of PAI-1, however the authors failed to address this issue further. In the other negative study, analysis was restricted to women younger than 45 years and thus the finding cannot be extrapolated beyond this group (Hindorff *et al.*, 2002).

2.4.8 Conclusions

This review of the literature has highlighted an important problem facing genetic association studies: a lack of reproducibility. There are several important reasons for this. Firstly, many of the studies were of inadequate sample size and consequently lacked the statistical power to determine a significant genetic risk. Secondly, the demographic and ethnic composition of the study populations (for both cases and controls) varied widely. For example, a genetic risk factor found in a population of young females with cryptogenic stroke may not necessarily be relevant to an older stroke population. Thirdly, the definition and classification of stroke was not uniform between studies. This is particularly important in studies of ischemic stroke, which is a pathologically heterogeneous entity and requires a well-defined and accurate classification. Fourthly, the studies varied in the methods of statistical analysis used and other aspects, such as measurement and adjustment for known cerebrovascular risk factors and ethnicity. Finally, the effect of interaction between potentially pathogenic SNP's with each other or with other cerebrovascular risk factors has not been addressed. Given the polygenic hypothesis of common ischemic stroke, it is probable that the latter mechanism will be the most important. The aim of this thesis is therefore not only to determine the relevance of eight SNP's to stroke, but also to evaluate the effects of interaction with each other and other known cerebrovascular risk factors.

Chapter 3

Thesis Aims and Rationale

3.1 Thesis Aims

Despite considerable evidence for a genetic predisposition to ischemic stroke, the precise genetic determinants remain unknown. Epidemiological studies to date have identified several genetic candidates, however their significance has remained controversial, as conflicting data exists. Furthermore, there is limited information on the genetic predisposition to ischemic stroke in an Australian population (Baker *et al.*, 2001). The purpose of this thesis is to determine the risk of ischemic stroke in an Australian population associated with the following eight candidate gene SNP's:

PON1 –107C/T
PON1 M54L
Glycoprotein 1b 145Thr/Met
Glycoprotein IIb/IIIa PlA1/A2
β fibrinogen –148 C/T
Prothrombin 20210 G /A
TPA –7,351 C/T
PAI-1 5G/4G

The specific aims of this thesis are as follows:

a) To determine the prevalence of each SNP in the general population (without a history of ischemic stroke) residing in metropolitan Adelaide, South Australia.

b) To determine the prevalence of each SNP in patients presenting with acute ischemic stroke to the major hospitals in metropolitan Adelaide, South Australia.

c) To determine the risk of ischemic stroke associated with each SNP.

d) To determine the risk of ischemic stroke stratified for stroke sub-type associated with each SNP.

e) To determine the magnitude of risk of ischemic stroke associated with each SNP (attributable risk).

f) To determine the risk of ischemic stroke associated with the interaction between two SNP's.

g) To determine the risk of ischemic stroke associated with the interaction between each SNP and the following cerebrovascular risk factors:

Family history of stroke
Cigarette smoking
Hypertension
Atrial Fibrillation
Diabetes Mellitus
Hypercholesterolemia

3.2 Selection of Study Design

Association and linkage studies comprise the two main strategies for the detection of a genetic SNP predisposing to ischemic stroke. The latter has proved to be an important application in the identification of genetic mutations underlying many monogenic disorders, including Huntington disease and cystic fibrosis. The methodology focuses on the familial segregation of genes and is based on the principle of linkage disequilibrium, whereby two alleles at separate loci (usually physically close together)

are inherited at a greater frequency than expected by chance. Typically, known DNA markers are applied to large pedigrees of affected and unaffected individuals, with identification of a DNA segment containing the mutant gene occurring when co-segregation with a disease phenotype is found. The great advantage of this method is its ability to identify novel genes. In multifactorial ischemic stroke, however, the application of linkage analysis becomes difficult for several reasons: (1) As linkage analysis relies on the availability of live relatives, the application to common ischemic stroke that occurs predominantly in the elderly is less feasible. (2) Unlike Mendelian disorders, common ischemic stroke is a genetically heterogenous entity with multiple genetic determinants predisposing to the same phenotype. This introduces significant complexity when linkage analysis is used, with the identification of multiple loci requiring a larger number of intact sibling or parental units. For example, if there are 10 known gene loci for ischemic stroke, it has been estimated that linkage analysis (using the affected sibling pair method) would require a sample size of 6,800 affected sibling pairs to show a threefold increase in the relative risk of ischemic stroke (Hassan *et al.*, 2002). This increases to 61,230 affected sibling pairs if 20 ischemic stroke genes were anticipated. (3) Linkage analysis also relies on complete gene penetrance, i.e. those who inherit the disease-causing allele will manifest the disease phenotype. In polygenic ischemic stroke, individual polymorphisms may not independently lead to disease expression, but contribute in a sub-clinical manner to the overall disease burden that is also influenced by environmental risk factors. This in turn limits the phenotypic segregation of DNA variants, an essential requirement for successful linkage analysis. (4) The multifactorial aetiology of common ischemic stroke may also result in phenocopy. This phenomenon occurs when ischemic stroke manifests in individuals who do not carry the disease-causing allele, suffer an ischemic stroke because of other confounding environmental influences. Application of linkage analysis in these cases may dilute the true association of a genetic mutation with the ischemic stroke phenotype (Hassan and Markus, 2000).

The case-control study has been the most widely accepted approach in the identification of genes that predispose to ischemic stroke (Cardon and Bell, 2001). It involves the association of disease with the difference in prevalence of the SNP

between cases and controls. There are several important advantages that this methodology has to offer: ease of subject attainment with more efficient recruitment of cases and controls when compared to family-based strategies; large sample sizes can be recruited; no requirement of genetic information from parents or siblings, allowing for late-onset diseases to be investigated; confounding environmental influences can be adjusted for readily and disease penetrance and attributable risk can be estimated simultaneously (Cardon and Palmer, 2003). Most importantly, case-control studies possess a greater power to identify genetic variants with modest effect when compared with linkage studies (Risch and Teng, 1998).

Despite these advantages, no stroke gene has been unequivocally demonstrated using this method. The inconsistency and lack of reproducibility of positive findings is often ascribed to undetected population stratification. This occurs when there is a difference in allelic frequency between cases and controls that is unrelated to disease outcome, but is attributable to the ethnic diversity of the population studied (Cardon and Palmer, 2003). Other factors responsible for the lack of reproducibility of positive findings include: an inadequate sample size affecting study power; differences in study recruitment; variation of population demographic factors between studies; variable levels of adjustment for confounding risk factors; differing methods of statistical analysis; and differing methods of ischemic stroke classification.

Several alternative strategies have been developed in an attempt to overcome the problem of undetected population stratification. First, implementation of a prospective cohort study essentially eliminates selection bias of a control population. This approach however, consumes significant resources and time, as it requires the recruitment of a substantial number of cases prior to development of disease. Other control sampling strategies include transmission disequilibrium testing (TDT) and sibling transmission disequilibrium testing (S-TDT) which match cases with internal family controls. Although robust in controlling for population stratification, these approaches require parental and sibling genotypes which may be difficult to obtain in late-onset diseases such as ischemic stroke. Furthermore, as TDT is based on transmission of the disease-causing allele from a heterozygous parent to the affected

proband, parental DNA with wild-type or homozygous genotype is disregarded. Consequently, a greater sample size is required to establish statistical significance. It has been estimated that for TDT, in order to demonstrate a doubling of risk of ischemic stroke, screening of 31,680 index cases is required to provide the required 414 cases with intact family units. The number required to screen in a S-TDT and for a matched case-control study to achieve a similar risk profile and sample size is 3,062 and 820 respectively (Hassan *et al.*, 2002).

Given the above information, time available and resource limitations, a well-designed case-control study was considered the best method to address the research objectives. A particular emphasis was placed on the recruitment of controls using measures to minimise the effects of population stratification (see Chapter 4).

3.3 Selection of Candidate SNP's

With over four million human SNP's listed on a public database to date (http://www.ncbi.nlm.nih.gov/SNP), a number of inclusion criteria were defined that allowed for the selection of SNP's for investigation. First, technical considerations limited the number of candidate SNP's that could be studied. A 96-well polymerase chain reaction (PCR) thermal cycling machine, which allowed for the simultaneous amplification of multiple DNA fragments containing polymorphic loci was available for this research. Therefore, for each PCR reaction, the genotyping capacity ranged from 1 SNP for 48 subjects, to 48 SNP's for 1 subject. After consideration of the required sample size (see below), number of PCR reactions required, funding and laboratory resources, genotyping of 8 SNP's (in 6 subjects per PCR) was considered feasible.

The choice of SNP's was also influenced by the availability of genomic sequence during the planning phase of this thesis. Genomic sequence containing the SNP was an essential requirement allowing for oligonucleotide primer design. Only SNP's publicly listed on the Entrez SNP Database (http://www.ncbi.nlm.nih.gov/SNP) were considered as potential candidates.

The most important criterion for selection was biological plausibility. Before consideration, it was essential that the SNP occurred in a gene of known function, with a direct influence on the pathogenesis of ischemic stroke by one or more of the following mechanisms: a) predisposing to atherosclerosis; b) enhancing platelet function; c) augmenting intrinsic coagulation and d) impairing fibrinolysis. Polymorphisms previously reported to predispose or increase the susceptibility to known cerebrovascular risk factors were not considered for selection. SNP's with functional significance that altered the expression or function of the relevant gene product were favoured. An emphasis was placed on SNP's that occurred within the regulatory regions of the gene, with experimental evidence of altered gene expression. In this latter group, consideration was strengthened if a correlation between an objective measure of the gene product (i.e. plasma level of the protein) and ischemic stroke had been reported. Although preference was given to novel SNP's that had not been previously investigated for their relevance to ischemic stroke, strong consideration was also given to SNP's with conflicting findings. In this latter group, consideration was strengthened if positive results came from studies of good epidemiological design. SNP's were not considered if their significance potentially occurred through linkage disequilibrium with another functional gene locus.

3.4 Sample Size Estimation

On fulfilment of biological plausibility criteria, the feasibility of investigating each SNP was then considered given the other limitations discussed above. Using previous hospital admission data and taking into account the clinical resources and recruitment time available, a sample size of up to 500 stroke patients and 500 controls was estimated as a feasible target. An estimate of the sample size required to achieve statistical significance for each SNP was then made, using prevalence data from previous studies. As the genotype distribution may vary between ethnic groups, preference was given to studies performed on Australian or Caucasian populations. SNP's were then considered as candidates if the calculated sample size in at least one estimation was within the pre-specified range of 500 per group.

Table 3.1
Sample size estimations for selected single nucleotide SNP's based on genotype frequency data in a Caucasian population

Polymorphism	Min OR	P* (%)	Sample Size Required**	Reference
PON1–107 C/T				
TT	2.0	42	133 (265)	(Leviev and James, 2000)
TT	2.0	31	166 (332)	(James *et al.*, 2000)
PON1 54M/L				
LL	2.0	42	133 (265)	(Schmidt *et al.*, 1998a)
LL	2.0	34	136 (271)	(James *et al.*, 2000)
LL	2.0	41	133 (265)	(Salonen *et al.*, 1999)
LL	2.0	42	133 (265)	(Leviev and James, 2000)
Gp1b 145Thr/Met				
Thr/Met or Met/Met	2.0	60	152 (305)	(Simsek *et al.*, 1994)
GpIIb/IIIa PlA1/A2				
PlA2 allele	2.0	15	205 (410)	(Newman *et al.*, 1989)
A1/A1 or A1/A2	2.0	19	176 (352)	(Weiss *et al.*, 1996)
β Fibrinogen –148 C/T				
CT or TT	2.0	45	133 (267)	(Schmidt *et al.*, 1998b)
T allele	2.0	19	176 (352)	(Thomas *et al.*, 1994)
T allele	2.0	21	166 (332)	(Behague *et al.*, 1996)
TPA –7351 C/T				
T allele	2.0	39	133 (266)	(Ladenvall *et al.*, 2000)
TT	2.0	12	241 (483)	(Ladenvall *et al.*, 2000)
PAI 5G/4G				
4G/4G	2.0	26	149 (298)	(Margaglione *et al.*, 1998)
4G allele	2.0	45	133 (267)	(Mikkelsson *et al.*, 2000)
4G allele	2.0	53	140 (279)	(Eriksson *et al.*, 1995)

**Background prevalence of the specified genotypes in control populations.*
***Total study sample size in brackets.*

Following the consideration and ranking of over 25 candidate SNP's, the eight highest ranked SNP's that fulfilled the study criteria were selected for investigation. The sample size estimations for these SNP's (and source of prevalence data) are listed in tables 3.1 and 3.2.

Table 3.2
Sample size estimations for selected single nucleotide SNP's based on previous positive case-control studies of ischemic stroke

Polymorphism	Cases (%)	Controls (%)	Sample Size*	Reference
Gp1b 145Thr/Met				
Thr/Met or Met/Met	22.1	10.6	158 (315)	(Gonzalez-Conejero *et al.*, 1998)
Thr/Met or Met/Met	26.5	14.2	174 (349)	(Sonoda *et al.*, 2000)
Gp IIb/IIIa PlA1/A2				
A1/A1 or A1/A2	70	28	22 (43)	(Wagner *et al.*, 1998)
Prothromb. 20210 G/A				
GA or AA	11	2	119 (237)	(De Stefano *et al.*, 1998)

** Estimated sample size per group. Total study sample size in brackets*

Calculations were performed using Win Episcope 2.0 statistical software (Epidecon, University of Edinburgh). Estimated sample sizes were determined using the sample size calculator for a matched case-control study. This was performed in the planning phase of this thesis in May 2000. Therefore calculations were based on the studies published prior to that time. In all estimates, the level of confidence was specified at 95% with a study power of 80%. For polymorphisms without previous reference to ischemic stroke, the sample size estimation was performed using the reported background frequency of the SNP in the general population, with the minimum detectable odds ratio specified at 2.0 (Table3.1). This value was selected, as it represented a risk level that was considered clinically relevant in a sample size of 500.

Where previous studies had found a positive association of the candidate SNP to ischemic stroke, the prevalence of the SNP in both cases and controls was used to estimate the required sample size (Table 3.2).

Chapter 4

Research Methods

4.1 Clinical Methods

4.1.1 Participating Hospitals

The research objectives were addressed using an incident, matched, case-control study. Cases were recruited from one of five hospitals in metropolitan Adelaide, South Australia: The Queen Elizabeth Hospital, The Lyell McEwin Health Service, Flinders Medical Centre, St Andrews Hospital and Memorial Hospital. The following clinical research ethics committees approved this research (respective hospitals and their private/public status in brackets):

The North West Adelaide Health Service, Ethics of Human Research Committee
(The Queen Elizabeth Hospital and Lyell McEwin Heath Service – both public hospitals)

The Adelaide Community Healthcare Alliance Ethics Committee
(Memorial Hospital – private hospital)

St Andrews Hospital Ethics Committee
(St Andrews Hospital – private hospital)

The Flinders Clinical Research Ethics Committee
(Flinders Medical Centre – public hospital)

Patients and controls were recruited over a 28-month period between August 2000 and November 2002.

4.1.2 Recruitment of Ischemic Stroke Cases

Cases considered consisted of inpatients of the above hospitals with an admission diagnosis of acute ischemic stroke. The method of identification of potentially eligible cases varied according to the participating hospital. For The Queen Elizabeth Hospital and Lyell McEwin Health Service, a daily review of the electronic admission databases was undertaken. For the Flinders Medical Centre, a list of potentially suitable patients was obtained from the stroke trials research nurse who would log new stroke admissions on a bi-weekly basis. For private hospitals (Memorial and St Andrews Hospital), the identification of potentially eligible patients was performed by weekly communication with the treating neurologist. Recruitment was performed in approximately fifteen 4-6 week time blocks. An attempt was made to minimise selection bias by approaching consecutive patients presenting to hospital during these time frames. However, as this study spanned across five hospitals and enrolment was undertaken by a single investigator, this could not be strictly adhered to.

In all cases, the researcher was given permission by the treating doctor or head of unit to approach the patient. A verbal consent from the patient (or next of kin if the patient was unable to communicate) was sought initially to review the patient's admission records and investigations. Confirmation of ischemic stroke with a brain computerised tomography (CT) or magnetic resonance imaging (MRI) scan was required to confirm eligibility. In those with normal findings, patients remained eligible if the clinical diagnosis (as made by the treating medical staff and confirmed by the researcher) was in keeping with the WHO definition for acute ischemic stroke (see Chapter 1.3.1) (Hatono, 1976). Those with primary intracerebral haemorrhage (including sub-arachnoid haemorrhage) on neuro-imaging were not eligible for this study. Patients with haemorrhagic transformation on neuro-imaging (i.e., secondary haemorrhage within a cerebral infarct), however, remained eligible. Eligibility was not restricted by age or gender.

Following fulfilment of the initial screening criteria, patients were re-approached by the researcher and invited to participate. This involved discussing the research in detail and providing an opportunity for the patient to review an information sheet (Appendix

II), ask questions of the researcher and discuss it with family members. Informed written consent was then obtained (Appendix III). For those who were unable to give informed consent, provision of study information and written consent was obtained from the next of kin. Although no standardised criteria were used to determine the patient's ability to give informed consent, this was considered carefully during the initial consultation together with the opinion of the treating doctor and family members. A conservative approach was taken when the patient's ability to give consent was questionable. In this situation, written consent was sought from the next of kin. Patients who were unable to give informed consent and who did not have a suitable next of kin to provide informed consent were not eligible for this study.

4.1.3 Questionnaire Assessment of Ischemic Stroke Cases

Following informed consent, patients were interviewed and information was recorded directly into a secured electronic database (Microsoft Access, Version 2000). For patients who were unable to communicate orally, information was obtained from a review of the hospital records and from the next of kin. The following demographic information was recorded:

(a) Name
(b) Gender
(c) Date of birth – age in years determined
(d) Address
(e) Ethnic origin – classified into the following categories:
Caucasian, Asian, Black, Hispanic, Arabic, Other
(f) Number of live siblings
(g) Contact details

A history of known cerebrovascular risk factors was then determined. A binary yes/no response was recorded to following questions:

(a) Has a doctor ever diagnosed you with a stroke?
(b) Has a doctor ever diagnosed you with high blood pressure?
(c) Has a doctor ever diagnosed you with a high cholesterol level?
(d) Has a doctor ever diagnosed you with diabetes?

(e) In the last 5 years, have you smoked cigarettes or tobacco?

(f) Do you have an artificial heart valve?

(g) Has anyone in your family ever suffered a stroke?

A five-year cut-off was chosen for cigarette smoking, as this is the time required for the risk of ischemic stroke attributable to this risk factor to return to a non-smoker's level (Colditz *et al.*, 1988, Wolf *et al.*, 1988).

All medications taken on a regular basis prior to admission were recorded. Medications were further classified into the following categories:

(a) Antiplatelets

(b) Anticoagulants

(c) Antihypertensives

(d) Lipid lowering agents

(e) Other

In all cases, validation of self-reported cerebrovascular risk factors and medications was performed by review of the medical records.

4.1.4 Clinical Assessment of Ischemic Stroke Cases

All cases underwent a neurological examination by a neurologist (JJ) and were classified into one of four clinical categories as defined by the Oxfordshire Community Stroke Project (OCSP) classification system (Bamford *et al.*, 1991):

(a) Total anterior circulation syndrome (TACS)

(b) Partial anterior circulation syndrome (PACS)

(c) Posterior circulation syndrome (PCS)

(d) Lacunar syndrome (LS)

The OCSP classification criteria are listed in Appendix I. Although often used alone, in this study, the recorded stroke sub-type was determined in conjunction with the known brain CT or MRI findings. This was considered important as patients were

examined several days following stroke onset, at which time many of the clinical signs may have resolved. Where the neuro-imaging was normal, patients were classified on the OCSP criteria alone. Where the clinical findings had resolved or were inconsistent with the neuro-imaging findings, the classification was based on site of acute infarction on the CT or MRI scan. For example, under strict OCSP criteria, striatocapsular infarction with no cortical signs would be coded as LS. As arterial occlusion occurs within the proximal middle cerebral artery, in this study it was coded as a PACS.

Cases were also classified into cardioembolic and non-cardioembolic causes. A cardioembolic origin was recorded if it was the most likely aetiology causing ischemic stroke, based on the review of cardiac investigations including echocardiography and electrocardiography. The research protocol did not call for echocardiography to be performed routinely and this was performed at the discretion of the treating neurologist. The presence of atrial fibrillation was determined in all cases by review of the twelve lead electrocardiogram(s) performed on admission. Atrial fibrillation was also deemed present if there was a reported history of paroxysmal atrial fibrillation, regardless of the cardiac rhythm on electrocardiography. All cases with atrial fibrillation were recorded as cardioembolic in origin except when other investigations (i.e. high grade internal carotid artery stenosis) showed otherwise. Patients with cardiac valve prostheses were also coded as cardioembolic in origin if there was no clinical or radiological evidence of an alternate thrombotic source.

Blood pressure was determined in accordance with the WHO recommendations for measuring blood pressure (Chalmers *et al.*, 1999). An average of four blood pressure measurements using a manual sphygmomanometer was determined. Blood pressure was taken in the supine or seated position. To avoid a spuriously elevated blood pressure caused by cerebral infarction or by secondary complications, measurements were taken at least three days after ischemic stroke onset or when the patient was considered by the treating doctor to be clinically stable. The latter was defined as patients who did not show evidence of clinical progression over the preceding 48 hours or have any evidence of a secondary complication (i.e. sepsis, pneumonia, thromboembolism, or raised intracranial pressure).

4.1.5 Case Procedures

Following clinical examination, a random, non-fasting, 20ml venous blood sample was withdrawn from the cubital vein using a needle and syringe under aseptic conditions. Although the researcher usually performed this at the time of recruitment, for patients who had endured repeated blood tests during their admission, blood was collected at the time of their next scheduled blood test by the hospital blood technicians. Although collection of a fasting blood sample for the determination of plasma glucose and cholesterol would have been preferable, this was not feasible with the resources available.

4.1.6 Recruitment of Community Controls

The names and contact details of community controls were obtained from The Population Research and Outcome Studies (PROS) Unit, South Australian Department of Human Services (DHS, 2002). Controls were selected from persons who had previously participated in PROS population surveys between 1999 and 2002. These surveys were conducted using the South Australian Social and Environmental Context Information System (SERCIS). The methodology used a computer-aided telephone interviewing system for selection of participants and data recording. The sample frame consisted of households listed in the South Australian electronic white pages and located within metropolitan Adelaide. Sample households were randomly selected from the electronic white pages. The response rate for SERCIS surveys was high, usually between 70-80% of the initial number of households sampled (DHS, 2002). Within households, the person who was last to have a birthday (aged 18 years and over) was selected for interview. If the selected person was unavailable, an alternative person was not accepted. Under this circumstance, regular callbacks were performed with an individual deemed a "non-contact" following six unsuccessful contact attempts. Following a successful contact with the selected person, verbal consent to participate in the survey was obtained. On conclusion of the survey, permission from the participant was sought to contact them again if the need arose. For this research, only those who had previously agreed to being contacted again were considered as potential community controls. Community controls were matched with cases of ischemic stroke for age and gender. Age was matched to within 5-year strata,

commencing from 21-25 years to 81-85 years. Those less than 21 years of age and greater than 85 years of age were classified in the "less than 21 years" and "greater than 85 years" age categories, respectively. PROS first identified all persons matched for age and gender who had given prior consent to be re-contacted. From this subgroup, persons were contacted randomly by telephone and requested to participate in this research. The request was performed by a professional telephone interviewer who:

(a) Explained the purpose of the study.
(b) Explained what their participation would involve.
(c) Gained agreement (or otherwise) for participation in the study.
(d) Established that the persons had never been diagnosed medically with a stroke.
(e) Explained that participants would be contacted by a member of the research team and established preferred contact details.

Following the identification of potential community controls by PROS, individuals who agreed to participate on initial screening were re-contacted by telephone by the current researcher. The study purpose and methods were then discussed in detail and a confirmation of their acceptance to participate was sought. A negative personal history for ischemic stroke also confirmed as a pre-request for eligibility. Those who agreed were posted a study information sheet (Appendix II) and were given an appointment to be seen on an outpatient basis. Control recruitment clinics were held in the outpatient departments of The Queen Elizabeth Hospital and Lyell McEwin Health Service between December 2000 and December 2002.

4.1.7 Questionnaire Assessment of Community Controls

Following informed written consent, controls were interviewed and information was recorded directly into a secure electronic database (Microsoft Access, Version 2000). Demographic information, prior history of stroke, personal history of cerebrovascular risk factors and medication history was determined as per the ischemic stroke cases (see 4.1.3).

4.1.8 Clinical Evaluation of Community Controls

Blood pressure was determined in accordance with the WHO recommendations for measuring blood pressure (Chalmers *et al.*, 1999). An average of four blood pressure measurements using a manual sphygmomanometer was determined. All blood pressure measurements were taken in the seated position.

4.1.9 Control Procedures

An electrocardiogram was performed on all controls and the presence of atrial fibrillation determined. Atrial fibrillation was also deemed present if there was a personal history of paroxysmal atrial fibrillation, regardless of the cardiac rhythm on electrocardiography. A random, non-fasting, 20ml venous blood sample was drawn from the cubital vein using a needle and syringe under aseptic conditions. To maintain consistency in the methodology, controls were not requested to fast for their blood sample.

4.2 Laboratory Methods

4.2.1 Glucose and Cholesterol Determination

10ml of venous blood was sent to the Institute of Veterinary and Medical Science clinical biochemistry laboratory (The Queen Elizabeth Hospital) for the determination of plasma glucose and total cholesterol (non-fasting).

4.2.2 DNA Extraction

DNA preparation was performed at the Australian Red Cross (South Australian Branch) Tissue-Typing Laboratory. DNA was extracted from 10mls of venous blood using a standard salting-out procedure (Miller *et al.*, 1988) and was stored at 4^0C until tested.

4.2.3 Genotyping Methods

The presence or absence of SNP's was determined using the polymerase chain reaction sequence-specific primer (PCR-SSP) method (Bunce *et al.*, 1995). In this method, the terminal 3'-nucleotide of a primer is designed to base pair with the target polymorphic site. A consensus primer is also designed to target a non-polymorphic region within

300-400 base pairs of the SNP. The use of a SNP-specific primer, under optimised conditions, allows for amplification only when the polymorphism is present. This occurs because Taq polymerase is deficient in exonuclease activity and therefore is unable to repair a mismatched 3'-terminal primer nucleotide, which would facilitate amplification of the wild-type allele (Chien *et al.*, 1976). In this study, two additional oligonucleotide primers were also used in multiplex reactions that generated a 600bp amplicon from the HLA-DRB3 gene. This acted as a positive control that was used to discriminate between failed and negative PCR reactions. PCR-SSP was performed in 96-well PCR plates using a Peltier PTC-200 thermal cycler.

4.2.4 Oligonucleotide Primer Design

Oligonucleotide primers were designed for all SNP's using sequences obtained from the Entrez Human Genome public database (www.ncbi.nlm.nih.gov/SNP) and were synthesised locally (Geneworks Pty Ltd, Adelaide, South Australia). Primers were designed to have a primer-template annealing temperature of 60-62^{0}C and an approximately 50% guanine (G) to cytosine (C) content. A SNP-dependent nucleotide was placed on the terminal 3'-terminal nucleotide of the primer in all cases apart from the PAI 5G/4G SNP, where an internal primer mismatch was created. The oligonucleotide primer sequences, genebank accession codes and PCR product sizes are listed in table 4.1. A BLAST search was performed on all oligonucleotide primer sequences (www.ncbi.nlm.nih.gov/BLAST) and checked for cross-reactivity with other DNA regions. The sequence of the two positive control primers coding for a 600 base pair fragment of the HLA-DRB3 gene is as follows:

Forward primer:	TGCCAAGTGGAGCACCCAA
Reverse primer:	GCATCTTGCTCTGTGCAGAT

Table 4.1
Oligonucleotide Primer Sequences, Genebank Accession Numbers and PCR Product Sizes

Polymorphism	Primer Sequence (5'-3')	Genebank Accession No.	Position (5'-3')	Primer Size (Mer)	Product Size (Mer)
PON1 –107C/T					
C	CCG ATT GGC CCG CCC CG	AF05133	927 - 911	17	
T	CCG ATT GGC CCG CCC CA	AF05133	927 - 911	17	
Consensus	CAA GGA CCG GGA TGG CAC	AF05133	654 - 671	18	274
PON1 M54L					
M	CAG AAA CTG GCT CTG AAG ACA	U55879	48 - 68	21	
L	CAG AAA CTG GCT CTG AAG ACT	U55879	48 - 68	21	
Consensus	AAG TGG GCA TGG GTA TAC AG	S56546	306 - 287	20	150
Gp1b145T/M					
T	CCA GCT TGG GTG TGG GCG	NM_000173	541 - 524	18	
M	CCA GCT TGG GTG TGG GCA	NM_000173	541 - 524	18	
Consensus	CGA AAG ACA CAA CCA TCC TC	NM_000173	179 - 198	20	363
GpIIb/IIIa Pl A1/A2					
A1	CAC AGC GAG GTG AGC CCA	M32672	1515 - 1498	18	
A2	CAC AGC GAG GTG AGC CCG	M32672	1515 - 1498	18	
Consensus	CTG TAC AAC GGT CCT AAG GG	M32672	1321 - 1340	20	195

Table 4.1 (continued)
Oligonucleotide Primer Sequences, Genebank Accesion Numbers and PCR Product Sizes

Polymorphism	Primer Sequence (5'-3')	Genebank Accession No.	Position (5'-3')	Primer Size (Mer)	Product Size (Mer)
β Fib –148 C/T					
C	GCA ACA TCT TCC CAG CAA AG	X05018	1371 - 1352	20	
T	GCA ACA TCT TCC CAG CAA AA	X05018	1371 - 1352	20	
Consensus	GTC ACG ATT TTA GTG GTT GCC	X05018	1112 - 1132	21	260
Proth. 20210G/A					
G	CCC AAT AAA AGT GAC TCT CAG CG	M17262	26762 - 26784	23	
A	CCC AAT AAA AGT GAC TCT CAG CA	M17262	26762 - 26784	23	
Consensus	GAG TGC TCG GAC TAC CAG CGT GC	M17262	26907 - 26685	23	146
TPA –7351 C/T					
C	ATG GCT GTG TCT GGG GCG	Z48484	2245-2228	19	
T	ATG GCT GTG TCT GGG CA	Z48484	2245-2228	19	
Consensus	ATT GGC GCA AAC TCC TCA	Z48484	1841-1858	18	405
PAI 5G/4G					
5G	ACA CGG CTG ACT CCC CCA	J03764	2504 - 2487	18	
4G	ATA CAC GGC TGA CTC CCC A	J03764	2506 - 2487	19	
Consensus	CCT AAA AGC ACA CCC TGC AA	J03764	2193 - 2212	20	5G=312, 4G=313

4.2.5 PCR-SSP Method

Oligonucleotide Primer Preparation

The final working concentrations (and manufacturers) of the components for a control primer stock containing the HLA-DRB3 positive control primers are listed below:

Cresol red	62µg/ml	(Fluka)
HLA-DRB3 forward primer	5µg/ml	(GeneWorks)
HLA-DRB3 reverse primer	5µg/ml	(GeneWorks)

SNP-specific primers were then added to the above. As the primers for each SNP varied in composition and hence DNA annealing capacity, an initial optimisation step was performed to determine the SNP-specific primer concentration that achieved optimal amplification of the target DNA fragment. The final primer concentrations per reaction are listed in table 4.2.

Table 4.2
Primer Reaction Mix for each SNP

Polymorphism	Allele-Specific Primer (µg/ml)	Consensus Primer (µg/ml)
(PON1) –107C/T	20	20
(PON1) M54L	20	20
Gp Ib 145 T/M	20	20
Gp IIb/IIIa PlA1/A2	10	10
β Fibrinogen –148 C/T	10	10
Prothrombin 20210G/A	10	10
TPA –7351C/T	20	20
PAI-1 5G/4G	10	10

5µl of each primer mix was then dispensed in separate wells of a 96-well PCR plate. 10µl of paraffin oil was added to each well to minimise primer and reagent evaporation during PCR thermal cycling.

DNA Preparation

The following reagent mixtures (final concentrations listed) were prepared in advance and used in the preparation of DNA:

(a) 10x PCR Buffer

Tris Base 67mM	(Sigma Chemicals)
Ammonium Sulphate 17mM	(BDH Laboratory Supplies)
pH 8.9 with concentrated hydrochloric acid	(BDH Laboratory Supplies)
5ml Tween 20 per 500ml	(BDH Laboratory Supplies)

(b) TMDH Mixture

6ml 10x PCR buffer	
6ml 10mM dNTP (nucleotides)	(Applied Biosytems)
5.1 ml 25mM Magnesium Chloride	(Applied Biosytems)
6ml autoclaved, deionised water	

The DNA solution was prepared using the following protocol:

100µl TMDH	
57µl autoclaved, deionised water	
4.25 Units AmpliTaq DNA Polymerase	(Applied Biosytems)
3µl DNA	

8µl of DNA solution was then added to each well containing the allele-specific primers. A colour change (pink - mediated by cresol red) confirmed the presence of DNA in the PCR reaction mix.

PCR Thermal Cycling

PCR was performed using a PTC-200 Peltier Thermal Cycler (MJ Research)

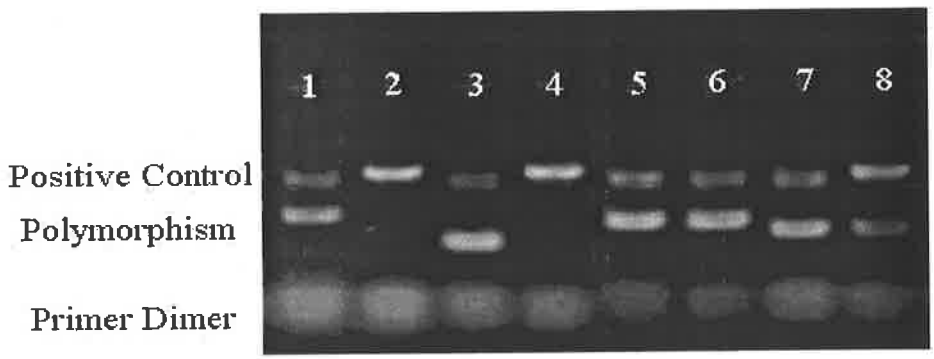
The following PCR cycling parameters were used:

5 cycles of:	96^0C for 25 seconds	70^0C for 45 seconds	72^0C for 45 seconds
21 cycles of:	96^0C for 25 seconds	65^0C for 50 seconds	72^0C for 45 seconds
4 cycles of:	96^0C for 25 seconds	55^0C for 60 seconds	72^0C for 125 seconds

4.2.6 Genotype Determination

PCR-SSP products were fractionated on 2% agarose gels and visualized under ultraviolet light following ethidium bromide staining (Welsh and Bunce, 1999). Gels were photographed using a high-resolution digital camera (Olympus ColourPix X990) and the genotype determined by visual inspection. The genotype determination of four SNP's in one subject is illustrated in figure 4.1.

Figure 4.1
Genotype Determination by Visual Inspection: Glycoprotein Ib T/M, Glycoprotein IIIaPlA1/A2, PAI 5G/4G and Fibrinogen 148 C/T Genotype In One Subject.



1. Glycoprotein Ib T (wild-type)
2. Glycoprotein Ib M (SNP)
3. Glycoprotein IIIa PlA1 (wild-type)
4. Glycoprotein IIIa PlA2 (SNP)
5. PAI 5G (wild-type)
6. PAI 4G (SNP)
7. Fibrinogen 148 C (wild-type)
8. Fibrinogen 148 T (SNP)

A larger sized positive control band indicates a successful PCR reaction. The absence of the Glycoprotein Ib methionine band indicates that this individual is homozygous for the wild-type allele. For fibrinogen 148 C/T, positive bands are present for both alleles, indicating a heterozygous state. The genotype for this individual is therefore: Glycoprotein Ib T/T, Glycoprotein IIIa PlA1/PlA1, PAI 5G/4G, Fibrinogen 148 C/T

The primer dimer band represents unused oligonucleotide primer.

4.2.7 Quality Assurance Measures

Two quality assurance measures were undertaken to ensure that genotype determination using SSP-PCR was accurate.

SSP-PCR Reproducibility

Two independent investigators genotyped the first 100 subjects for all eight SNP's. Investigators were blinded to the source of DNA. Genotyping using this approach was fully concordant.

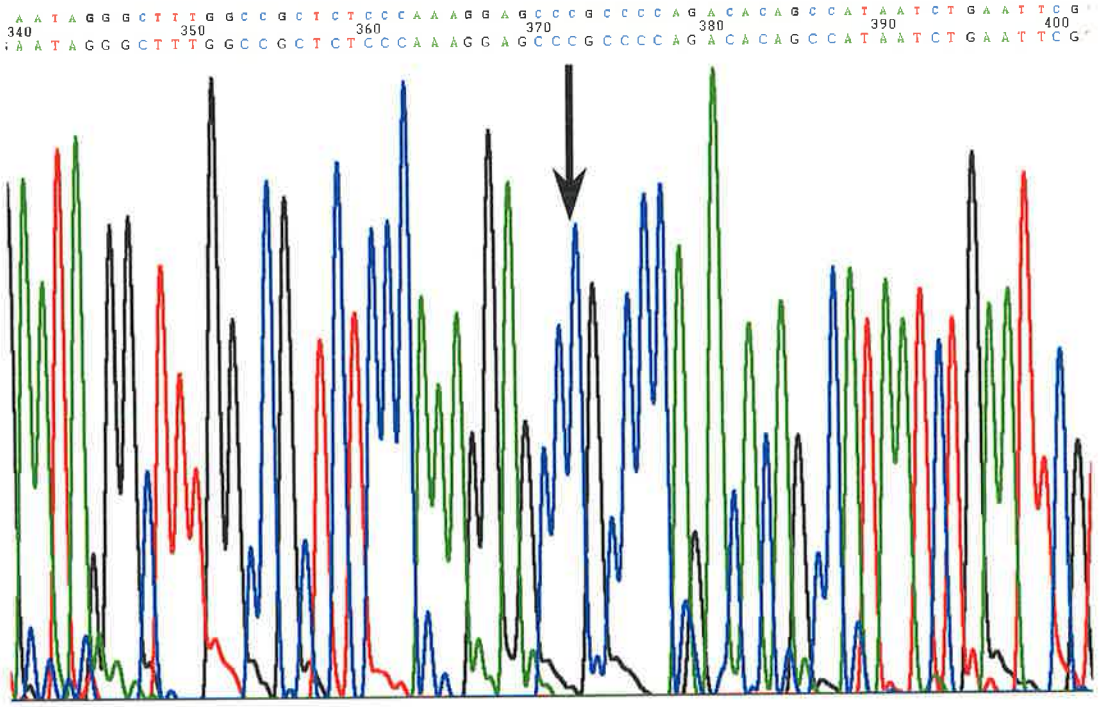
SSP-PCR Product Sequencing

The allele-specific bands were visualised on 2% agarose gels to check for correct amplicon size. Products were also sequenced to confirm amplification of the specific target. This process was performed for each polymorphism.

Amplicons to be sequenced were extracted from agarose gel using the Wizard PCR Preps DNA Purification System (Promega). SSP-PCR products were ligated into a pGEM–T bacterial vector using a pGEM–T Easy Vector System II (Promega) and then transformed into JM109 competent bacteria by heat shock. Transformed cells were grown on ampicillin (+) agar plates treated with IPTG (0.5mM) and X-Gal (80μg/ml). Positive transformants were selected by blue/white colony screening. Selected colonies were streaked onto another plate and grown overnight at 37^0C. Luria broth media with ampicillin (25mg/ml) was then inoculated with a single colony and grown overnight at 37^0C with shaking. Plasmid minipreps were performed using the QIAprep Spin Miniprep Kit (Qiagen).

DNA sequencing of the purified pGEM-T plasmid was performed by the Sequencing Facility, Flinders Medical Centre, South Australia. The sequence was then crosschecked with the known sequence of the intended DNA target containing the SNP. An example of the resultant chromatogram file (TPA –7351 C/T) is illustrated in Figure 4.1.

Figure 4.2
Chromatogram Showing DNA Sequence of TPA –7351 C/T SNP



The TPA –7351 C/T SNP is located at position 372 on the chromatogram (arrow). The C allele is shown.

4.3 Statistical Methods

Statistical analysis was performed by Covance Pty. Ltd. (Canberra, ACT) using STATA (Version 7) statistical software. All tests are two-sided. Statistical significance was defined at a p value of < 0.05.

4.3.1 Descriptive Statistics

The age and gender for both consenting and non-consenting stroke patients was determined that allowed for the identification of case selection bias. A descriptive comparison of other demographic variables and cerebrovascular risk factors was not performed, as this was considered unethical. The proportion of controls that refused to participate in this study after initial and subsequent screening was not disclosed by

PROS. Consequently, a similar descriptive analysis between consenting and non-consenting controls could not be performed.

4.3.2 Univariate Analysis

Unconditional logistic regression was used to determine the univariate odds ratios for stroke for each of the following demographic factors:

(a) Age

Age was condensed into six age-group strata: <40, 40-49, 50-59, 60-69, 70-79, 80+

(b) Gender

(c) Ethnic Origin

A dichotomous variable for Caucasian and non-Caucasian origin was created.

Logistic regression was also used to determine the univariate odds ratio of stroke for the following cerebrovascular risk factors:

(d) Cigarette smoking - defined as a history of smoking in the past 5 years

(e) Hypertension - defined as a self-reported history of hypertension or a measured blood pressure >140mmHg systolic or >90mmHg diastolic

(f) Hypercholesterolemia - defined as a self-reported of hypercholesterolemia or a random cholesterol level >5.5mmol/l

(g) Diabetes - defined as a self-reported history of diabetes or a random blood glucose >11.1mmol/l

(h) Atrial fibrillation – defined as a self-reported history of paroxysmal atrial fibrillation or presence of atrial fibrillation on electrocardiography

(i) Family history of stroke - defined as a self-reported of stroke in a first-degree relative

A univariate odds ratio for the use of the following medication sub-types was determined: antiplatelet, anticoagulants, antihypertensives, lipid lowering medication.

Unconditional logistic regression analysis was used to obtain estimates of odds ratios for stroke for each polymorphism. Three separate parameterisations were fitted for each polymorphism:

(a) By genotype using a three level categorical variable: wild-type, heterozygous, homozygous; wild-type used as the reference category.

(b) By clustering using a two level categorical variable: wild-type, heterozygous and homozygous combined; wild-type used as the reference category.

(c) By allelic frequency using a two level categorical variable representing either allele.

Separate logistic regression models were formed to assess the univariate risk of stroke associated with each polymorphism stratified for the following stroke subtypes:

(a) Lacunar

(b) Non-Lacunar Stroke: Defined as those with a total anterior circulation, partial anterior circulation or posterior circulation syndrome.

(c) Cardioembolic Stroke

(d) Non-cardioembolic Stroke

4.3.3 Multivariate Analysis

Unconditional logistic regression was used to determine the risk of stroke associated with each SNP following adjustment for known cerebrovascular risk factors. Two methods were used to determine which variables would be included in the final multivariate models. Firstly, a set of "important variables" with respect to each polymorphism was identified. This was achieved using a bivariate analysis that determined the effect of adjusting for each of the variables listed in section 4.3.2. Although age and gender were matched for cases and controls, these variables were also considered as independent, potentially confounding variables. A bivariate analysis for each SNP and the following continuous variables was also performed: systolic blood pressure; diastolic blood pressure; random glucose level and a non-fasting cholesterol level. Variables that altered the univariate odds ratio (associated with either the heterozygous or homozygous state for the pathogenic allele) by at least 10% were then incorporated into a multivariate model.

The second method involved incorporation of all potential confounding variables into an unconditional logistic regression model for each SNP. This method was performed

because these factors considered together may result in considerable confounding despite individually having a small influence.

4.3.4 Interactions

The above multivariate models were also used to determine the effect of possible gene-environment and gene-gene interactions. Given the number of variables assessed (and therefore the numerous interaction permutations), interaction assessment was limited to the SNP's that were found to be independently associated with ischemic stroke. This involved creating interaction variables that were incorporated into the final models and allowed for the determination of the following:

(a) Interaction between the pathogenic SNP(s) and each of the risk factors listed in 4.3.2.

(b) Interaction between the pathogenic SNP(s) and the other SNP's investigated.

Chapter 5

Results

5.1 Study Population

Within the study recruitment period between August 2000 and December 2002, 201 patients with a diagnosis of acute ischemic stroke were invited to take part in this study. 182 (90.5%) patients (or their next of kin) gave informed, written consent to participate. 137 (75%) patients presented with their first-ever ischemic stroke. Most of those who were not enrolled were severely cognitively impaired and did not have a suitable next of kin to provide informed consent on their behalf (10/19 cases). Three cases died within 72 hours, despite an earlier intention to participate. The remaining non-consenters made an informed decision not to participate. There was no difference in terms of age and gender between consenters (both cases and controls) and non-consenters (table 5.1). Although the ethnic origin and cerebrovascular risk factor profile for non-consenters was available to the researcher, it was not recorded, as this was considered unethical.

The number of ischemic stroke patients recruited was considerably lower than that anticipated. This was mainly due to the withdrawal from participation by a major public hospital in metropolitan Adelaide, subsequent to the commencement of this study. This hospital manages approximately 300 stroke patients per year and was considered an important source of patients during the planning phase of the study. The withdrawal occurred as a satisfactory agreement on authorship from publications generated from this thesis could not be reached.

The contact details of 320 potential community controls were provided by The Population Research and Outcome Studies (PROS) Unit (South Australian Department of Human Services). Of these, 301 (94%) gave informed, written consent to participate

between August 2000 and December 2002. A high consent rate was achieved as PROS only forwarded the contact details of individuals who had voiced interest in this study during prior screening (see Chapter 4.1.6 – Recruitment of Community Controls). The proportion of individuals that refused to participate during the initial screening process was not disclosed by PROS and therefore this information is unavailable. Controls were matched to stroke cases for age (within 5-year strata) and gender. As the number of ischemic stroke cases was lower than anticipated, an effort was made to recruit as many controls as possible within the recruitment period in order to maximise study power. Consequently, this study achieved a patient to control ratio of 1:1.65.

Table 5.1
Gender and Age Characteristics of Consenters and Non-Consenters

Variable	Consent n (%)	Non-Consent n (%)
Gender		
Female	214 (44)	10 (53)
Male	269 (56)	9 (44)
Age		
<40	8 (2)	0 (0)
40-49	16 (3)	0 (0)
50-59	35 (7)	2 (11)
60-69	81 (17)	3 (16)
70-79	184 (38)	7 (37)
80+	159 (33)	7 (37)
Total*	483 (100)**	19 (100)

** Includes both patients and controls. ** Comprised of 182 patients and 301 controls.*

The demographic characteristics and prevalence of cerebrovascular risk factors for cases and controls are shown in table 5.2. No significant differences were observed between cases and controls in terms of demographic variables. 71% of cases were above 70 years of age, with more males (56%) than females (45%). Most were of Caucasoid origin. Although more non-Caucasoid cases were recruited than controls (2% vs 0.3%), the absolute number was small and the difference was not statistically significant (OR 6.7, 95%CI 0.8-60.6).

Table 5.2
Demographic Characteristics Between Controls and Cases.

Variable	Controls n (%)	Cases n (%)	Odds Ratio (95%CI)	p
Gender				
Female	134 (45)	80 (44)		
Male	167 (55)	102 (56)	1.0 (0.7-1.5)	0.9
Age (Years)				
<40	6 (2)	2 (1)	1.0	
40-49	10 (4)	6 (3)	1.8 (0.3-12.0)	0.5
50-59	22 (7)	13 (7)	1.8 (0.3-10.1)	0.5
60-69	49 (16)	32 (18)	2.0 (0.4-10.3)	0.4
70-79	120 (40)	64 (35)	1.6 (0.3-8.2)	0.6
80+	94 (31)	65 (36)	2.1 (0.4-10.6)	0.4
Ethnic Origin				
Caucasoid	300 (99.7)	178 (98)	1.0	
Non-Caucasoid*	1 (0.3)	4 (2)	6.7 (0.8-60.6)	0.1
Total	**301 (100)**	**182 (100)**		

** Aboriginal, Hispanic, Arabic, Asian*

5.2 Univariate Analysis: Stroke Risk Factors

The risk of ischemic stroke associated with each of the known cerebrovascular risk factors is shown in table 5.3. Of these, atrial fibrillation was associated with the highest risk of ischemic stroke, with 23% of cases versus 3% of controls affected (OR 8.5, 95% CI 4.1-17.4). A history of smoking within the last five years (OR 3.1, 95% CI 1.9-5.2) and diabetes (OR 2.9, 1.8 - 4.7) were also found to be associated significantly with ischemic stroke. No association was observed between ischemic stroke and a history of stroke in a first-degree relative. In contrast, the prevalence of hypertension and hypercholesterolemia was lower in cases than controls, with the difference being statistically significant (hypertension OR 0.6, 95%CI 0.4-0.9, hypercholesterolemia OR 0.5, 95%CI 0.3-0.7).

Table 5.3
Cerebrovascular Risk Factor Characteristics Between Controls and Cases.

Risk Factor	Controls n (%)	Cases n (%)	Odds Ratio (95%CI)	p
Smoking*				
No	271 (90)	135 (74)		
Yes	30 (10)	47 (26)	3.1 (1.9-5.2)	<0.0001
Hypertension**				
No	53 (18)	50 (27)		
Yes	248 (82)	132 (73)	0.6 (0.4-0.9)	0.01
Hypercholesterolemia†				
No	110 (37)	99 (54)		
Yes	191 (63)	83 (46)	0.5 (0.3-0.7)	<0.0001
Diabetes Mellitus††				
No	267 (89)	132 (73)		
Yes	34 (11)	49 (27)	2.9 (1.8-4.7)	<0.0001
Atrial Fibrillation				
No	291 (97)	141 (77)		
Yes	10 (3)	41 (23)	8.5 (4.1-17.4)	<0.0001
Family History‡				
No	201 (67)	122 (67)		
Yes	100 (33)	60 (33)	1.0 (0.7-1.5)	1.0
Total	**301 (100)**	**182 (100)**		

** Defined as cigarette smoking in the last 5 years*
*** Defined as a self-reported history of hypertension or blood pressure>140mmHg systolic or >90mmHg diastolic*
† Defined as a self-reported history of hypercholesterolemia or random cholesterol level > 5.5mmol/l
†† Defined as a self-reported history of diabetes or random blood glucose>11.1 mmol/l
‡ Defined as a history of stroke in a first-degree relative

To further evaluate the apparent protective effect of hypertension, a univariate analysis for several definitions of hypertension was performed (table 5.4).

Table 5.4
Hypertension Between Controls and Cases.

Risk Factor	Controls n (%)	Cases n (%)	Odds Ratio (95%CI)	p
Hypertension (Composite)*				
No	53 (18)	50 (27)		
Yes	248 (82)	132 (73)	0.6 (0.4-0.9)	0.01
Hypertension Grade**				
None	81 (27)	95 (52)		
Grade 1	22 (7)	6 (3)	0.2 (0.1-0.6)	0.003
Grade 2	18 (6)	6 (3)	0.3 (0.1-0.8)	0.01
Grade 3	21 (7)	4 (2)	0.2 (0.1-0.5)	0.001
Isolated Systolic	159 (53)	71 (39)	0.4 (0.3-0.6)	<0.0001
Hypertension (History)†				
No	164 (54)	87 (48)		
Yes	137 (46)	95 (52)	1.3 (0.9-1.9)	0.2
Total	**301 (100)**	**182 (100)**		

** Defined as a self-reported history of hypertension or blood pressure>140mmHg systolic or >90mmHg diastolic*
*** Hypertension grade as defined by the WHO classification (Chalmers et al., 1999)*
† *Defined as a self-reported history of hypertension alone*

Stratification according to the World Health Organization blood pressure criteria (Chalmers *et al.*, 1999) showed a predominance of isolated systolic hypertension amongst controls. This suggests a "white coat" hypertensive effect that may have led to the misclassification of hypertension in these subjects (see Chapter 6.8 – "Study Limitations" for further discussion). When a history of hypertension alone is considered, the prevalence of hypertension amongst cases and controls is reversed, although the difference is not statistically significant (OR 1.3, 95%CI 0.9-1.9).

Table 5.5 provides information regarding the apparent protective effect of hypercholesterolemia. When a history of hypercholesterolemia alone is considered, the apparent protective effect of hypercholesterolemia is lost, with no significant difference between cases and controls. As cholesterol levels were determined in the

non-fasting state, this observation suggests that misclassification bias may have occurred (see Chapter 6.8 - "Study limitations" for further explanation).

Table 5.5
Hypercholesterolemia Between Controls and Cases.

Risk Factor	Controls n (%)	Cases n (%)	Odds Ratio (95%CI)	p
Hypercholesterolemia (Composite)*				
No	110 (37)	99 (54)		
Yes	191 (63)	83 (46)	0.5 (0.3-0.7)	<0.0001
Hypercholesterolemia (History)**				
No	182 (60)	117 (64)		
Yes	119 (40)	65 (36)	0.9 (0.6-1.2)	0.4
Total	301 (100)	182 (100)		

** Defined as a self-reported history of hypercholesterolemia or random cholesterol level > 5.5mmol/l*
*** Defined as a self-reported history of hypercholesterolemia alone*

5.3 Univariate Analysis: Medication

Table 5.6 illustrates the distribution of medication use between cases and controls. No statistically significant difference was observed between cases and controls for the use of antiplatelet, anticoagulant, lipid lowering or antihypertensive medication.

Table 5.6
Medication Use Between Controls and Cases.

Medication Class	Controls n (%)	Cases n (%)	Odds Ratio (95%CI)
Antiplatelet			
No	189 (63)	105 (58)	
Yes	112 (27)	77 (42)	1.24 (0.9-1.8)
Lipid Lowering			
No	226 (75)	139 (76)	
Yes	75 (25)	43 (24)	0.9 (0.6-1.4)
Anticoagulant			
No	295 (98)	171 (94)	
Yes	6 (2)	11 (6)	3.2 (1.2-8.3)
Antihypertensive			
No	165 (55)	90 (49)	
Yes	136 (45)	92 (51)	1.2 (0.9-1.8)
Total	301 (100)	182 (100)	

5.4 Univariate Analysis: SNP's

Two cases died after enrolment in the study and prior to venous blood sampling, thus genetic analysis could not be performed in these cases. Table 5.7 shows the distribution of the SNP genotypes for the remaining 180 cases and 301 controls. No cases were identified to be homozygous for the Glycoprotein 1b HPA 2b allele. Similarly, no individuals (cases or controls) possessed the Prothrombin 20210 AA genotype. Using logistic regression for each SNP stratified for genotype, no significant association with ischemic stroke was identified. The stratification of each SNP by allelic frequency is shown in table 5.8. Amongst the control group, the prevalence of pathogenic alleles ranged from 1% (Prothrombin 20210A allele) to 65% (PON1 54 L). The univariate analysis for each SNP by allelic frequency did not reveal any significant association with ischemic stroke.

Table 5.7
Univariate Analysis of SNP's (by genotype) between Controls and Cases.

Genotype	Controls n (%)	Cases n (%)	Odds Ratio (95%CI)	p
PON 1 –107T/C				
TT	61 (20)	44 (24)	1.0	
TC	200 (66)	97 (54)	0.7 (0.4-1.1)	0.1
CC	40 (13)	39 (22)	1.4 (0.8-2.4)	0.3
PON 1 M54L				
MM	36 (12)	20 (11)	1.0	
ML	138 (46)	82 (46)	1.1 (0.6-2.0)	0.8
LL	127 (42)	78 (43)	1.1 (0.6-2.0)	0.7
Gp 1b HPA 2a/b				
aa	256 (85)	158 (88)	1.0	
ab	43 (14)	22 (12)	0.8 (0.5-1.4)	0.5
bb	2 (1)	0 (0)	-	-
Gp IIb/IIIa PlA1/A2				
A1A1	213 (71)	116 (64)	1.0	
A1A2	81 (27)	61 (34)	1.4 (0.9-2.1)	0.1
A2A2	7 (2)	3 (2)	0.8 (0.2-3.1)	0.7
β Fibrinogen –148C/T				
CC	190 (63)	114 (63)	1.0	
CT	102 (34)	58 (32)	0.9 (0.6-1.4)	0.8
TT	9 (3)	8 (4)	1.5 (0.6-4.0)	0.4
Prothrom. 20210G/A				
GG	293 (97)	174 (97)	1.0	
GA	8 (3)	6 (3)	1.3 (0.4-3.7)	0.7
AA	0 (0)	0 (0)	-	-
TPA –7351C/T				
CC	137 (46)	74 (41)	1.0	
CT	136 (45)	83 (46)	1.1 (0.8-1.7)	0.5
TT	28 (9)	23 (13)	1.5 (0.8-2.8)	0.2
PAI-1 5G/4G				
5G5G	61 (20)	42 (23)	1.0	
5G4G	158 (52)	91 (51)	0.8 (0.5-1.3)	0.5
4G4G	82 (27)	47 (26)	0.8 (0.5-1.4)	0.5
Total	301 (100)	180 (100)		

Table 5.8
Univariate Analysis of SNP's (by allele) between Controls and Cases.

Allele	Controls n (%)	Cases n (%)	Odds Ratio (95%CI)	p
PON 1 –107T/C				
T	322 (53)	185 (51)	1.0	
C	280 (47)	175 (49)	1.0 (0.8-1.4)	0.5
PON 1 M54L				
M	210 (35)	122 (34)	1.0	
L	392 (65)	238 (66)	1.1 (0.8-1.4)	0.8
Gp 1b HPA 2a/b				
a	555 (92)	338 (94)	1.0	
b	47 (8)	22 (6)	0.8 (0.5-1.3)	0.3
Gp IIb/IIIa PlA1/A2				
A1	507 (84)	293 (81)	1.0	
A2	95 (16)	67 (19)	1.2 (0.9-1.7)	0.3
β Fib –148C/T				
C	482 (80)	286 (79)	1.0	
T	120 (20)	74 (21)	1.0 (0.8-1.4)	0.8
Prothromb 20210G/A				
G	594 (99)	354 (99)	1.0	
A	8 (1)	6 (2)	1.3 (0.4-3.7)	0.7
TPA –7351C/T				
C	410 (68)	231 (64)	1.0	
T	192 (32)	129 (36)	1.2 (0.9-1.6)	0.2
PAI-1 5G/4G				
5G	280 (47)	175 (49)	1.0	
4G	322 (53)	185 (51)	0.9 (0.7- 2.0)	0.5
Total	602 (100)	360 (100)		

5.5 Bivariate Analysis: Identification of Confounders

A bivariate analysis for each SNP and known cerebrovascular risk factors was performed. Both categorical and continuous risk factor variables were included in the analysis (methods described in section 4.3.3). The determination of 'important' confounding variables is shown in Appendix IV. A summary of the 'important' confounding variables is listed below:

PON 1 –107T/C: nil

PON 1 M54L: smoking, DBP, atrial fibrillation

Gp 1b HPA 2a/b:	atrial fibrillation
Gp IIb/IIIa PlA1/A2:	SBP, hypertension, cholesterol, ethnic origin, glucose
β Fib –148C/T:	SBP, DBP, cholesterol, glucose, diabetes, atrial fibrillation
Prothrombin 20210G/A:	glucose, diabetes, atrial fibrillation
TPA –7351C/T:	atrial fibrillation
PAI-1 5G/4G:	DBP, atrial fibrillation

5.6 Multivariate Analysis: SNP's

Two statistical methods were used to adjust for the confounding effects of known cerebrovascular risk factors. First, factors that were identified as 'important' variables on bivariate analysis (as above) were incorporated into an unconditional logistic regression model for each SNP. The second method involved incorporation of all potential confounding variables into an unconditional logistic regression model for each SNP (see section 4.3.3 – "Multivariate Analysis"). The latter method incorporated the following variables in separate logistic regression models for each SNP: age, gender, ethnicity, smoking, history of hypertension, history of hypercholesterolemia, history of diabetes, atrial fibrillation and family history of stroke.

Table 5.9 shows the risk of ischemic stroke for each SNP (stratified for genotype) after adjustment for "important' confounding variables as determined in section 5.5. A 1.9-fold increased risk of ischemic stroke was associated with the TPA –7351 TT genotype (OR 1.9, 95%CI 1.0-3.6, p=0.049).

Although the prevalence of the TPA –7351 CT genotype was higher for stroke patients, this was not statistically significant (OR 1.2, 95%CI 0.8-1.9, p=0.3). No other significant association was observed. Table 5.10 illustrates the risk of ischemic stroke associated with each SNP (stratified by allelic frequency), adjusted for 'important' confounding variables as determined in 5.4. No significant association was observed, however, the risk of ischemic stroke associated with the TPA –7351 T allele fell just short of statistical significance at the 5% level (OR 1.3, 95%CI 1.0-1.8).

Table 5.9
Multivariate Analysis of SNP's (by genotype) between Controls and Cases Adjusted for 'Important' Confounders.

Genotype	Controls n (%)	Cases n (%)	Odds Ratio* (95%CI)	p
PON 1 –107T/C				
TT	61 (20)	44 (24)	1.0	
TC	200 (66)	97 (54)	0.7 (0.4-1.1)	0.1
CC	40 (13)	39 (22)	1.4 (0.8-2.4)	0.3
PON 1 M54L				
MM	36 (12)	20 (11)	1.0	
ML	138 (46)	82 (46)	1.2 (0.6-2.4)	0.6
LL	127 (42)	78 (43)	1.0 (0.5-2.1)	1.0
Gp 1b HPA 2a/b				
aa	256 (85)	158 (88)	1.0	
ab	43 (14)	22 (12)	0.7 (0.4-1.3)	0.3
bb	2 (1)	0 (0)	-	-
Gp IIb/IIIa PlA1/A2				
A1A1	213 (71)	116 (64)	1.0	
A1A2	81 (27)	61 (34)	1.5 (0.9-2.4)	0.1
A2A2	7 (2)	3 (2)	0.6 (0.1-3.0)	0.5
β Fib –148C/T				
CC	190 (63)	114 (63)	1.0	
CT	102 (34)	58 (32)	1.1 (0.7-1.8)	0.6
TT	9 (3)	8 (4)	1.7 (0.5-5.9)	0.4
Prothromb 20210G/A				
GG	293 (97)	174 (97)	1.0	
GA	8 (3)	6 (3)	1.1 (0.3-8.8)	0.9
AA	0 (0)	0 (0)	-	-
TPA –7351C/T				
CC	137 (46)	74 (41)	1.0	
CT	136 (45)	83 (46)	1.2 (0.8-1.9)	0.3
TT	28 (9)	23 (13)	1.9 (1.0-3.6)	0.05
PAI-1 5G/4G				
5G5G	61 (20)	42 (23)	1.0	
5G4G	158 (52)	91 (51)	1.1 (0.7-1.8)	0.6
4G4G	82 (27)	47 (26)	0.9 (0.5-1.7)	0.4
Total	301 (100)	180 (100)		

**Adjusted for 'important' cerebrovascular risk factors as determined in Appendix IV*

Table 5.10
Multivariate Analysis of SNP's (by allele) between Controls and Cases Adjusted for 'Important' Confounders

Allele	Controls n (%)	Cases n (%)	Odds Ratio* (95%CI)	p
PON 1 –107T/C				
T	322 (53)	185 (51)	1.0	
C	280 (47)	175 (49)	1.1 (0.8-1.4)	0.5
PON 1 M54L				
M	210 (35)	122 (34)	1.0	
L	392 (65)	238 (66)	1.0 (0.7-1.3)	0.8
Gp 1b HPA 2a/b				
a	555 (92)	338 (94)	1.0	
b	47 (8)	22 (6)	0.7 (0.4-1.2)	0.2
Gp IIb/IIIa PlA1/A2				
A1	507 (84)	293 (81)	1.0	
A2	95 (16)	67 (19)	1.2 (0.8-1.9)	0.3
β Fib –148C/T				
C	482 (80)	286 (79)	1.0	
T	120 (20)	74 (21)	1.2 (0.8-1.7)	0.4
Prothromb 20210G/A				
G	594 (99)	354 (99)	1.0	
A	8 (1)	6 (2)	1.1 (0.3-3.8)	0.9
TPA –7351C/T				
C	410 (68)	231 (64)	1.0	
T	192 (32)	129 (36)	1.3 (1.0-1.8)	0.06
PAI-1 5G/4G				
5G	280 (47)	175 (49)	1.0	
4G	322 (53)	185 (51)	1.0 (0.7-1.3)	0.7
Total	602 (100)	360 (100)		

** Adjusted for 'important' cerebrovascular risk factors as determined in Appendix IV*

When all potential confounding variables are considered, no statistical association between ischemic stroke and each of the SNP's was observed (Table 5.11).

Table 5.11
Multivariate Analysis of SNP's (by genotype) Between Controls and Cases Adjusted for all Potential Confounders

Genotype	Controls n (%)	Cases n (%)	Odds Ratio* (95%CI)	p
PON 1 –107T/C				
TT	61 (20)	44 (24)	1.0	
TC	200 (66)	97 (54)	0.7 (0.4-1.2)	0.2
CC	40 (13)	39 (22)	1.6 (0.8-3.1)	0.2
PON 1 M54L				
MM	36 (12)	20 (11)	1.0	
ML	138 (46)	82 (46)	1.7 (0.8-3.4)	0.2
LL	127 (42)	78 (43)	1.4 (0.7-2.8)	0.4
Gp 1b HPA 2a/b				
aa	256 (85)	158 (88)	1.0	
ab	43 (14)	22 (12)	0.8 (0.4-1.4)	0.4
bb	2 (1)	0 (0)	-	-
Gp IIb/IIIa PlA1/A2				
A1A1	213 (71)	116 (64)	1.0	
A1A2	81 (27)	61 (34)	1.4 (0.9-2.2)	0.1
A2A2	7 (2)	3 (2)	0.5 (0.1-2.8)	0.5
β Fib –148C/T				
CC	190 (63)	114 (63)	1.0	
CT	102 (34)	58 (32)	0.95 (0.6-1.5)	0.8
TT	9 (3)	8 (4)	1.9 (0.7-5.4)	0.2
Prothromb 20210G/A				
GG	293 (97)	174 (97)	1.0	
GA	8 (3)	6 (3)	1.7 (0.5-5.4)	0.4
AA	0 (0)	0 (0)	-	-
TPA –7351C/T				
CC	137 (46)	74 (41)	1.0	
CT	136 (45)	83 (46)	1.2 (0.8-1.9)	0.4
TT	28 (9)	23 (13)	1.8 (0.9-3.4)	0.1
PAI-1 5G/4G				
5G5G	61 (20)	42 (23)	1.0	
5G4G	158 (52)	91 (51)	1.0 (0.6-1.8)	0.9
4G4G	82 (27)	47 (26)	1.0 (0.6-1.9)	0.9
Total	301 (100)	180 (100)		

**Adjusted for age, gender, ethnicity, smoking, history of hypertension, history of hypercholesterolemia, history of diabetes, atrial fibrillation and family history of stroke.*

In particular, the TPA –7351 TT genotype was not associated with an increased risk of ischemic stroke. Stratification by allelic status did not alter the outcome of the above

findings. When all potential confounding variables were considered, the TPA –7351 T allele was not associated with an increase in risk of ischemic stroke (Table 5.12).

Table 5.12
Multivariate Analysis of SNP's (by allele) between Controls and Cases Adjusted for all Potential Confounders

Allele	Controls n (%)	Cases n (%)	Odds Ratio* (95%CI)	p
PON 1 –107T/C				
T	322 (53)	185 (51)	1.0	
C	280 (47)	175 (49)	1.2 (0.9-1.5)	0.3
PON 1 M54L				
M	210 (35)	122 (34)	1.0	
L	392 (65)	238 (66)	1.1 (0.8-1.4)	0.7
Gp 1b HPA 2a/b				
a	555 (92)	338 (94)	1.0	
b	47 (8)	22 (6)	0.7 (0.4-1.2)	0.2
Gp IIb/IIIa PlA1/A2				
A1	507 (84)	293 (81)	1.0	
A2	95 (16)	67 (19)	1.2 (0.8-1.8)	0.3
β Fib –148C/T				
C	482 (80)	286 (79)	1.0	
T	120 (20)	74 (21)	1.1 (0.8-1.6)	0.6
Prothromb 20210G/A				
G	594 (99)	354 (99)	1.0	
A	8 (1)	6 (2)	1.6 (0.5-5.2)	0.4
TPA –7351C/T				
C	410 (68)	231 (64)	1.0	
T	192 (32)	129 (36)	1.3 (0.9-1.7)	0.1
PAI-1 5G/4G				
5G	280 (47)	175 (49)	1.0	
4G	322 (53)	185 (51)	1.0 (0.8-1.4)	0.9
Total	602 (100)	360 (100)		

**Adjusted for age, gender, ethnicity, smoking, history of hypertension, history of hypercholesterolemia, history of diabetes, atrial fibrillation and family history of stroke.*

5.7 Subgroup Analysis: Lacunar Stroke

Table 5.13 illustrates the risk of lacunar stroke associated with each SNP. 44 cases (24%) were classified with lacunar stroke. The risk of lacunar stroke associated with the glycoprotein 1b HPA2 b/b and prothrombin 20210 GA or AA genotypes could not be determined as there were no lacunar stroke cases with this genotype. On univariate analysis, the TPA –7351 TT genotype was associated with a 2.6-fold increase in risk of lacunar stroke (OR 2.6, 95%CI 1.1-6.4).

The corresponding attributable risk of lacunar stroke associated with this genotype was 13% (range 2%-17%) (table 5.14). A higher prevalence of the PON1 –107CC genotype also occurred in the lacunar stroke cohort compared to controls, however, the difference only just achieved statistical significance (OR 2.4, 95%CI 1.01-5.9). The corresponding attributable risk of lacunar stroke associated with the PON1 –107CC genotype was 21% (0%-29%) (table 5.14). The remaining SNP's were not associated with an increased risk of lacunar stroke.

Table 5.13
Lacunar Stroke: Univariate Analysis of SNP's Between Controls and Cases

Genotype	Controls n (%)	Lacunar Stroke n (%)	Odds Ratio (95%CI)	p
PON 1 –107T/C				
TT	61 (20)	10 (23)	1.0	
TC	200 (66)	18 (41)	0.6 (0.2-1.3)	0.2
CC	40 (13)	16 (36)	2.4 (1.01-5.9)	0.05
PON 1 M54L				
MM	36 (12)	3 (7)	1.0	
ML	138 (46)	25 (57)	2.2 (0.6-7.6)	0.2
LL	127 (42)	16 (36)	1.5 (0.4-5.5)	0.5
Gp 1b HPA 2a/b				
aa	256 (85)	40 (91)	1.0	
ab	43 (14)	4 (9)	0.6 (0.2-1.8)	0.3
bb	2 (1)	0 (0)	-	-
Gp IIb/IIIa PlA1/A2				
A1A1	213 (71)	33 (75)	1.0	
A1A2	81 (27)	10 (23)	0.8 (0.4-1.7)	0.6
A2A2	7 (2)	1 (2)	0.9 (0.1-7.7)	0.9
β Fib –148C/T				
CC	190 (63)	26 (59)	1.0	
CT	102 (34)	15 (34)	1.1 (0.5-2.1)	0.8
TT	9 (3)	3 (7)	2.4 (0.6-9.6)	0.2
Prothrombin 20210G/A				
GG	293 (97)	44 (100)	1.0	
GA	8 (3)	0 (0)	-	-
AA	0 (0)	0 (0)	-	-
TPA –7351C/T				
CC	137 (46)	17 (39)	1.0	
CT	136 (45)	18 (41)	1.1 (0.5-2.2)	0.9
TT	28 (9)	9 (20)	2.7 (1.1-6.4)	0.04
PAI-1 5G/4G				
5G5G	61 (20)	12 (27)	1.0	
5G4G	158 (52)	20 (45)	0.6 (0.3-1.4)	0.3
4G4G	82 (27)	12 (27)	0.7 (0.3-1.8)	0.5
Total	301 (100)	44 (100)		

Table 5.14
Attributable Risk of Lacunar Stroke Associated With TPA –7351TT and PON1–107CC Genotypes

	Parameter	LS Estimate	Lower Limit	Upper Limit
TPA –7351 TT				
Relative risk estimate	OR	2.69	1.08	6.69
Proportion cases exposed	P	0.20	0.20	0.20
Attributable Risk	P(OR-1)/OR	0.13	0.02	0.17
PON1 –107 CC				
Relative risk estimate	OR	2.44	1.01	5.91
Proportion cases exposed	P	0.36	0.36	0.36
Attributable Risk	P(OR-1)/OR	0.21	0.00	0.29

Table 5.15 shows the risk of non-lacunar stroke associated with each SNP (stratified for genotype). Non-lacunar stroke was defined as cases classified with TACS, PACS or PCS. 136 cases (76%) were classified with non-lacunar stroke. No association was observed between non-lacunar stroke and each of the SNP's.

Table 5.15
Non-Lacunar Stroke: Univariate Analysis of SNP's Between Controls and Cases

Genotype	Controls n (%)	Non-Lacunar stroke* n (%)	Odds Ratio (95%CI)	p
PON 1 –107T/C				
TT	61 (20)	34 (25)	1.0	
TC	200 (66)	79 (58)	0.7 (0.4-1.2)	0.2
CC	40 (13)	23 (17)	1.0 (0.5-2.0)	0.9
PON 1 M54L				
MM	36 (12)	17 (13)	1.0	
ML	138 (46)	57 (42)	0.9 (0.5-1.7)	0.7
LL	127 (42)	62 (46)	1.0 (0.5-2.0)	0.9
Gp 1b HPA 2a/b				
aa	256 (85)	118 (87)	1.0	
ab	43 (14)	18 (13)	0.9 (0.5-1.6)	0.8
bb	2 (1)	0 (0)	-	-
Gp IIb/IIIa PlA1/A2				
A1A1	213 (71)	83 (61)	1.0	
A1A2	81 (27)	51 (38)	1.6 (1.0-2.5)	0.06
A2A2	7 (2)	2 (1)	0.7 (0.2-3.6)	0.7
β Fib –148C/T				
CC	190 (63)	88 (65)	1.0	
CT	102 (34)	43 (32)	0.9 (0.6-1.4)	0.7
TT	9 (3)	5 (4)	1.2 (0.4-3.7)	0.8
Prothromb 20210G/A				
GG	293 (97)	130 (96)	1.0	
GA	8 (3)	6 (4)	1.7 (0.6-5.0)	0.3
AA	0 (0)	0 (0)	-	-
TPA –7351C/T				
CC	137 (46)	57 (42)	1.0	
CT	136 (45)	65 (48)	1.2 (0.8-1.8)	0.5
TT	28 (9)	14 (10)	1.2 (0.6-2.5)	0.6
PAI-1 5G/4G				
5G5G	61 (20)	30 (22)	1.0	
5G4G	158 (52)	71 (52)	0.9 (0.5-1.5)	0.7
4G4G	82 (27)	35 (26)	0.9 (0.5-1.6)	0.6
Total	301 (100)	136 (100)		

5.8 Subgroup Analysis: Cardioembolic Stroke

Table 5.16 illustrates the risk of cardioembolic stroke associated with each SNP. A cardioembolic source of ischemic stroke occurred in 24 (13%) cases. On univariate analysis, no significant genetic association predisposing to cardioembolic stroke was found. A similar finding was observed for cases coded as non-cardioembolic stroke (N=156), (table 5.17).

Table 5.16
Cardioembolic (CE) Stroke: Univariate Analysis of SNP's Between Controls and Cases

Genotype	Controls n (%)	CE Stroke n (%)	Odds Ratio (95%CI)	p
PON 1 –107T/C				
TT	61 (20)	6 (25)	1.0	
TC	200 (66)	13 (54)	0.7 (0.2-1.8)	0.4
CC	40 (13)	5 (21)	1.3 (0.4-4.4)	0.7
PON 1 M54L				
MM	36 (12)	4 (17)	1.0	
ML	138 (46)	10 (42)	0.7 (0.2-2.2)	0.5
LL	127 (42)	10 (42)	0.7 (0.2-2.4)	0.6
Gp 1b HPA 2a/b				
aa	256 (85)	20 (83)	1.0	
ab	43 (14)	4 (17)	1.2 (0.4-3.7)	0.8
bb	2 (1)	0 (0)	-	-
Gp IIb/IIIa PlA1/A2				
A1A1	213 (71)	14 (58)	1.0	
A1A2	81 (27)	10 (42)	1.9 (0.8-4.4)	0.1
A2A2	7 (2)	0 (0)	-	-
β Fib –148C/T				
CC	190 (63)	21 (88)	1.0	
CT	102 (34)	3 (13)	0.3 (0.1-1.0)	0.05
TT	9 (3)	0 (0)	-	-
Prothromb 20210G/A				
GG	293 (97)	24 (100)	1.0	
GA	8 (3)	0 (0)	-	-
AA	0 (0)	0 (0)	-	-
TPA –7351C/T				
CC	137 (46)	11 (46)	1.0	
CT	136 (45)	13 (54)	1.2 (0.5-2.8)	0.7
TT	28 (9)	0 (0)	-	-
PAI-1 5G/4G				
5G5G	61 (20)	7 (29)	1.0	
5G4G	158 (52)	9 (38)	0.5 (0.2-1.4)	0.2
4G4G	82 (27)	8 (33)	0.9 (0.3-2.5)	0.8
Total	301 (100)	24 (100)		

Table 5.17
Non-Cardioembolic Stroke: Univariate Analysis of SNP's Between Controls and Cases

Genotype	Controls n (%)	Non-CE Stroke n (%)	Odds Ratio (95%CI)	p
PON 1 –107T/C				
TT	61 (20)	38 (24)	1.0	
TC	200 (66)	84 (54)	0.7 (0.4-1.1)	0.1
CC	40 (13)	34 (22)	1.4 (0.7-2.5)	0.3
PON 1 M54L				
MM	36 (12)	16 (10)	1.0	
ML	138 (46)	72 (46)	1.2 (0.6-2.3)	0.6
LL	127 (42)	68 (44)	1.2 (0.6-2.3)	0.6
Gp 1b HPA 2a/b				
aa	256 (85)	138 (88)	1.0	
ab	43 (14)	18 (12)	0.8 (0.4-1.4)	0.4
bb	2 (1)	0 (0)	-	-
Gp IIb/IIIa PlA1/A2				
A1A1	213 (71)	102 (65)	1.0	
A1A2	81 (27)	51 (33)	1.3 (0.9-2.0)	0.2
A2A2	7 (2)	3 (2)	0.9 (0.2-3.5)	0.9
β Fib –148C/T				
CC	190 (63)	93 (60)	1.0	
CT	102 (34)	55 (35)	1.1 (0.7-1.7)	0.6
TT	9 (3)	8 (5)	1.8 (0.7-4.9)	0.2
Prothromb 20210G/A				
GG	293 (97)	150 (96)	1.0	
GA	8 (3)	6 (4)	1.5 (0.5-4.3)	0.5
AA	0 (0)	0 (0)	-	-
TPA –7351C/T				
CC	137 (46)	63 (40)	1.0	
CT	136 (45)	70 (45)	1.1 (0.7-1.7)	0.6
TT	28 (9)	23 (15)	1.8 (1.0-3.3)	0.07
PAI-1 5G/4G				
5G5G	61 (20)	35 (22)	1.0	
5G4G	158 (52)	82 (53)	0.9 (0.6-1.5)	0.7
4G4G	82 (27)	39 (25)	0.8 (0.5-1.5)	0.5
Total	301 (100)	156 (100)		

5.9 Gene-Risk Factor Interactions

A total of 56 separate logistic regression models (8 SNP's by 7 risk factors) assessed whether a genotype effect is more pronounced in the presence of other risk factors. Interaction variables contained the pathogenic genotype of the relevant SNP (homozygous state for the polymorphic allele except for prothrombin 20210G/A and glycoprotein 1b HPA 2a/b SNP's where the heterozygous state was used) and each of the ischemic stroke risk factors (as defined in table 5.3). An interaction variable for each SNP and the presence of multiple risk factors (possession of one to all six risk factors) was also created. A further 56 logistic regression models evaluated the effect of interaction between the pathogenic allele of each SNP and the known cerebrovascular risk factors defined in table 5.3. The p values determined by each model are presented in table 5.18. A p value of <0.05 occurred for the following interactions:

β Fibrinogen –148TT genotype and hypercholesterolemia (p=0.001)
β Fibrinogen –148 T allele and hypercholesterolemia (p=0.03)
β Fibrinogen –148C/T and multiple risk factors (p=0.03)
Glycoprotein 1b HPA 2a/b genotype and hypercholesterolemia (p=0.04)
Glycoprotein IIb/IIIa PlA2/A2 genotype and hypercholesterolemia (p=0.01)
Glycoprotein IIb/IIIa PlA2/A2 genotype and diabetes (p=0.04)

The number of significant findings (6) is approximately what would be expected by chance alone, and therefore should be considered with caution. Furthermore, the corresponding odds ratios associated with these interaction variables were very small and reflect an inadequate sample size and study power to perform an interaction analysis. In conclusion, no meaningful gene-environment interaction was observed.

Table 5.18
Probability Values for Gene-Risk Factor Interactions

	PON1107C/T	PON154M/L	Gp1bHPA2a/b	GpPlA1/A2
Smoking				
Genotype	0.1	0.9	0.6	1.0
Allele	0.6	0.9	0.3	0.9
Hypertension				
Genotype	0.5	0.9	0.3	0.2
Allele	0.5	0.9	0.5	0.3
Hypercholesterolemia				
Genotype	0.3	0.8	0.04	0.01
Allele	0.9	0.8	0.07	0.06
Diabetes				
Genotype	0.3	0.1	0.3	0.04
Allele	0.9	1.0	0.2	0.2
Family History				
Genotype	0.8	0.4	0.6	1.0
Allele	0.5	0.6	0.7	0.9
Atrial Fibrilation				
Genotype	0.05	0.5	0.9	0.06
Allele	0.2	0.2	0.8	0.8
Multiple Risk Factors				
Genotype	0.2	0.3	0.6	0.2
Allele	0.6	0.3	0.6	0.1

Table 5.18
Probability Values for Gene-Risk Factor Interactions (cont.)

	βFib-148C/T	Pro.20210A	TPA -7351C/T	PAI 5G/4G
Smoking				
Genotype	0.3	-	0.1	0.2
Allele	0.1	-	0.4	0.1
Hypertension				
Genotype	0.3	0.6	0.9	0.4
Allele	0.1	0.6	0.9	0.2
Hypercholesterolemia				
Genotype	0.001	0.8	0.5	0.9
Allele	0.03	0.8	0.6	0.7
Diabetes				
Genotype	0.05	0.5	0.2	0.9
Allele	0.4	0.5	0.07	0.8
Family History				
Genotype	0.3	0.07	0.3	0.9
Allele	0.2	0.1	0.2	1.0
Atrial Fibrillation				
Genotype	0.5	-	0.09	0.6
Allele	0.6	-	0.2	0.5
Multiple Risk Factors				
Genotype	0.03	0.3	1.0	1.0
Allele	0.2	0.4	0.7	1.0

5.10 Gene-Gene Interactions

As a significant association between the TPA –7351 TT genotype and ischemic stroke was shown previously, logistic regression was used to assess whether this risk was more pronounced in the presence of any one of the other SNP's. The results from 7 logistic regression models each containing the TPA –7351 TT genotype are presented in table 5.19. The effect of interaction between the TPA –7351 TT genotype and the glycoprotein IIb/IIIa Pl A2A2 and β Fibrinogen –148 TT genotypes could not be determined, as this haplotype combination did not occur in our study sample. No significant interaction with the TPA –7351 TT genotype was observed. Importantly,

the interaction with the PAI 5G/4G, a SNP also influencing fibrinolysis, was not statistically significant (OR 1.1, 95%CI 0.2-5.3).

Table 5.19
Probability Values for Gene-Gene Interactions

Genotype	OR (95%CI)	p
PON1 –107 CC	1.7 (1.0-2.9)	0.07
TPA –7351 TT	1.6 (0.8-3.2)	0.2
Interaction (CC/TT)	1.1 (0.3-4.7)	0.9
PON1 54LL	0.8 (0.5-1.3)	0.3
TPA –7351 TT	1.1 (0.5-2.7)	0.8
Interaction (LL/TT)	2.8 (0.8-10.7)	0.1
Gp 1b HPA2 ab	0.7 (0.4-1.4)	0.3
TPA –7351 TT	1.7 (0.9-3.3)	0.1
Interaction (ab/TT)	1.0 (0.2-5.6)	1.0
Gp 1Ib/IIIa Pl A2A2	-	-
TPA –7351 TT	-	-
Interaction (A2A2/TT)	-	-
β Fib –148 TT	-	-
TPA –7351 TT	-	-
Interaction (TT/TT)	-	-
Prothrombin 20210 GA	0.7 (0.2-3.0)	0.6
TPA –7351 TT	1.4 (0.7-2.7)	0.4
Interaction (GA/TT)	5.7 (0.3-114.0)	0.3
PAI 4G4G	0.9 (0.5-1.4)	0.6
TPA –7351 TT	1.7 (0.9-3.5)	0.1
Interaction (4G4G/TT)	1.1 (0.2-5.3)	0.9

5.11 Results Summary

This study population consisted of 182 acute ischemic stroke patients and 301 age and gender-matched community controls. There was no significant difference between the two groups in terms of demographic factors, a family history of stroke or use of stroke preventing medication. Atrial fibrillation, smoking and diabetes mellitus were shown

to be independent risk factors for ischemic stroke. Although a protective effect for ischemic stroke was apparent for hypertension and hypercholesterolemia, this was reversed when a history alone for these risk factors was considered.

In a univariate analysis, an association between ischemic stroke and each of the SNP's was not apparent. However, after adjustment for known cerebrovascular risk factors, the TPA –7351TT genotype was significantly associated with ischemic stroke (OR 1.9 95%CI 1.0-3.6, p=0.049). This association persisted at a 10% significance level when all potential confounding influences were considered. Stratification for stroke subtype showed a significant association with lacunar but not non-lacunar stroke (adjusted OR 2.6 95%CI 1.1-6.4, p=0.04). The corresponding attributable risk of lacunar stroke associated with the TPA –7351 TT genotype was 13% (2%-17%). Although no statistical association was observed for the PON1 -107CC genotype and ischemic stroke on univariate or multivariate analysis (adjustment for both important and all potential confounders), stratification for stroke sub-type showed a 2.4-fold increased risk of lacunar stroke. The 95% confidence intervals were, however, wide and approached unity. No other significant association was observed for the entire stroke cohort or following stratification for stroke subtype or source of occluding thrombus.

Similarly, analysis for gene-risk factor and gene-gene interactions did not reveal any significant findings beyond that expected by chance alone.

Chapter 6

Discussion

6.1 Introduction

This chapter will approach the study findings in two ways. First, the results will be interpreted from a functional perspective, whereby a biologically plausible explanation for a genetic association with ischemic stroke (or lack of) will be addressed. This approach will endeavour to incorporate the current study findings with data from previous laboratory and epidemiological studies.

Second, the results of this study will be interpreted from an epidemiological and statistical perspective. In this section, the discussion will focus on the study design, with the significance of both positive and negative associations being addressed following consideration of the study limitations that were encountered.

6.2 TPA –7351 C/T SNP and Ischemic Stroke

In this study, a novel association between the TPA –7351 SNP and ischemic stroke was shown. When the entire stroke cohort was considered, the risk of ischemic stroke associated with the homozygous genotype (TT) nearly doubled. A gene-dose effect for the T allele was also apparent, with a higher prevalence of the heterozygous genotype (CT) in the stroke subgroup. However, this was not statistically significant. In a subgroup analysis stratified for stroke subtype, a 2.7-fold increased risk of lacunar stroke was observed. This finding corresponded to an attributable risk of 13%. In contrast, no association between the TPA –7351 SNP and ischemic stroke of large vessel or cardioembolic aetiology was observed. The findings suggest that impaired fibrinolysis may play an important role in the pathogenesis of lacunar stroke.

Although the pathogenesis of lacunar stroke remains controversial (Millikan and Futrell, 1990), two distinct small vessel pathologies underlie most cases: lipohyalinosis and micro-atheroma (Fisher, 1979). This constituted the 'Lacunar Hypothesis' that has been subsequently validated. In a prospective study of the Northern Manhattan Stroke cohort, lipohyalinosis or micro-atheroma were the most common causes of lacunar stroke, with cardio-embolism and other rare aetiologies accounting for only 5% and 2% respectively (Gan *et al.*, 1997). Both small vessel lipohyalinosis and micro-atheroma lead to slow luminal narrowing causing hemodynamic compromise that may initially present with the Capsular Warning Syndrome (Donnan *et al.*, 1993). Subsequent lacunar infarction occurs following non-thrombotic obliteration of diseased small vessels or from occlusive thrombus (Fisher, 1979).

The acute release of TPA by vascular endothelium is compromised by atherosclerosis, with the reduction being proportional to the degree of atherosclerotic burden (Newby *et al.*, 1999). Therefore, factors that further diminish fibrinolytic potential may be critical in predisposing to thrombotic occlusion. The results of this study are in keeping with this hypothesis, that in the presence of small vessel disease, possession of the TPA –7351TT genotype significantly increases the risk of thrombotic occlusion and presentation with lacunar stroke.

The findings also support the hypothesis that factors influencing fibrinolysis may play a more important role in the pathogenesis of lacunar stroke than in ischemic stroke associated with large vessel disease. This is further supported by a functional study of fibrinolytic potential in patients with ischemic stroke (Kilpatrick *et al.*, 1993). In this study, the euglobulin lysis time, a measure of fibrinolytic activity, was impaired in a larger proportion of patients with lacunar stroke compared to those with cortical infarction secondary to large vessel disease (67% vs 20%). Furthermore, thrombin activity, as determined by plasma fibrinopeptide A levels, was increased in patients with lacunar stroke, suggesting that the pathogenesis of small vessel occlusion involves an interaction of pro-thrombotic and hypo-fibrinolytic forces favouring thrombus formation.

Immunohistochemical studies of primate and rat brains provide further evidence supporting the importance of fibrinolysis in small vessel disease (Levin and del Zoppo, 1994, Schreiber *et al.*, 1998). Under normal circumstances, in rat and primate brains, TPA expression is limited to the endothelial cells of pre-capillary arterioles and post-capillary venules, with no expression by the endothelium of large arteries. The former observation suggests that these vessels rely on a continuous release of TPA to maintain vascular patency and may be particularly vulnerable to thrombotic occlusion by factors that impair fibrinolytic potential. To date, the equivalent histological study in human brain has not been reported.

Recent studies have shown that cigarette smoking significantly reduces the acute release of TPA by vascular endothelium (Newby *et al.*, 1999). Although the molecular mechanisms remain unclear, the association has been implicated in the pathogenesis of coronary artery disease and would explain why TPA thrombolytic therapy for myocardial infarction is more effective in smokers (Newby *et al.*, 1999). In the current study, an objective measurement of acute TPA release was not performed, therefore, the relative impact of smoking and the TPA –7351C/T SNP could not be determined. Cigarette smoking, however, did not confound the risk of ischemic stroke associated with the TPA –7351C/T SNP and no interaction with the TPA –7351TT genotype was observed.

The findings are consistent with those of a recent study showing an increased risk of ischemic stroke associated with the D allele of the TPA Alu-repeat insertion/deletion SNP (Austin *et al.*, 2002). This study, however, did not stratify for stroke subtype and was performed on a younger stroke population. The TPA Alu-repeat insertion/deletion SNP is in tight linkage disequilibrium with the TPA –7351C/T SNP, with a pairwise maximum likelihood of linkage disequilibrium estimate of 0.71 (Ladenvall *et al.*, 2000). Functional significance, however, is greatest for the TPA –7351C/T polymorphism which accounts for 18% of the phenotypic variance in TPA release rates, a finding that is almost double that associated with the TPA Alu-repeat insertion/deletion polymorphism (10%) (Ladenvall *et al.*, 2001). It is therefore possible that the positive association observed between the TPA Alu-repeat insertion/deletion

polymorphism and ischemic stroke occurred through linkage with the functional TPA –7351C/T locus.

Functional significance for the TPA –7351C/T SNP has also been shown at a molecular level. The rate-limiting step governing the transcription of the TPA gene occurs at the initiation of transcription. This process involves the assembly of a pre-initiation complex, of which transcription factor IID (TFIID) is an important constituent (Kooistra *et al.*, 1994). For many genes, TFIID complex binds directly to DNA via a 'TATA box' that is located adjacent to the transcription start site. Following DNA binding, the TFIID complex then interacts with RNA polymerase II allowing the commencement of DNA transcription. For TPA however, TFIID complexes with another transcription factor, Sp1, which then binds to DNA at a region rich in guanine and cytosine, known as the GC-box I (Kooistra *et al.*, 1994). The functional significance of the TPA –7351 C/T SNP occurs because it is located within this GC-box. Possession of the T allele at position -7351 results in a reduced binding affinity for Sp1, which has a dominant effect on TPA release rate. Consequently, both CT and TT genotypes result in significantly less DNA transcription corresponding to less than half the TPA release rate than that observed with the CC genotype (Ladenvall *et al.*, 2000).

High plasma TPA levels have been associated with ischemic stroke in both cross-sectional (Johansson *et al.*, 2000, Kain *et al.*, 2001, Margaglione *et al.*, 1994, Lindgren *et al.*, 1996, Kristensen *et al.*, 1999, Macko *et al.*, 1999) and prospective cohort studies (Ridker, 1994, Smith *et al.*, 1998). This contrasts with the findings of this study of an association between low endothelial TPA release rates and ischemic stroke. There are two possible explanations for the difference. Firstly, TPA secretion by vascular endothelium occurs via two pathways: regulated and constitutional secretion (Kooistra *et al.*, 1994). In regulated secretion, newly synthesized TPA is stored in intracellular storage granules and is rapidly released only under stimulated conditions, such as exposure to thrombin (Kooistra *et al.*, 1994). The rapid release of TPA provides a means of achieving a high intraluminal TPA concentration that is critical for successful local fibrinolysis.

In constitutive secretion, TPA is continuously produced by the Golgi apparatus and transported to the cell surface via specific transport vesicles. The process provides a low-grade secretion of TPA in amounts insufficient to mediate fibrinolysis. This pathway, however, is the primary determinant of plasma TPA levels (Kooistra *et al.*, 1994). Importantly, basal TPA secretion is higher in atherosclerotic arteries than in normal blood vessels (Underwood and De Bono, 1993). Therefore, higher systemic TPA levels may reflect a higher 'atherosclerotic burden' and co-exist with factors that impair local TPA release. This is further illustrated in a recent case-control study investigating the relevance of the TPA-7351C/T SNP in patients with myocardial infarction (Ladenvall *et al.*, 2002). Although there was no significant difference in the plasma level of TPA between cases and controls stratified for TPA-7351C/T genotype, the greatest risk of myocardial infarction was observed in those with both the T allele (decreased TPA activity) and a high plasma TPA level (high atherosclerotic burden), implicating that both these variables had independent prognostic significance.

Alternatively, this discrepancy can be explained by the kinetics and metabolism of TPA. Once released into the blood stream, TPA is very rapidly inactivated by PAI-1 and circulates in the blood in this complex form. As the hepatic clearance of TPA/PAI-1 complex occurs at a much slower rate than that of free TPA, the total TPA level is largely determined by the plasma level of PAI. In the above-mentioned epidemiological studies, TPA was measured in the form bound to PAI-1, therefore, the higher TPA levels associated with ischemic stroke may reflect high PAI-1 levels rather than higher rates of TPA secretion.

In the current study, patients with hypertensive-type intracerebral haemorrhage were not investigated. However, the finding of a genetic predisposition for lacunar stroke may have important implications for the former stroke subtype. Both lacunar stroke and hypertensive-type intracerebral haemorrhage share the same vascular pathology, namely fibrinoid necrosis and lipohyalinosis (Masuda *et al.*, 1983). What determines the clinical phenotype of this vascular pathology is unknown, as both are associated with similar cerebrovascular risk factor profiles. The results of the current study suggest that the TPA-7351C/T SNP may be an important determinant of the

phenotypic expression of lipohyalinosis, with possession of the TT genotype favouring thrombosis and presentation with lacunar infarction. Whether the TPA–7351CC genotype, which corresponds to higher TPA release rates, predisposes to subcortical intracerebral haemorrhage in the presence of lipohyalinosis, in unknown and warrants further investigation.

6.3 PAI-1 5G/4G SNP and Ischemic Stroke

In this study, an association between the PAI-1 5G/4G SNP and ischemic stroke was not apparent. The lack of association persisted following stratification for stroke subtype and thrombotic source. Importantly, as PAI-1 is the primary inhibitor of TPA, the risk of ischemic stroke was not significantly increased in those with both the PAI-1 4G/4G and TPA-7351 TT genotypes. The results suggest that the PAI-1 5G/4G SNP does not play an important role in the pathogenesis of ischemic stroke.

This finding is not in keeping with the strong evidence showing that the PAI-1 5G/4G SNP plays an important role in the regulation of PAI-1 transcription and plasma levels (Eriksson *et al.*, 1995). The SNP is located within the promotor region of the PAI-1 gene, with both alleles contributing to DNA binding sites capable of binding to transcriptional activators. The PAI-1 5G allele, however, additionally binds a transcriptional repressor (Eriksson *et al.*, 1995). Consequently, the extent of PAI-1 transcription depends on the number of 4G alleles present, with the PAI-1 4G/4G genotype associated with a 6-fold increase in PAI-1 mRNA expression (Dawson *et al.*, 1993) and a doubling of PAI-1 activity compared to its wild-type counterpart (5G/5G) (Eriksson *et al.*, 1995). Therefore, the possession of the PAI-1 4G/4G genotype (corresponding to higher PAI-1 expression) would be expected to augment the inhibition of TPA and predispose to ischemic stroke, a hypothesis that did not hold true in this study.

The lack of association between ischemic stroke and the PAI-1 5G/4G SNP may be explained by the contrasting effects of PAI-1 mediated TPA inhibition on fibrinolysis and atherosclerotic plaque stability. The initiation of atherosclerotic plaque rupture is well described and involves the degradation of the plaque's fibrous cap, a process

mediated primarily by activated metalloproteinases (MMP's) (Shah *et al.*, 1995). MMP's are activated by plasmin, which is itself generated from the activation of plasminogen by unbound, active TPA. An increase in PAI-1 activity associated with possession of the PAI-1 4G allele may therefore limit TPA activation and ultimately reduce MMP mediated fibrous cap degradation (Roest *et al.*, 2000). In keeping with this hypothesis, two epidemiological studies have found a protective effect for ischemic stroke associated with the PAI-1 4G allele (Endler *et al.*, 2000, Roest *et al.*, 2000). In the largest of these involving 12,239 women aged 52 to 67 years, who were followed for up to 18 years, the relative risk of death caused by stroke associated with the PAI-1 4G/4G genotype was 0.4 (95%CI 0.2-0.7) (Roest *et al.*, 2000). The study, however, did not routinely perform brain neuro-imaging and therefore the relevance of this finding to either ischemic or haemorrhagic stroke could not be determined. In the current study, the pro-thrombotic influence of the PAI-4G allele may have been offset by its stabilising effect on atherosclerotic plaque, resulting in a negative association between PAI-1 5G/4G genotype and ischemic stroke.

Although the plasma level of PAI is significantly influenced by the PAI-1 5G/4G SNP (Dawson *et al.*, 1993), other factors including age, gender, lipid profile, body mass index, smoking and alcohol consumption have also been shown to be important determinants (Margaglione *et al.*, 1998). In the current study, atrial fibrillation and diastolic blood pressure were the only important confounders determined on bivariate analysis. There was no apparent confounding effect due to age, gender, smoking or plasma cholesterol level (see section 5.4). Plasma PAI levels, body mass index and an estimation of alcohol consumption, however, were not determined thus an adjustment for these possible confounding influences could not be performed. It therefore remains possible that a type 2 error may have occurred.

6.4 PON1 SNP's and Ischemic Stroke

Human serum PON1 is a calcium dependent esterase that hydrolyses organophosphates such as paraoxon, sarin and phenyl acetate (Aviram *et al.*, 1998). The virtual absence of this enzyme in insects and relative deficiency in birds and fish is responsible for the susceptibility to pesticides seen in these animals (Mackness and

Durrington, 1995). In humans, PON1 has also been implicated in the maintenance of arterial wall integrity by hydrolysing biologically active lipids in LDL before they accumulate and form atherogenic LDL particles (Watson *et al.*, 1995). The regulation of paroxonase gene expression is mostly influenced by the PON1 -107T/C SNP (Leviev and James, 2000, Brophy *et al.*, 2001). The SNP occurs within promoter region of the PON1 gene and forms a binding site for the transcriptional activator, Sp1. Possession of the T allele causes a lower affinity for Sp1 binding corresponding to less transcriptional activation. Consequently, both the PON mass concentration and enzymatic activity are approximately half of that associated with the C allele. Therefore, the PON1 -107CC genotype, which is associated with a higher PON1 activity, should theoretically protect against ischemic stroke.

In contrast, the present study found a 2.4-fold increase in risk of lacunar stroke associated with the PON1 -107CC genotype (OR 2.4, 95%CI 1.01-5.9, p=0.048). A plausible explanation for this contradiction can be extrapolated from studies of the PON1 54 M/L SNP, a SNP that is in strong linkage disequilibrium with the PON1 -107C/T SNP (Brophy *et al.*, 2001). Although the PON1 54LL and -107CC genotypes (genotypes which are in strong linkage disequilibrium with each other) are associated with the highest level of PON1 expression and paraoxon hydrolytic activity (Brophy *et al.*, 2001), they in contrast, provide the least protection to LDL against oxidative modification (Mackness *et al.*, 1998). The underlying mechanism causing this inverse relationship is poorly understood but research suggests that the mechanism providing protection against LDL oxidative modification is independent of that associated with paraoxon hydrolysis. Alternatively, the higher enzymatic activity and mass concentration associated with the PON1 54LL and -107CC genotypes may lead to the overproduction of toxic moieties that may facilitate LDL oxidation. In keeping with this hypothesis, paraoxonase has been implicated in the production of harmful lysolipids (eg lysophosphatidylcholine - lysolethecine), which have been shown to facilitate the oxidative modification of LDL (Mackness and Durrington, 1995). Furthermore, in an immunohistochemical study of atherosclerotic arteries, PON1 immunoreactivity was prominent in atherosclerotic lesions, with the intensity corresponding to the degree of disease severity (Mackness *et al.*, 1997).

In contrast to that observed with lacunar stroke, an association between the PON1 -107CC genotype and ischemic stroke caused by large vessel occlusion was not apparent. This finding suggests that PON1 may have specific relevance to the pathogenesis of small vessel disease. In keeping with this hypothesis, the PON1 54 LL genotype (which is in linkage disequilibrium with the PON1 -107CC genotype) has been associated with the progression of subcortical cerebral white matter lesions (Schmidt *et al.*, 2000) and diabetic retinopathy (Kao *et al.*, 1998). The reason underlying the predilection for small vessel disease is unclear. A plausible reason is that PON1 may predispose to both microatheroma and lipohyalinosis, thus exerting a stronger effect on the small-vessel-disease-related phenotypes. To date, the molecular pathophysiology of lipohyalinosis is unknown, however, vascular endothelial dysfunction causing disruption of the blood-brain barrier and leakage of toxic substances into the surrounding neural tissue has been hypothesized recently (Wardlaw *et al.*, 2003). Although this challenges the traditional notion of vascular occlusion by thrombus or hemodynamic compromise, common to both hypotheses is endothelial dysfunction. As PON1 is an important molecule regulating oxidative modification of small vessel endothelium, the finding of an association between a genetic determinant of PON1 gene expression and lacunar stroke is in keeping with the hypothesis of underlying endothelial dysfunction.

In the current study, no association was observed between PON1 54 M/L SNP and ischemic stroke. The lack of association persisted following stratification for stroke subtype and aetiology. This finding is consistent with the known genetic influences on PON1 transcription, with the greatest functional significance attributed to the PON1 107T/C SNP, which accounts for 22.8% of the variability of PON1 activity. The corresponding value for the PON1 54M/L SNP is 5% (Brophy *et al.*, 2001). Therefore, it is likely that the previously reported association between the PON1 54M/L SNP and small vessel cerebrovascular disease occurs through linkage with the more functionally significant PON1 -107T/C SNP. In keeping with this hypothesis, the PON1 54M/L and 107T/C SNP's have been shown to be in significant linkage disequilibrium (Brophy *et al.*, 2001). Although hypertension is strongly associated with the pathogenesis of lipohyalinosis and lacunar stroke, no interaction between this and the PON1 SNP's

was observed. Furthermore, in those with the pro-thrombotic PON1 -107CC genotype, the risk of ischemic stroke was not further increased in the presence of the TPA –7351TT genotype. The lack of interaction however should be interpreted with caution, as the absolute number of stroke cases in each subgroup was small (PON1 -107CC: n=40, TPA -7351TT: n=28).

6.5 β Fibrinogen -148C/T SNP and Ischemic Stroke

Hyperfibrinogenemia is an established risk factor for ischemic stroke (Qizilbash *et al.*, 1991, Resch *et al.*, 1992, Wu *et al.*, 1992, Kofoed *et al.*, 2003). Although the predominant mechanism of action is unknown, fibrinogen has been shown to influence mural clot susceptibility to fibrinolytic factors and promote atherosclerosis (Koenig, 2003). To date, eleven SNP's on the β fibrinogen gene have been identified. The greatest functional significance has been shown for the -148C/T SNP, which is located in the 5' promoter region of the β fibrinogen gene (Behague *et al.*, 1996).

In this study, the fibrinogen -148C/T SNP was not associated with an increased risk of ischemic stroke. This also held following stratification for stroke subtype, aetiology and assessment for gene-gene and gene-environment interactions. The most likely conclusion is that the β fibrinogen -148C/T SNP does not play an important role in the pathogenesis of ischemic stroke, a finding that is in keeping with that of the Physician's Health Study (Blake *et al.*, 2001).

The lack of association between ischemic stroke and the β fibrinogen -148C/T SNP contradicts its strong association with hyperfibrinogenemia. There are several possible reasons for this discrepancy. First, fibrinogen may be acting as a marker of some other process, such as inflammation, that arises secondary to ischemic stroke. In keeping with this suggestion, fibrinogen has been shown to be an acute phase reactant, with a significant rise in plasma level following infection, injury and trauma (Fey and Fuller, 1987). Alternatively, fibrinogen may play a role in the pathogenesis of ischemic stroke, but the effect of the fibrinogen -148C/T SNP is small when compared to other modifiers of plasma fibrinogen levels. Consequently, the minor changes in plasma fibrinogen levels attributed to the β fibrinogen -148C/T SNP are insufficient to

influence the clinical outcome. This hypothesis is consistent with the findings of the Physician's Health Study, that showed a lack of association between the β fibrinogen -148C/T SNP and ischemic stroke even though the plasma fibrinogen levels were significantly correlated with the β fibrinogen -148C/T SNP genotype. Further support to this hypothesis is provided by Humphries *et al.*, who concluded that genetic variation at the fibrinogen gene locus accounted for only 15% of the phenotypic variance of plasma fibrinogen levels. Other factors implicated in the determination of plasma fibrinogen levels include acute physiological stress, diabetes, increasing age, obesity and smoking (Humphries *et al.*, 1987). In the current study, fibrinogen levels were not determined and therefore the effect of the fibrinogen -148C/T SNP amongst the study participants is unknown.

Gene-environment interactions provided a third explanation for the discrepancy between the risk of ischemic stroke associated with fibrinogen and the β fibrinogen -148C/T SNP. Under this scenario, the effect of the β fibrinogen -148C/T SNP on both plasma fibrinogen levels and clinical phenotype becomes apparent only in the presence of some other factor. This phenomenon was illustrated by the ECTIM Study, a prospective case-control study of 565 patients with myocardial infarction and 668 controls (Behague *et al.*, 1996). In their population, a positive correlation between plasma fibrinogen level and the β fibrinogen -148C/T genotype was limited to those who smoked cigarettes. Cigarette smoking stimulates hepatic interleukin 6 synthesis (Thomas *et al.*, 1991), which inturn promotes fibrinogen transcription via an interleukin 6 responsive element in the promoter region of the fibrinogen gene. The β fibrinogen -148C/T SNP is located immediately adjacent to this region and plays an important role in the mediating the effect of interleukin 6 (Behague *et al.*, 1996). This provides a mechanism linking cigarette smoking with atherosclerosis and offers an explanation for the smoking/β fibrinogen -148TT genotype interaction on plasma fibrinogen levels. In the current study, fibrinogen levels were not determined and therefore this interaction could not be evaluated. However, no interaction between the β fibrinogen -148TT genotype and smoking on the risk of ischemic stroke was observed, a finding that should be interpreted with caution as the number of patients with the β fibrinogen -148TT genotype was small (n=8).

6.6 Prothrombin 20210G/A SNP and Ischemic Stroke

Prothrombin, or coagulation factor II, is a vitamin K-dependent zymogen that forms the precursor of thrombin. Upon activation by the coagulation pathway, prothrombin is converted to thrombin, which has a central role in the conversion of fibrinogen to fibrin. It is therefore reasonable to consider any factor that increases prothrombin activity to predispose to an augmented haemostatic response and ischemic stroke. In 1996, Poort *et al.* sequenced the coding and flanking regions of the prothrombin gene and found a G to A polymorphism at position 20210 which resulted in a 30% increase in the plasma level of prothrombin (Poort *et al.*, 1996). The SNP lies within the 3' untranslated region of the gene and is thought to influence prothrombin levels by enhancing mRNA stability during protein translation. The same investigators showed that this SNP was an independent risk factor for venous thrombosis. In a group of 471 unselected patients presenting with their first, objectively confirmed, deep vein thrombosis, the prevalence of the prothrombin 20210 A allele was 6.2%, compared with 2.3% in 474 age and gender matched controls, conferring a near 3-fold increase in risk of venous thrombosis (OR 2.8, 95%CI 1.4-5.6) (Poort *et al.*, 1996).

The relevance of the prothrombin 20210G/A SNP to arterial disease and in particular, to ischemic stroke remains controversial, with conflicting data existing. Of the negative studies, the US Physicians Health Study is the largest, consisting of 259 incident stroke cases and 1774 age and smoking-status matched controls. The study, however, has been criticized (Andreotti *et al.*, 2000), as it did not show an increase in risk of venous thrombosis, a finding that was convincingly shown in the original study by Poort *et al.* (Poort *et al.*, 1996). In the current study, the prothrombin 20210G/A SNP was not associated with an increased risk of ischemic stroke. The lack of association persisted following stratification of stroke subtype and aetiology. Furthermore, no significant interaction with the known cerebrovascular risk factors was observed. This is in keeping with the findings of seven other studies that have concluded that the prothrombin 20210G/A SNP does not play an important role in the pathogenesis of acute ischemic stroke.

The lack of association between the prothrombin 20210G/A SNP and ischemic stroke does not necessarily imply that plasma prothrombin activity does not play an important part in the pathogenesis of ischemic stroke. This is well illustrated by a recent study amongst 49 patients with a transient ischemic attack (TIA) or ischemic stroke and 87 asymptomatic blood donors (Gomez-Garcia *et al.*, 2002). In this study, the prevalence of the prothrombin 20210G/A SNP was not statistically different between the two groups (OR 2.3, 95% CI 0.6-0.8). However, after exclusion of those with the prothrombin variant, a significantly higher prothrombin activity was observed in the TIA/stroke group (1.11 U/ml vs 0.97 U/ml, p=0.0003). Furthermore, a 3.2-fold increase in risk of cerebral ischemia was associated with a prothrombin activity of greater than 1.10 U/ml (Gomez-Garcia *et al.*, 2002). The findings suggest that prothrombin does play an important role in the pathogenesis of ischemic stroke and that factors other than the prothrombin 20210G/A SNP may be important in determining prothrombin activity.

One important difference between the current study and others (Longstreth *et al.*, 1998, De Stefano *et al.*, 1998), is the age of the stroke patients studied. In the latter, stroke patients were less that 50 years of age and without known cerebrovascular risk factors such including diabetes, hypertension and hyperlipidemia. Why the risk of ischemic stroke associated with the prothrombin 20210G/A SNP is limited to young stroke patients of cryptogenic aetiology and not to an older stroke population, is not well understood. One possible explanation is the higher incidence of paradoxical embolism in the former subgroup. Atrial septal defects including a patent foramen ovale and atrial septal aneurysms are common in cases of young cryptogenic ischemic stroke (Mas *et al.*, 2001). These structural abnormalities permit the passage of venous thrombi from the right to left atrium and subsequent embolic occlusion of distal intracerebral vessels. Therefore, in this subgroup, factors that promote venous thrombosis can also be considered risk factors for ischemic stroke. In keeping with this hypothesis, the relevance of the prothrombin 20210G/A SNP in this setting has been shown recently (Lichy *et al.*, 2003). In a study involving 220 young ischemic stroke patients associated with a patent foramen ovale and 362 healthy controls, heterozygosity for the prothrombin 20210A allele conferred a 3.7-fold increase in risk

of ischemic stroke (OR 3.66, 95%CI 1.25-10.75, p=0.01). By contrast, there was no significant difference in the prevalence of the prothrombin 20210G/A SNP between healthy controls and a second cohort of 196 ischemic stroke patients, none of whom had an atrial septal abnormality (OR 1.5, 95%CI 0.42-5.41, p=0.5) (Lichy *et al.*, 2003). This suggests that the prothrombin 20210G/A SNP mediates its effect through an increased predisposition for venous rather than arterial thrombosis, with subsequent cerebral embolization occurring in the setting of an atrial septal defect.

6.7 Platelet Glycoprotein SNP's and Ischemic Stroke

Whilst platelet aggregation plays a central role in normal haemostasis, an exaggerated response has been implicated in pathogenesis of acute thrombosis causing ischemic stroke (Uchiyama *et al.*, 1983, Uchiyama *et al.*, 1994). Platelet aggregation is initiated following a breech of vascular endothelium, which permits the binding of platelet glycoprotein 1b/IX/V complex to sub-endothelial tissues. A cascade of intracellular events then follows, resulting in platelet activation and cross-linking of platelets via fibrinogen bound to platelet glycoprotein IIb/IIIa (Berndt *et al.*, 1995). Platelet aggregation can be stimulated in vitro by various agonists including adenosine diphosphate (ADP), adrenaline, collagen and thrombin. Other important determinants of increased platelet aggregation include increasing age, elevated cholesterol levels, presence of diabetes or hypertension, cigarette smoking, consumption of a high fat diet and excessive alcohol (O'Donnell *et al.*, 2001). The significance of these influences was recently determined in a study of the Framingham cohort (O'Donnell *et al.*, 2001). In this study, platelet function was compared between sibling and spouse pairs and determined by measuring the threshold concentration of adrenalin and ADP required to produce a platelet aggregation response. The study found that genetic factors had the greatest influence on platelet function, accounting for 30 percent of the variability of platelet aggregation, whilst the environmental factors listed above accounted for percent of the variability (O'Donnell *et al.*, 2001). To date, the specific genes mediating this effect are unknown.

The current study did not demonstrate a significant association between ischemic stroke and the glycoprotein IIb/IIIa P1A1/A2 SNP. The lack of association persisted

following stratification for stroke subtype or aetiology. The results of this study support the negative findings made in five previous studies (Carlsson *et al.*, 1997, Ridker *et al.*, 1997, Kekomaki *et al.*, 1999, Reiner *et al.*, 2000, van Goor *et al.*, 2002) and disagree with four others that showed a positive association (Carter *et al.*, 1998, Wagner *et al.*, 1998, Pongracz *et al.*, 2001, Streifler *et al.*, 2001). In the largest of these, involving 504 stroke patients and 402 controls, a positive association was limited to a subgroup of non-smokers (OR 2.37, 95%CI 1.19-4.74) and in those less than 50 years of age (OR 1.68, 95%CI 1.00-2.82) (Carter *et al.*, 1998). When the entire cohort of stroke patients and controls was included in the analysis, no statistically significant association between the glycoprotein IIb/IIIa P1A1/A2 SNP and ischemic stroke was observed. The reason why a greater risk of ischemic stroke was apparent in non-smokers remains unclear, however the authors suggest that the greater risk of ischemic stroke associated with smoking may have "masked" the effect of the Glycoprotein IIb/IIIa P1A1/A2 SNP. An alternative explanation is that a spurious association was made (type 1 error). In the current study, the risk of ischemic stroke associated with the glycoprotein IIb/IIIa P1A1/A2 SNP was not influenced by smoking status on bivariate analysis. Furthermore, no interaction between the glycoprotein IIb/IIIa P1A1/A2 SNP and smoking was observed. The findings of the current study suggest that the glycoprotein IIb/IIIa P1A1/A2 SNP does not play an important role in the pathogenesis of acute thrombosis and ischemic stroke. This is further supported by the platelet aggregation study performed on the Framingham cohort (O'Donnell *et al.*, 2001). Despite genetic factors accounting for 30 percent of the variability of platelet aggregation, the glycoprotein IIb/IIIa P1A1/A2 SNP contributed less than 1 percent to the overall variance.

In the current study, the glycoprotein 1b HPA2a/b SNP was not associated with ischemic stroke. This was also true following stratification for stroke subtype and aetiology. Furthermore, no interactions between the known cerebrovascular risk factors and the glycoprotein 1b HPA2a/b SNP was observed. Glycoprotein 1b plays an important role in the initial adhesion of platelets to exposed subendothelial tissues (Berndt *et al.*, 1995). The process acts as a "platelet brake" that slows down the circulating platelet and allows for other critical binding interactions to occur.

Glycoprotein 1b also mediates platelet activation and thrombosis that is precipitated by high shear stress conditions, such as those that are encountered across high-grade arterial stenoses (Douglas *et al.*, 2000). Whether the glycoprotein 1b HPA2a/b SNP predisposes to acute thrombosis and ischemic stroke in this specific setting is uncertain and could not be demonstrated by this study, as vascular imaging was not performed. The functional relevance of the glycoprotein 1b HPA2a/b SNP in a general stroke population, however, has been recently determined (Meiklejohn *et al.*, 2001). In this study involving 77 patients with ischemic stroke and a similar number of controls matched for age and gender, platelet P-selectin and fibrinogen binding were used as markers of platelet activation and compared between cases and controls stratified for glycoprotein 1b HPA2a/b SNP genotype. The study found that although platelet activation was strongly associated with ischemic stroke ($p<0.001$), it remained independent of glycoprotein 1b HPA2a/b SNP genotype. This supports the findings of the current study that the glycoprotein 1b HPA2a/b SNP does not play an important role in the pathogenesis of ischemic stroke.

6.8 Study Limitations

The potential limitations of this study warrant consideration. Most importantly, the study sample size was considerably smaller than that anticipated during the planning phase of the study. This limited the statistical power to detect an association between the pathogenic genotypes and ischemic stroke. The lack of statistical power was particularly illustrated when all potential confounding influences were incorporated into a multivariate model for the TPA –7351 SNP. Although a statistically significant association between the TPA –7351TT genotype and ischemic stroke was observed when the "important" cerebrovascular risk factors were considered in the multivariate model, when all potential confounding variables were incorporated, this association was weakened to be significant only at a 10% level.

The impact of the small sample size is further illustrated in table 6.1, which shows the estimated number of cases and controls required to establish a statistically significant association (at the 5% significance level) between each SNP and ischemic stroke. The estimates are calculated using the adjusted odds ratios and the background genotype

frequencies of each SNP as determined by this study. When a one to one patient-control ratio is considered, only the sample size estimate for the TPA-7351 SNP approached the number of controls recruited in this study. Statistical significance for this SNP, amongst the entire stroke cohort, was achieved by increasing the patient-control ratio to 1:1.65.

Table 6.1
Sample Size Estimation Based on Adjusted OR's Determined for Each SNP

SNP Genotype	Adjusted OR	Controls (%)	Sample Size*
PON1 54LL	1.0	42	7,047
PAI-1 4G4G	0.9	27	7,359
PON1 107CC	1.4	13	1,090
TPA –7351TT	1.9	9	366
β Fib –148TT	1.7	3	1,513
Prothrombin 20210GA	1.1	3	56,823
GpIIIa P1A1A2	1.5	27	477
GpIbHPA 2a2b	0.7	14	1,175

Sample sizes were estimated using Winepiscope2.0 – sample size estimation of an unmatched case-control study. Estimations are based on OR's for each SNP following adjustment for confounding variables (Table 5.8). All estimates are based on a study power of 80%, a 95% level of confidence and one to one matching of cases with controls.
**Sample size estimations are per group – i.e. require an equivalent number of cases and controls*

The estimates also show that for the remaining SNP's, the recruited number of patients and controls fell considerably short of that required to determine a statistically significant association. The negative outcomes for each of these SNP's may therefore represent type 2 errors, which require much larger studies of greater power to confirm the absence of association.

The small sample size may have also led to a spurious positive association. Clinical relevance for both the TPA –7351C/T and the PON1 –107C/T SNP's arose following stratification for stroke subtype. Consequently, these associations occurred in an even smaller population, resulting in relatively wide confidence intervals that approached unity. Therefore, the possibility of a type 1 error cannot be excluded and a study of greater power is required to confirm the strength of association. The inherent design of

a prospective case-control study also permits investigation of stroke survivors alone. Consequently, an underestimation of the pro-thrombotic polymorphisms amongst our stroke population is possible. Recruitment of these individuals, however, would strengthen any association, rather than dilute it.

Various sources of misclassification bias may have also arisen. Both hypertension and hypercholesterolemia appeared to be protective of ischemic stroke, with the effect being statistically significant (hypertension OR 0.6, 95%CI 0.4-0.9, hypercholesterolemia OR 0.5, 95%CI 0.3-0.7). These findings contradict those of previous large-scale, population-based epidemiological studies (Davis *et al.*, 1987, Grau *et al.*, , Simons *et al.*, 1998, D'Agostino *et al.*, 1994). In the current study, hypertension was coded using a composite definition that incorporated a past medical history of hypertension, use of anti-hypertensive medication and measured blood pressure. This definition was adopted in an attempt to accurately classify those with no personal history of hypertension who were objectively hypertensive at the time of recruitment. This, however, led to the misclassification of hypertension in some situations. When a past history of hypertension alone was considered, the apparent protective effect of hypertension was lost, suggesting that there was an over-representation of newly diagnosed hypertension amongst the control group. This is further illustrated following stratification of blood pressure according to the World Health Organization criteria (table 5.4). The prevalence of isolated systolic hypertension is significantly higher for cases than controls and is suggestive of a greater "white coat" effect amongst this group. Controls were recruited on an outpatient basis, with recruitment clinics held concurrently with other outpatient clinics at either The Queen Elizabeth or Lyell McEwin Hospitals. Therefore, the greater isolated systolic hypertension observed amongst controls may reflect the anxiety induced by a busy hospital outpatient environment. In contrast, for stroke patients, blood pressure was measured in a relaxed state at their bedside, typically in a quiet ward.

This study did not measure fasting cholesterol levels and for most controls, venous sampling was performed after breakfast. Restriction of oral food intake in stroke

patients, however, was not uncommon, and implemented to minimise the risk of aspiration pneumonia. Therefore, the higher prevalence of hypercholesterolemia amongst the control group most likely represents this disparity in fasting status at the time of venous blood collection. This is further supported by the lack of association between hypercholesterolemia and ischemic stroke when a history of hypercholesterolemia alone is considered (Table 5.5).

The limitations associated with the OCSP classification system could also have led to misclassification of stroke subtype. For example, the OCSP cannot differentiate between lacunar infarction (small vessel disease) and striato-capsular infarction without cortical signs (large vessel disease). The OCSP classification system, however, has been validated, with concordance between clinical subtype and brain neuro-imaging reported in 75% of cases *(Whisnant et al., 1996)*. Furthermore, in the present study, clinical subtype was made in conjunction with neuro-imaging, thus limiting this error.

When a case-control sample is derived from an ethnically mixed population, or even in an apparent homogenous population where ethnic mixing has occurred over several generations, non-random associations can occur at loci that are independent of a disease phenotype. This occurs because allele frequencies in both disease and non-disease causing genes vary widely amongst different ethnic groups. As a result, there will be inherent differences in allele frequency within subsets of a mixed population, irrespective of an association with a disease phenotype. Confounding in case-control studies can therefore arise if the study groups differ considerably in ancestral origin. This phenomenon is known as population stratification and is cited as the primary cause of non-replication of positive associations in case-control studies (Cardon and Palmer, 2003).

In the current study, the confounding effect of population stratification was not determined and a type 1 error may have occurred. One approach to dealing with this problem has been to match cases and controls for ethnicity (Reich and Goldstein, 2001). Although this was not performed in this study, the ethnic origin of participants

was recorded. Due to the lack of interpreting services and differences in cultural attitudes to participation in medical research, the study sample consisted chiefly of Caucasians, with no significant difference in the proportion of non-Caucasians between cases and controls. Furthermore, in the control group, the genotype distributions (for the SNP's found to be associated with ischemic stroke) were similar to that found in other Caucasian populations (Table 6.2).

Table 6.2
Variation of PON1 107 T/C and TPA-7351 C/T Genotype Distributions Amongst Caucasian Populations

SNP	**Genotype Distribution (%)**			**Reference**
	TT	TC	CC	
PON1 -107 T/C	20	67	13	Present Study
PON1 -107 T/C	21	56	23	(James *et al.*, 2000)
PON1 -107 T/C	32	44	24	(Leviev and James, 2000)
PON1 -107 T/C	25	50	25	(Brophy *et al.*, 2001)
	CC	CT	TT	
TPA –7351 C/T	46	45	9	Present Study
TPA –7351 C/T	51	39	10	(Ladenvall *et al.*, 2002)

Although this may appear favourable, the interpretation that it represents a uniform ethnic population should be made with caution, as this population is derived from a diverse range of ethnic backgrounds, with genetic admixture occurring over several generations. Thus, the impact of population stratification may still be playing an important role.

Association by linkage disequilibrium also warrants consideration. The finding of a SNP associated with a disease phenotype has one of three explanations. Firstly, the association is spurious and a type 1 error has occurred. Secondly, the association is true, however, it has arisen through linkage disequilibrium with a nearby genetic locus of functional significance; or thirdly, the SNP in question has a true pathogenic role in the development of the disease phenotype. The latter explanation is most likely for both the PON1 -107T/C and TPA –7351C/T SNP's, as strong evidence showing biological significance for these SNP's exists (Leviev and James, 2000, Ladenvall *et al.*, 2000). The SNP's occur within regulatory regions of their respective genes, with mRNA transcription and gene product activity highly dependent on genotype. Therefore, the positive associations determined in this study are unlikely to be secondary to linkage disequilibrium with another gene locus, although this cannot be totally excluded.

Chapter 7

Future Directions

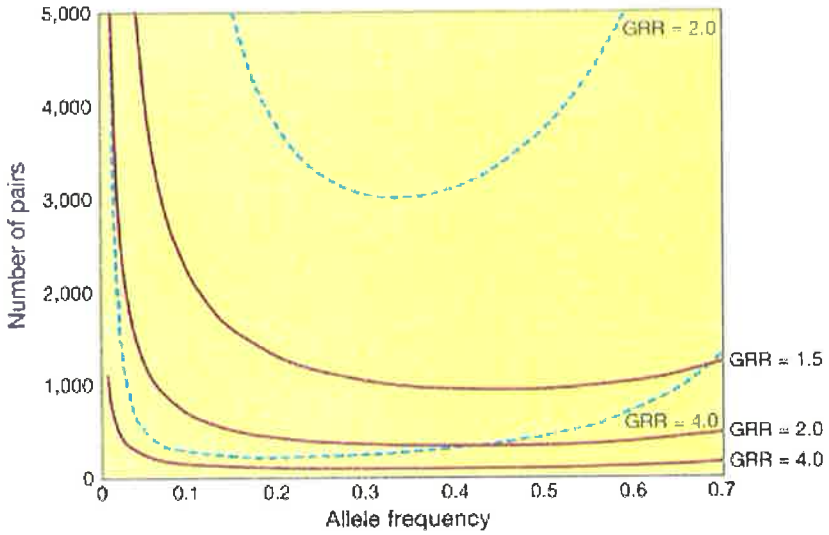
This thesis has identified two common SNP's in the TPA and PON1 genes that may play an important role in the pathogenesis of lacunar stroke. The association, however, was made in a study of relatively small sample size, with some methodological limitations. As such, a larger, well-designed study of greater statistical power is required to confirm the strength of association. This chapter will discuss key issues in the implementation of such a study.

7.1 Selection of Appropriate Study Design

In the planning of any study, the selection of the most appropriate study design that will address the hypothesis is critical. For genetic studies, there are two main approaches: Linkage analysis and genetic association. Linkage analysis relies on the identification of genetic markers that are in linkage disequilibrium with a disease locus. This can be used to examine co-segregation with a disease phenotype amongst sibling pairs. Despite this method being used in identifying mutations associated with single-gene disorders, its application to diseases of complex genetic inheritance has been less effective, as it relies greatly on mutations with a strong phenotypic effect and high penetrance (Cardon and Bell, 2001). For non-Mendelian diseases such as ischemic stroke, this situation does not occur, with the genetic predisposition arising through the interaction and accumulation of numerous genetic loci, each of which have a smaller phenotypic effect and weaker penetrance. Thus, the application of linkage analysis in this setting is less powerful in detecting a genetic predisposition. This is further illustrated by a recent study that developed a biometric model comparing linkage analysis (using affected sibling pairs) with the case-control study design (Risch, 2000). In this model, it was assumed that risk associated with the homozygous state for the disease-causing allele was the square of the risk associated with the

heterozygous state (Figure 7.1). The model showed that application of linkage analysis is feasible for alleles that possess a strong phenotypic effect (i.e. genotype relative risk = 4.0). However, for alleles with weaker effects (genotype relative risk = 2.0), linkage analysis is insufficiently powered to detect a statistical association, requiring unrealistically large sample sizes. In contrast, case-control studies are suitably powered to detect an association for alleles of weaker penetrance (figure 7.1).

Figure 7.1
Comparison of Linkage (dashed lines) and Association Analysis (solid lines) for Detecting Genetic Effects



Comparison of linkage (dashed lines) with association analysis (solid lines) for detecting genetic effects.
GRR = Genotype Relative Risk
From (Risch, 2000)

In the current study, the risk of lacunar stroke associated with the PON1 -107CC and TPA –7,351TT genotype was increased two to three-fold. If confirmed by a larger study, the true increase in risk may be even smaller, as the magnitude of genetic effect is often overestimated by small association studies (Ioannidis *et al.*, 2003). This relatively weak genetic effect is therefore most amenable to investigation with a case-control study design.

7.2 The Importance of Stroke Sub-typing

Accurate sub-typing of ischemic stroke is of paramount importance, as the underlying pathophysiology is heterogenous. This is particularly relevant to the genetic predisposition of ischemic stroke. For example, the genes predisposing to intra-cardiac thrombosis are likely to be quite different to those causing atherosclerosis or lipohyalinosis. The OCSP classification system is often chosen as the clinical sub-typing tool (including the current study), as it has a low resource requirement, is easily implemented, has high reproducibility and low inter-observer error. However, it is poor at differentiating for subcortical large vessel disease, posterior circulation small vessel disease and cardio-embolism. A far superior tool is that developed by the TOAST investigators (Kolominsky-Rabas *et al.*, 2001). Although it requires significantly greater resources, it provides a robust means to differentiate between stroke subtypes. For small vessel disease, however, genetic association using the TOAST criteria may still be sub-optimal, as the criteria do not allow differentiation between lacunar stroke caused by lipohyalinosis, microatheroma or embolism, mechanisms in which different genes are likely to be responsible.

The importance of accurate stroke sub-typing for the purpose of establishing genetic risk has recently been confirmed (Jerrard-Dunne *et al.*, 2003). In this study, a family history of stroke was recorded in 1000 consecutive white subjects with ischemic stroke and 800 age- and gender-matched controls. All patients were sub-typed using the TOAST criteria. The study made several very important findings. First, after adjustment for conventional cerebrovascular risk factors, a family history of stroke did not appear to be a significant risk factor for ischemic stroke (OR 1.22, 95%CI 0.90-1.39). However, when the analysis was limited to those with a history of stroke occurring in a relative less than 65 years of age, a statistically significant association with stroke was found (OR 1.38, 95%CI 1.01-1.90). Furthermore, following stratification for stroke subtype, an association between ischemic stroke and a family history of stroke (less than 65 years) was limited to those classified with small or large vessel ischemic stroke (small vessel disease OR 1.93, 95%CI 1.25-2.97, large vessel

disease OR 2.24, 95%CI 1.49-3.36). A family history of stroke was not an important risk factor for cardio-embolic stroke or ischemic stroke that could not be classified.

This study also showed a strong interaction between a family history of stroke (less than 65 years of age) and the age of stroke onset. The risk of small and large vessel ischemic stroke associated with a family history (less than 65 years of age) increased as the age of stroke onset decreased (Table 7.1). For example, in those presenting with small vessel disease less than 55 years of age, a family history of stroke (less than 65 years of age) was associated with a 4-fold increase in risk. The corresponding odds ratio for those presenting with small vessel-related ischemic stroke that were less than 65 years of age was 2.69 and for those less than 80 years, was 1.55 (Table 7.1).

Table 7.1
Relationship Between Age of Stroke and Positive Family History of Stroke < 65 Years for Small and Large Vessel Disease

	Small Vessel Disease	Large Vessel Disease
	OR (95%CI)	**OR (95%CI)**
Stroke <55 y	3.99 (1.25-12.7)	4.46 (1.03-19.3)
Stroke <60 y	2.70 (1.18-6.18)	2.55 (1.04-6.25)
Stroke <65 y	2.69 (1.46-4.96)	2.34 (1.21-4.52)
Stroke <70 y	1.91 (1.11-3.28)	1.86 (1.10-3.12)
Stroke <75 y	1.55 (0.94-2.53)	1.88 (1.18-3.00)
Stroke <80 y	1.55 (0.97-2.48)	1.82 (1.15-2.86)
All Strokes	1.49 (0.94-2.37)	1.67 (1.08-2.66)

From (Jerrard-Dunne et al., 2003)

The findings illustrate the importance of accurate stroke sub-typing and suggest that the greatest likelihood of establishing a genetic determinant for ischemic stroke will occur when the study sample consists of a younger stroke population presenting with small or large vessel occlusion and having positive family history for early onset stroke. Patients with cardio-embolic or unclassified stroke should be excluded from analysis, as their recruitment is likely to weaken the strength of any association.

Table 7.2
Estimated Sample Size Requirement for Studies Using Specific Stroke Subtypes and Age Groups

		No. of Cases in Case-Control Study*		
	OR (95%CI) associated with SNP**	**Allele Frequency 5%**	**Allele Frequency 10%**	**Allele Frequency 15%**
All Stroke	1.38 (1.01-1.90)	2761	1481	1065
All Stroke <65yrs	2.21 (1.41-3.47)	369	205	152
Large vessel	1.67 (1.08-2.66)	996	542	395
Large vessel <65yrs	2.34 (1.21-4.52)	313	175	131
Small-vessel	1.49 (0.94-2.37)	1739	936	676
Small-vessel <65yrs	2.69 (1.46-4.96)	217	123	93

** Estimated sample sizes required to detect an effect equal to the observed OR's and 95% CI with a P value of <0.05 and 80% power.*
***OR's are based on the observed multivariate OR for a positive family history of stroke <65years*
From (Jerrard-Dunne et al., 2003)

These authors also showed that targeting a specific stroke sub-group defined by age, family history of stroke and stroke sub-type, might impact on the sample size required to establish a significant association. This is illustrated in table 7.2. In subgroups where the impact of a genetic predisposition is high, the sample size required to establish an association is small. For example, in a study of patients of less than 65 years of age, who present with small vessel disease and have a positive family history of early-onset stroke (less than 65 years), the estimated sample size required to establish an association for a genotype of 10% prevalence is 123 (range 39-1049) (Jerrard-Dunne *et al.*, 2003). This figure is particularly relevant to the present study which showed the background frequency rate of the PON1 107CC and TPA –7,351TT genotypes to be 13% and 9% respectively. The methodology used to estimate the sample sizes, however, deserves some discussion. The estimates were based on the odds ratios determined for a positive family history of stroke that serves as a surrogate for the 'total' genetic predisposition. Using this to estimate the sample size required to establish association with one SNP, assumes that this SNP accounts totally for the

genetic risk of stroke. This assumption is unlikely and contradicts the polygenic model of ischemic stroke.

Although in theory, restriction of the case-control study to a pre-specified stroke subgroup appears ideal, application of such a study is difficult, as the number of stroke patients required to be screened is high. This problem, however, could be overcome if recruitment was derived from large prospective population-based incidence study that involved multiple centres and spanned wide geographical regions. This study design would also allow the identification of a large number of controls that could be matched for demographic characteristics and ethnicity.

7.3 Adjustment for Population Stratification

Bias from the effect of population stratification is often cited as the reason for non-replication of SNP associations with disease. Population stratification bias occurs when an apparent allelic association with a disease phenotype arises not due a true association, but due to differences in ethnic composition between cases and controls. The selection of controls is therefore critical, as any systematic differences in allele frequency may manifest as a disease association. Although matching of cases and controls for ethnic origin may minimise this bias, the potential for spurious association persists, as considerable ethnic admixture may still exist even in an apparent homogeneous population. Therefore strategies are required to either eliminate, or alternatively, quantify and adjust for the confounding effects of this bias.

This bias can be minimised largely by using a prospective cohort study (Cardon and Bell, 2001). This requires the recruitment of a large number of asymptomatic individuals who are monitored throughout the study for the incidence of ischemic stroke. As patients are then matched with stroke-free individuals from the same population, there is no bias for the selection of controls. The approach, however, requires significantly more resources and can take many years before an adequate number of incident strokes can be collected. The method could be applied in a large-scale multi-centre study, which would span wide geographical regions. Another significant advantage of this study is the ability to prospectively record and obtain a

more accurate profile of the known cerebrovascular risk factors. An alternative strategy has been to select controls from families of affected stroke patients. The most common approach used is the Transmission Disequilibrium Test, which analyses the transmission of disease-causing alleles from living parents to their affected offspring (Hassan *et al.*, 2002). The methodology, however, relies on the availability of heterozygous parents and therefore a significantly larger number of cases require to be screened. Furthermore, the requirement for parental genotypes makes its use in late-onset diseases such as ischemic stroke less feasible.

More recently, a method allowing for quantification and adjustment for population stratification bias has been devised (Reich and Goldstein, 2001). The method involves genotyping a number of unlinked SNP's that have no theoretical relationship with the disease phenotype, in addition to the candidate SNP. In the ideal situation, where similar allele frequencies exist between cases and controls, no statistical association between the unlinked markers and the disease phenotype should arise. However, when ethnic diversity between the groups occurs, the prevalence of unlinked markers will vary, with the average difference being detectable and representative of the magnitude of population stratification bias. The ability of this methodology to detect population stratification bias will consequently depend on the number of unlinked SNP's used, with a range of 20 – 40 SNP's being most optimal (Reich and Goldstein, 2001). This method provides a robust and feasible means to quantify and adjust for population stratification that can be applied readily to relatively small genetic association studies.

7.4 Alternate Disease Phenotypes

The relevance of the TPA –7,351TT and PON1 -107CC genotypes to other small vessel disease phenotypes such as subcortical leukoaraiosis also warrants further investigating. Extrapolation of the findings to subcortical leukoaraiosis should be made with caution, as the underlying pathology is mainly due to arteriosclerosis rather than microatheroma or lipohyalinosis (Dozono *et al.*, 1991). Despite this, a close correlation between lacunar infarction and subcortical arteriololsclerosis has been observed (Dozono *et al.*, 1991). Furthermore, in the Austrian Stroke Prevention Study, the PON1 54ML SNP was shown to be significantly associated with progression of subcortical vascular

lesions (Schmidt *et al.*, 1998a). It remains feasible, therefore, that the TPA and PON1 SNP's may influence the development and progression of this pathology and may be an important determinant of vascular disease-related cognitive decline. Subcortical leukoaraiosis also serves as an intermediate phenotype, as the affected individual may be asymptomatic. A case control study would be ideal to determine the relevance of the TPA and PON1 SNP's to this phenotype, however, this would also require neuro-imaging of controls to exclude subclinical disease. An alternate approach is a prospective cohort study, whereby each asymptomatic participant is followed with serial neuro-imaging and a comparison made for the prevalence of the pathogenic alleles between those who showed objective evidence of progression with those who did not. Such a method would also eliminate any population stratification bias.

The association between lacunar stroke and the TPA –7,351C/T SNP leads to the hypothesis that this SNP may also have a role in the development of hypertensive-type intracerebral haemorrhage. This arises as lipohyalinosis underlies both phenotypes. The findings of this study suggest that the prothrombotic TPA–7,351 TT genotype tips the balance towards thrombosis, with a subsequent phenotypic expression of lacunar infarction. Whether the profibrinolytic TPA–7,351CC genotype predisposes to the haemorrhagic phenotype warrants further investigation. A well-designed case-control study would be necessary to pursue this research question.

7.5 Final Considerations

Table 7.3 lists the factors required to confirm the strength of association of the findings of this study. The case-control study remains a powerful research tool when issues relating to stroke sub-typing, sample size estimation and population stratification bias are considered. Furthermore, as cases and controls are recruited prospectively, DNA samples can be stored and used for future candidate gene studies. This is particularly important as high-throughput genome-wide SNP analysis technologies such as that developed by Sequenom (www.sequenom.com) become available. This technology will enable a detailed genetic profile to be determined that can be used to stratify individuals with the highest risk of first and subsequent ischemic stroke. Genome-wide SNP analysis can also be applied to a case-control study to find novel genetic loci

associated with ischemic stroke. This gene mapping strategy uses approximately 100,000 SNP's that span the entire human genome and relies on the differential clustering of SNP's around a disease locus between cases and controls (Hassan and Markus, 2000). High-throughput genotyping systems will also enhance our understanding of the underlying pathophysiological mechanisms and pave the way for new therapeutic interventions.

Table 7.3
Recommendations for the Model Genetic Association Study

1. Case-control study design (sub-study of a larger prospective incidence study ideal).
2. Selection of SNP's that play an important role in the pathogenesis of ischemic stroke.
3. Avoidance of candidate genes that are in linkage disequilibrium with other SNP's of functional significance.
4. Measurement and adjustment for factors that may influence gene product expression.
5. Validation of genetic effect, i.e.- correlation of genotype with plasma levels of gene product. Selection of candidate genes with the greatest functional significance only.
6. Large sample size – at least 1000 cases and 1000 controls. Best performed by multi-centre collaboration.
7. Significance level set at $p<0.01$ or less to avoid type 1 errors.
8. Standardized assessment of known cerebrovascular risk factors.
9. Adoption of WHO definition of stroke (Hatono, 1976).
10. Accurate stroke sub-typing using TOAST criteria.
11. Neuro-imaging (preferably MRI) confirmation of ischemic stroke.
12. Exclusion of cardio-embolic and unclassified stroke sub-types.
13. Exclusion of patients over 65 years of age.
14. Eligibility restricted to patients with a history of stroke in a 1st degree relative.
15. Random selection of population controls.
16. Controls matched for age, gender and ethnic origin.
17. Genotyping of 20-40 unlinked SNP's for the determination and adjustment of population stratification bias.
18. Neuro-imaging (preferably MRI) confirmation of normality in controls. Particularly important in studies of lacunar stroke to exclude subclinical disease and avoid misclassification.
19. Avoidance of sub-group analysis.
20. Establishment of DNA bank and database for future investigation.

This thesis has identified two SNP's that may play an important role in the predisposition of lacunar stroke, as well as factors that can significantly impact on the outcome of a case-control study. As the study is of limited statistical power, the findings should be considered as hypothesis generating. Confirmation, with a larger study of greater power is necessary to address all the issues raised in this chapter. With the recent advances in biotechnology and understanding of population genetics, a study of this calibre is now feasible. We have the know-how; the challenge to confirm these findings and apply them to other cerebrovascular phenotypes now lies ahead.

Bibliography

ABS (2001). 3303.0 Causes of Death, Australia, Australian Bureau of Statistics.

Andreotti, F., Burzotta, F., De Stefano, V., Maseri, A. and Iacoviello, L. (2000). The G20210A prothrombin mutation and the Physicians' Health Study. *Circulation.* **101:** E207-8.

Austin, H., Chimowitz, M. I., Hill, H. A., Chaturvedi, S., Wechsler, L. R., Wityk, R. J., Walz, E., Wilterdink, J. L., Coull, B., Sila, C. A., Mitsias, P., Evatt, B., Hooper, W. C., Genetics and Stroke in the Young Study Group (2002). Cryptogenic stroke in relation to genetic variation in clotting factors and other genetic polymorphisms among young men and women. *Stroke.* **33:** 2762-8.

Aviram, M., Billecke, S., Sorenson, R., Bisgaier, C., Newton, R., Rosenblat, M., Erogul, J., Hsu, C., Dunlop, C. and La Du, B. (1998). Paraoxonase active site required for protection against LDL oxidation involves its free sulfhydryl group and is different from that required for its arylesterase/paraoxonase activities: selective action of human paraoxonase allozymes Q and R. *Arteriosclerosis, Thrombosis & Vascular Biology.* **18:** 1617-24.

Bak, S., Gaist, D., Sindrup, S. H., Skytthe, A. and Christensen, K. (2002). Genetic liability in stroke: a long-term follow-up study of Danish twins. *Stroke.* **33:** 769-74.

Baker, R. I., Eikelboom, J., Lofthouse, E., Staples, N., Afshar-Kharghan, V., Lopez, J. A., Shen, Y., Berndt, M. C. and Hankey, G. (2001). Platelet glycoprotein Ib{alpha} Kozak polymorphism is associated with an increased risk of ischemic stroke. *Blood.* **98:** 36-40.

Bamford, J., Sandercock, P., Dennis, M., Burn, J. and Warlow, C. (1991). Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet.* **337:** 1521-6.

Bang, C. O., Park, H. K., Ahn, M. Y., Shin, H. K., Hwang, K. Y. and Hong, S. Y. (2001). 4G/5G polymorphism of the plasminogen activator inhibitor-1 gene and insertion/deletion polymorphism of the tissue-type plasminogen activator gene in atherothrombotic stroke. *Cerebrovascular Diseases.* **11:** 294-9.

Behague, I., Poirier, O., Nicaud, V., Evans, A., Arveiler, D., Luc, G., Cambou, J. P., Scarabin, P. Y., Bara, L., Green, F. and Cambien, F. (1996). Beta fibrinogen gene polymorphisms are associated with plasma fibrinogen and coronary artery disease in patients with myocardial infarction. The ECTIM Study. *Circulation.* **93:** 440-9.

Berndt, M. C., Ward, C. M., De Luca, M., Facey, D. A., Castaldi, P. A., Harris, S. J. and Andrews, R. K. (1995). The molecular mechanism of platelet adhesion. *Australian & New Zealand Journal of Medicine.* **25:** 822-30.

Blake, G. J., Schmitz, C., Lindpaintner, K. and Ridker, P. M. (2001). Mutation in the promoter region of the beta-fibrinogen gene and the risk of future myocardial infarction, stroke and venous thrombosis.[comment]. *European Heart Journal.* **22:** 2262-6.

Boysen, G., Nyboe, J., Appleyard, M., Sorensen, P. S., Boas, J., Somnier, F., Jensen, G. and Schnohr, P. (1988). Stroke incidence and risk factors for stroke in Copenhagen, Denmark. *Stroke.* **19:** 1345-53.

Brass, L. M., Isaacsohn, J. L., Merikangas, K. R. and Robinette, C. D. (1992). A study of twins and stroke. *Stroke.* **23:** 221-3.

Brass, L. M. and Shaker, L. A. (1991). Family history in patients with transient ischemic attacks. *Stroke.* **22:** 837-41.

Bray, P. F., Weiss, E. J., Tayback, M. and Goldschmidt-Clermont, P. J. (1997). PlA1/A2 polymorphism of platelet glycoprotein IIIa and risk of cardiovascular disease. *Lancet.* **349:** 1100-1.

Brockmann, M. J., Schwendemann, G. and Stief, T. W. (1991). Plasminogen activator inhibitor in acute stroke. *Molecular & Chemical Neuropathology.* **14:** 143-50.

Brophy, V. H., Jampsa, R. L., Clendenning, J. B., McKinstry, L. A., Jarvik, G. P. and Furlong, C. E. (2001). Effects of 5' regulatory-region polymorphisms on paraoxonase-gene (PON1) expression. *American Journal of Human Genetics.* **68:** 1428-36.

Bunce, M., O'Neill, C. M., Barnardo, M. C., Krausa, P., Browning, M. J., Morris, P. J. and Welsh, K. I. (1995). Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP). *Tissue Antigens.* **46:** 355-67.

Cardon, L. R. and Bell, J. I. (2001). Association study designs for complex diseases. *Nature Reviews Genetics.* **2:** 91-9.

Cardon, L. R. and Palmer, L. J. (2003). Population stratification and spurious allelic association. *Lancet.* **361:** 598-604.

Carlsson, L. E., Greinacher, A., Spitzer, C., Walther, R. and Kessler, C. (1997). 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* **28:** 1392-5.

Carmelli, D., DeCarli, C., Swan, G. E., Jack, L. M., Reed, T., Wolf, P. A. and Miller, B. L. (1998). Evidence for genetic variance in white matter hyperintensity volume in normal elderly male twins. *Stroke.* **29:** 1177-81.

Carrieri, P. B., Orefice, G., Maiorino, A., Provitera, V., Balzano, G. and Lucariello, A. (1994). Age-related risk factors for ischemic stroke in Italian men. *Neuroepidemiology.* **13:** 28-33.

Carter, A. M., Catto, A. J., Bamford, J. M. and Grant, P. J. (1998). Platelet GP IIIa PlA and GP Ib variable number tandem repeat polymorphisms and markers of platelet activation in acute stroke. *Arteriosclerosis, Thrombosis & Vascular Biology.* **18:** 1124-31.

Catto, A. J., Carter, A. M., Stickland, M., Bamford, J. M., Davies, J. A. and Grant, P. J. (1997). Plasminogen activator inhibitor-1 (PAI-1) 4G/5G promoter polymorphism and levels in subjects with cerebrovascular disease. *Thrombosis & Haemostasis.* **77:** 730-4.

Cesari, M., Sartori, M. T., Patrassi, G. M., Vettore, S. and Rossi, G. P. (1999). Determinants of plasma levels of plasminogen activator inhibitor-1 : A study of normotensive twins. *Arteriosclerosis, Thrombosis & Vascular Biology.* **19:** 316-20.

Chalmers, J. and Chapman, N. (2001). Progress in reducing the burden of stroke. *Clinical & Experimental Pharmacology & Physiology.* **28:** 1091-5.

Chalmers, J., MacMahon, S., Mancia, G., Whitworth, J., Beilin, L., Hansson, L., Neal, B., Rodgers, A., Ni Mhurchu, C. and Clark, T. (1999). 1999 World Health Organization-International Society of Hypertension Guidelines for the management of hypertension. Guidelines sub-committee of the World Health Organization. *Clinical & Experimental Hypertension (New York).* **21:** 1009-60.

Chien, A., Edgar, D. B. and Trela, J. M. (1976). Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. *Journal of Bacteriology.* **127:** 1550-7.

Colditz, G. A., Bonita, R., Stampfer, M. J., Willett, W. C., Rosner, B., Speizer, F. E. and Hennekens, C. H. (1988). Cigarette smoking and risk of stroke in middle-aged women. *New England Journal of Medicine.* **318:** 937-41.

Constantinides, P. (1967). Pathogenesis of cerebral artery thrombosis in man. *Archives of Pathology.* **83:** 422-428.

Corral, J., Gonzalez-Conejero, R., Lozano, M. L., Rivera, J., Heras, I. and Vicente, V. (1997). The venous thrombosis risk factor 20210 A allele of the prothrombin gene is not a major risk factor for arterial thrombotic disease. *British Journal of Haematology.* **99:** 304-7.

D'Agostino, R. B., Wolf, P. A., Belanger, A. J. and Kannel, W. B. (1994). Stroke risk profile: adjustment for antihypertensive medication. The Framingham Study. *Stroke. 25(1):40-3.*

Davis, P. H., Dambrosia, J. M., Schoenberg, B. S., Schoenberg, D. G., Pritchard, D. A., Lilienfeld, A. M. and Whisnant, J. P. (1987). Risk factors for ischemic stroke: a prospective study in Rochester, Minnesota. *Annals of Neurology.* **22:** 319-27.

Dawson, S. J., Wiman, B., Hamsten, A., Green, F., Humphries, S. and Henney, A. M. (1993). The two allele sequences of a common polymorphism in the promoter of the plasminogen activator inhibitor-1 (PAI-1) gene respond differently to interleukin-1 in HepG2 cells. *Journal of Biological Chemistry.* **268:** 10739-45.

De Stefano, V., Chiusolo, P., Paciaroni, K., Casorelli, I., Rossi, E., Molinari, M., Servidei, S., Tonali, P. A. and Leone, G. (1998). Prothrombin G20210A mutant genotype is a risk factor for cerebrovascular ischemic disease in young patients. *Blood.* **91:** 3562-5.

Del Zoppo, G. J. (1998). The role of platelets in ischemic stroke. *Neurology.* **51:** S9-14.

Dewey, H., Macdonell, R., Donnan, G., McNeil, J., Freeman, E., Thrift, A. and Sharples, C. (2001a). Inter-rater reliability of stroke sub-type classification by neurologists and nurses within a community-based stroke incidence study. *Journal of Clinical Neuroscience.* **8:** 14-17.

Dewey, H. M., Thrift, A. G., Mihalopoulos, C., Carter, R., Macdonell, R. A., McNeil, J. J. and Donnan, G. A. (2001b). Cost of stroke in Australia from a societal perspective: results from the North East Melbourne Stroke Incidence Study (NEMESIS). *Stroke.* **32:** 2409-16.

Department of Human Services, South Australia (2002). Centre for Population Studies in Epidemiology.

Diaz, J. F., Hachinski, V. C., Pederson, L. L. and Donald, A. (1986). Aggregation of multiple risk factors for stroke in siblings of patients with brain infarction and transient ischemic attacks. *Stroke.* **17:** 1239-42.

Donnan, G. A., O'Malley, H. M., Quang, L., Hurley, S. and Bladin, P. F. (1993). The capsular warning syndrome: pathogenesis and clinical features. *Neurology.* **43:** 957-62.

Douglas, H., Gorog, D. and Ahmed, N. (2000). Functional significance of GP1b alpha receptor polymorphism in patients with coronary artery disease. *European Heart Journal.* **21 (Suppl):** 66.

Dozono, K., Ishii, N., Nishihara, Y. and Horie, A. (1991). An autopsy study of the incidence of lacunes in relation to age, hypertension, and arteriosclerosis. *Stroke.* **22:** 993-6.

Endler, G., Lalouschek, W., Exner, M., Mitterbauer, G., Haring, D. and Mannhalter, C. (2000). The 4G/4G genotype at nucleotide position -675 in the promotor region of the plasminogen activator inhibitor 1 (PAI-1) gene is less frequent in young patients with minor stroke than in controls. *British Journal of Haematology.* **110:** 469-71.

Eriksson, P., Kallin, B., van 't Hooft, F. M., Bavenholm, P. and Hamsten, A. (1995). Allele-specific increase in basal transcription of the plasminogen-activator inhibitor 1 gene is associated with myocardial infarction. *Proceedings of the National Academy of Sciences of the United States of America.* **92:** 1851-5.

Ferraresi, P., Marchetti, G., Legnani, C., Cavallari, E., Castoldi, E., Mascoli, F., Ardissino, D., Palareti, G. and Bernardi, F. (1997). The heterozygous 20210 G/A prothrombin genotype is associated with early venous thrombosis in inherited thrombophilias and is not increased in frequency in artery disease. *Arteriosclerosis, Thrombosis & Vascular Biology.* **17:** 2418-22.

Fey, G. H. and Fuller, G. M. (1987). Regulation of acute phase gene expression by inflammatory mediators. *Molecular Biology & Medicine.* **4:** 323-38.

Fisher, C. (1979). Capsular Infarcts-the underlying vascular lesions. *Archives of Neurol*ogy. **36:** 65-73.

Fortunato, G., Rubba, P., Panico, S., Trono, D., Tinto, N., Mazzaccara, C., De Michele, M., Iannuzzi, A., Vitale, D. F., Salvatore, F. and Sacchetti, L. (2003). A paraoxonase gene polymorphism, PON 1 (55), as an independent risk factor for increased carotid intima-media thickness in middle-aged women. *Atherosclerosis* **167:** 141-8.

Gan, R., Sacco, R. L., Kargman, D. E., Roberts, J. K., Boden-Albala, B. and Gu, Q. (1997). Testing the validity of the lacunar hypothesis: the Northern Manhattan Stroke Study experience. *Neurology.* **48:** 1204-11.

Giordano, P., De Lucia, D., Coppola, B. and Iolascon, A. (1999). Homozygous prothrombin gene mutation and ischemic cerebrovascular disease. *Acta Haematologica.* **102:** 101-103.

Gomez-Garcia, E. B., van Goor, M. P., Leebeek, F. W., Brouwers, G. J., Koudstaal, P. J. and Dippel, D. W. (2002). Elevated prothrombin is a risk factor for cerebral arterial ischemia in young adults. *Clinical Neurology & Neurosurgery.* **104:** 285-8.

Gonzalez-Conejero, R., Lozano, M. L., Rivera, J., Corral, J., Iniesta, J. A., Moraleda, J. M. and Vicente, V. (1998). Polymorphisms of platelet membrane glycoprotein Ib associated with arterial thrombotic disease. *Blood* **92:** 2771-6.

Graffagnino, C., Gasecki, A. P., Doig, G. S. and Hachinski, V. C. (1994). The importance of family history in cerebrovascular disease. *Stroke* **25:** 1599-604.

Grau, A. J., Weimar, C., Buggle, F., Heinrich, A., Goertler, M., Neumaier, S., Glahn, J., Brandt, T., Hacke, W. and Diener, H. C. (2001). Risk factors, outcome, and treatment in subtypes of ischemic stroke: the German stroke data bank. *Stroke. 32(11): 2559-66*

Gubitz, G., Sandercock, P. and Counsell, C. (2003). *Cochrane Database of Systematic Reviews* **1:** 1.

Hademenos, G. J., Alberts, M. J., Awad, I., Mayberg, M., Shepard, T., Jagoda, A., Latchaw, R. E., Todd, H. W., Viste, K., Starke, R., Girgus, M. S., Marler, J., Emr, M. and Hart, N. (2001). Advances in the genetics of cerebrovascular disease and stroke. *Neurology* **56:** 997-1008.

Hankey, G. J., Eikelboom, J. W., van Bockxmeer, F. M., Lofthouse, E., Staples, N. and Baker, R. I. (2001). Inherited thrombophilia in ischemic stroke and its pathogenic subtypes. *Stroke.* **32:** 1793-9.

Hankey, G. J., Jamrozik, K., Broadhurst, R. J., Forbes, S. and Anderson, C. S. (2002). Long-term disability after first-ever stroke and related prognostic factors in the Perth Community Stroke Study, 1989-1990. *Stroke.* **33:** 1034-40.

Hassan, A. and Markus, H. S. (2000). Genetics and ischaemic stroke. *Brain* **123:** 1784-812.

Hassan, A., Sham, P. C. and Markus, H. S. (2002). Planning genetic studies in human stroke: sample size estimates based on family history data. *Neurology.* **58:** 1483-8.

Hatono, S. (1976). Experience from a multicentre stroke register: a preliminary report. *Bulletin WHO.* **54:** 541-553.

Herman, B., Schmitz, P. I., Leyten, A. C., Van Luijk, J. H., Frenken, C. W., Op De Coul, A. A. and Schulte, B. P. (1983). Multivariate logistic analysis of risk factors for stroke in Tilburg, The Netherlands. *American Journal of Epidemiology.* **118:** 514-25.

Hindorff, L. A., Schwartz, S. M., Siscovick, D. S., Psaty, B. M., Longstreth, W. T., Jr. and Reiner, A. P. (2002). The association of PAI-1 promoter 4G/5G insertion/deletion polymorphism with myocardial infarction and stroke in young women. *Journal of Cardiovascular Risk.* **9:** 131-7.

Humphries, S. E., Cook, M., Dubowitz, M., Stirling, Y. and Meade, T. W. (1987). Role of genetic variation at the fibrinogen locus in determination of plasma fibrinogen concentrations. *Lancet.* **1:** 1452-5.

Imai, Y., Morita, H., Kurihara, H., Sugiyama, T., Kato, N., Ebihara, A., Hamada, C., Kurihara, Y., Shindo, T., Oh-hashi, Y. and Yazaki, Y. (2000). Evidence for association between paraoxonase gene polymorphisms and atherosclerotic diseases. *Atherosclerosis* **149:** 435-42.

Ioannidis, J. P., Trikalinos, T. A., Ntzani, E. E. and Contopoulos-Ioannidis, D. G. (2003). Genetic associations in large versus small studies: an empirical assessment. *Lancet.* **361:** 567-71.

Iso, H., Folsom, A. R., Winkelmann, J. C., Koike, K., Harada, S., Greenberg, B., Sato, S., Shimamoto, T., Iida, M. and Komachi, Y. (1995). Polymorphisms of the beta fibrinogen gene and plasma fibrinogen concentration in Caucasian and Japanese population samples. *Thrombosis & Haemostasis.* **73:** 106-11.

James, R. W., Leviev, I., Ruiz, J., Passa, P., Froguel, P. and Garin, M. C. (2000). Promoter polymorphism T(-107)C of the paraoxonase PON1 gene is a risk factor for coronary heart disease in type 2 diabetic patients. *Diabetes* **49:** 1390-3.

Jamrozik, K., Broadhurst, R. J., Anderson, C. S. and Stewart-Wynne, E. G. (1994). The role of lifestyle factors in the etiology of stroke. A population-based case-control study in Perth, Western Australia. *Stroke.* **25:** 51-9.

Jamrozik, K., Broadhurst, R. J., Lai, N., Hankey, G. J., Burvill, P. W. and Anderson, C. S. (1999). Trends in the incidence, severity, and short-term outcome of stroke in perth, Western Australia. *Stroke.* **30:** 2105-11.

Jern, C., Ladenvall, P., Wall, U. and Jern, S. (1999). Gene polymorphism of t-PA is associated with forearm vascular release rate of t-PA. *Arteriosclerosis, Thrombosis & Vascular Biology.* **19:** 454-9.

Jerrard-Dunne, P., Cloud, G., Hassan, A. and Markus, H. S. (2003). Evaluating the genetic component of ischemic stroke subtypes: a family history study. *Stroke.* **34:** 1364-9.

Johansson, L., Jansson, J. H., Boman, K., Nilsson, T. K., Stegmayr, B. and Hallmans, G. (2000). Tissue plasminogen activator, plasminogen activator inhibitor-1, and tissue plasminogen activator/plasminogen activator inhibitor-1 complex as risk factors for the development of a first stroke. *Stroke.* **31:** 26-32.

Jousilahti, P., Rastenyte, D., Tuomilehto, J., Sarti, C. and Vartiainen, E. (1997). Parental history of cardiovascular disease and risk of stroke. A prospective follow-up of 14371 middle-aged men and women in Finland. *Stroke.* **28:** 1361-6.

Kain, K., Catto, A. J., Carter, A. M., Young, J., Bamford, J., Bavington, J. and Grant, P. J. (2001). Decreased fibrinolytic potential in South Asian women with ischaemic cerebrovascular disease. *British Journal of Haematology.* **114:** 155-61.

Kain, K., Young, J., Bamford, J., Bavington, J., Grant, P. J. and Catto, A. J. (2002). Determinants of plasminogen activator inhibitor-1 in South Asians with ischaemic stroke. *Cerebrovascular Diseases.* **14:** 77-83.

Kao, Y. L., Donaghue, K., Chan, A., Knight, J. and Silink, M. (1998). A variant of paraoxonase (PON1) gene is associated with diabetic retinopathy in IDDM. *Journal of Clinical Endocrinology & Metabolism.* **83:** 2589-92.

Kekomaki, S., Hamalainen, L., Kauppinen-Makelin, R., Palomaki, H., Kaste, M. and Kontula, K. (1999). Genetic polymorphism of platelet glycoprotein IIIa in patients with acute myocardial infarction and acute ischaemic stroke. *Journal of Cardiovascular Risk.* **6:** 13-7.

Kessler, C., Spitzer, C., Stauske, D., Mende, S., Stadlmuller, J., Walther, R. and Rettig, R. (1997). The apolipoprotein E and beta-fibrinogen G/A-455 gene polymorphisms are associated with ischemic stroke involving large-vessel disease. *Arteriosclerosis, Thrombosis & Vascular Biology.* **17:** 2880-4.

Khaw, K. T. and Barrett-Connor, E. (1986). Family history of stroke as an independent predictor of ischemic heart disease in men and stroke in women. *American Journal of Epidemiology.* **123:** 59-66.

Kiely, D. K., Wolf, P. A., Cupples, L. A., Beiser, A. S. and Kannel, W. B. (1994). Physical activity and stroke risk: the Framingham Study.[erratum appears in Am J Epidemiol 1995 Jan 15;141(2):178]. *American Journal of Epidemiology.* **140:** 608-20.

Kiely, D. K., Wolf, P. A., Cupples, L. A., Beiser, A. S. and Myers, R. H. (1993). Familial aggregation of stroke. The Framingham Study. *Stroke.* **24:** 1366-71.

Kilpatrick, T. J., Matkovic, Z., Davis, S. M., McGrath, C. M. and Dauer, R. J. (1993). Hematologic abnormalities occur in both cortical and lacunar infarction. *Stroke.* **24:** 1945-50.

Koenig, W. (2003). Fibrin(ogen) in cardiovascular disease: an update. *Thrombosis & Haemostasis.* **89:** 601-9.

Kofoed, S. C., Wittrup, H. H., Sillesen, H. and Nordestgaard, B. G. (2003). Fibrinogen predicts ischaemic stroke and advanced atherosclerosis but not echolucent, rupture-prone carotid plaques. The Copenhagen City Heart Study. *European Heart Journal.* **24:** 567-76.

Kolominsky-Rabas, P. L., Weber, M., Gefeller, O., Neundoerfer, B. and Heuschmann, P. U. (2001). Epidemiology of ischemic stroke subtypes according to TOAST criteria: incidence, recurrence, and long-term survival in ischemic stroke subtypes: a population-based study. *Stroke.* **32:** 2735-40.

Konstantopoulos, K., Grotta, J. C., Sills, C., Wu, K. K. and Hellums, J. D. (1995). Shear-induced platelet aggregation in normal subjects and stroke patients. *Thrombosis & Haemostasis.* **74:** 1329-34.

Kooistra, T., Schrauwen, Y., Arts, J. and Emeis, J. J. (1994). Regulation of endothelial cell t-PA synthesis and release. *International Journal of Hematology.* **59:** 233-55.

Kristensen, B., Malm, J., Nilsson, T. K., Hultdin, J., Carlberg, B., Dahlen, G. and Olsson, T. (1999). Hyperhomocysteinemia and hypofibrinolysis in young adults with ischemic stroke. *Stroke.* **30:** 974-80.

Kristensen, B., Malm, J., Nilsson, T. K., Hultdin, J., Carlberg, B. and Olsson, T. (1998). Increased fibrinogen levels and acquired hypofibrinolysis in young adults with ischemic stroke. *Stroke.* **29:** 2261-7.

Kuijpers, R. W., Faber, N. M., Cuypers, H. T., Ouwehand, W. H. and von dem Borne, A. E. (1992). NH2-terminal globular domain of human platelet glycoprotein Ib alpha has a methionine 145/threonine145 amino acid polymorphism, which is associated with the HPA-2 (Ko) alloantigens. *Journal of Clinical Investigation.* **89:** 381-4.

Kullo, I. J., Edwards, W. D. and Schwartz, R. S. (1998). Vulnerable plaque: pathobiology and clinical implications. *Annals of Internal Medicine.* **129:** 1050-60.

Ladenvall, P., Johansson, L., Jansson, J. H., Jern, S., Nilsson, T. K., Tjarnlund, A., Jern, C. and Boman, K. (2002). Tissue-type plasminogen activator -7,351C/T enhancer polymorphism is associated with a first myocardial infarction. *Thrombosis & Haemostasis.* **87:** 105-9.

Ladenvall, P., Wall, U., Jern, S. and Jern, C. (2000). Identification of eight novel single-nucleotide polymorphisms at human tissue-type plasminogen activator (t-PA) locus: association with vascular t-PA release in vivo. *Thrombosis & Haemostasis.* **84:** 150-5.

Ladenvall, P., Wall, U., Jern, S. and Jern, C. (2001). Eight single-nucleotide polymorphisms (SNPs) at the human tissue-type plasminogen activator (t-PA) locus. *Journal of Human Genetics.* **46:** 737-8.

Lander, E. S. and Schork, N. J. (1994). Genetic dissection of complex traits. *Science.* **265:** 2037-48.

Leus, F. R., Wittekoek, M. E., Prins, J., Kastelein, J. J. and Voorbij, H. A. (2000). Paraoxonase gene polymorphisms are associated with carotid arterial wall thickness in subjects with familial hypercholesterolemia. *Atherosclerosis.* **149:** 371-7.

Leviev, I. and James, R. W. (2000). Promoter polymorphisms of human paraoxonase PON1 gene and serum paraoxonase activities and concentrations. *Arteriosclerosis, Thrombosis & Vascular Biology.* **20:** 516-21.

Leviev, I., Righetti, A. and James, R. W. (2001). Paraoxonase promoter polymorphism T(-107)C and relative paraoxonase deficiency as determinants of risk of coronary artery disease. *Journal of Molecular Medicine.* **79:** 457-63.

Levin, E. G. and del Zoppo, G. J. (1994). Localization of tissue plasminogen activator in the endothelium of a limited number of vessels. *American Journal of Pathology.* **144:** 855-61.

Liao, D., Myers, R., Hunt, S., Shahar, E., Paton, C., Burke, G., Province, M. and Heiss, G. (1997). Familial history of stroke and stroke risk. The Family Heart Study. *Stroke.* **28:** 1908-12.

Lichy, C., Reuner, K. H., Buggle, F., Litfin, F., Rickmann, H., Kunze, A., Brandt, T. and Grau, A. (2003). Prothrombin G20210A mutation, but not factor V Leiden, is a risk factor in patients with persistent foramen ovale and otherwise unexplained cerebral ischemia. *Cerebrovascular Diseases.* **16:** 83-7.

Lindgren, A., Lindoff, C., Norrving, B., Astedt, B. and Johansson, B. B. (1996). Tissue plasminogen activator and plasminogen activator inhibitor-1 in stroke patients. *Stroke.* **27:** 1066-71.

Liu, Y., Pan, J., Wang, S., Li, X. and Huang, Y. (2002). beta-fibrinogen gene -455A/G polymorphism and plasma fibrinogen level in Chinese stroke patients. *Chinese Medical Journal.* **115:** 214-6.

Longstreth, W. T., Jr., Rosendaal, F. R., Siscovick, D. S., Vos, H. L., Schwartz, S. M., Psaty, B. M., Raghunathan, T. E., Koepsell, T. D. and Reitsma, P. H. (1998). Risk of stroke in young women and two prothrombotic mutations: factor V Leiden and prothrombin gene variant (G20210A). *Stroke.* **29:** 577-80.

Lopaciuk, S., Bykowska, K., Kwiecinski, H., Mickielewicz, A., Czlonkowska, A., Mendel, T., Kuczynska-Zardzewialy, A., Szelagowska, D., Windyga, J., Schroder, W., Herrmann, F. H. and Jedrzejowska, H. (2001). Factor V Leiden, prothrombin gene G20210A variant, and methylenetetrahydrofolate reductase C677T genotype in young adults with ischemic stroke. *Clinical & Applied Thrombosis/Hemostasis.* **7:** 346-50.

Mackness, B., Durrington, P. N. and Mackness, M. I. (1999). Polymorphisms of paraoxonase genes and low-density lipoprotein lipid peroxidation. *Lancet.* **353:** 468-9.

Mackness, B., Durrington, P. N. and Mackness, M. I. (2002). The paraoxonase gene family and coronary heart disease. *Current Opinion in Lipidology.* **13:** 357-62.

Mackness, B., Hunt, R., Durrington, P. N. and Mackness, M. I. (1997). Increased immunolocalization of paraoxonase, clusterin, and apolipoprotein A-I in the human artery wall with the progression of atherosclerosis. *Arteriosclerosis, Thrombosis & Vascular Biology.* **17:** 1233-8.

Mackness, B., Mackness, M. I., Arrol, S., Turkie, W. and Durrington, P. N. (1998). Effect of the human serum paraoxonase 55 and 192 genetic polymorphisms on the protection by high density lipoprotein against low density lipoprotein oxidative modification. *FEBS Letters.* **423:** 57-60.

Mackness, M. I. and Durrington, P. N. (1995). Paraoxonase: another factor in NIDDM cardiovascular disease. *Lancet.* **346:** 856.

Macko, R. F., Kittner, S. J., Epstein, A., Cox, D. K., Wozniak, M. A., Wityk, R. J., Stern, B. J., Sloan, M. A., Sherwin, R., Price, T. R., McCarter, R. J., Johnson, C. J., Earley, C. J., Buchholz, D. W. and Stolley, P. D. (1999). Elevated tissue plasminogen activator antigen and stroke risk: The Stroke Prevention In Young Women Study. *Stroke.* **30:** 7-11.

Malin, R., Jarvinen, O., Sisto, T., Koivula, T. and Lehtimaki, T. (2001). Paraoxonase producing PON1 gene M/L55 polymorphism is related to autopsy-verified artery-wall atherosclerosis. *Atherosclerosis.* **157:** 301-7.

Margaglione, M., Cappucci, G., d'Addedda, M., Colaizzo, D., Giuliani, N., Vecchione, G., Mascolo, G., Grandone, E. and Di Minno, G. (1998). PAI-1 plasma levels in a general population without clinical evidence of atherosclerosis: relation to environmental and genetic determinants. *Arteriosclerosis, Thrombosis & Vascular Biology.* **18:** 562-7.

Margaglione, M., Di Minno, G., Grandone, E., Vecchione, G., Celentano, E., Cappucci, G., Grilli, M., Simone, P., Panico, S. and Mancini, M. (1994). Abnormally high circulation levels of tissue plasminogen activator and plasminogen activator inhibitor-1 in patients with a history of ischemic stroke. *Arteriosclerosis & Thrombosis.* **14:** 1741-5.

Martiskainen, M., Pohjasvaara, T., Mikkelsson, J., R, M. A., Kunnas, T., Laippala, P., Ilveskoski, E., Kaste, M., Karhunen, P. J. and Erkinjuntti, T. (2003). Fibrinogen gene promoter -455 A allele as a risk factor for lacunar stroke. *Stroke.* **34:** 886-91.

Mas, J. L., Arquizan, C., Lamy, C., Zuber, M., Cabanes, L., Derumeaux, G., Coste, J., Patent Foramen, O. and Atrial Septal Aneurysm Study, G. (2001). Recurrent cerebrovascular events associated with patent foramen ovale, atrial septal aneurysm, or both. *New England Journal of Medicine.* **345:** 1740-6.

Masuda, J., Tanaka, K., Omae, T., Ueda, K. and Sadoshima, S. (1983). Cerebrovascular diseases and their underlying vascular lesions in Hisayama, Japan--a pathological study of autopsy cases. *Stroke.* **14:** 934-40.

Mathers, C. D., Vos, E. T., Stevenson, C. E. and Begg, S. J. (2000). The Australian Burden of Disease Study: measuring the loss of health from diseases, injuries and risk factors. *Medical Journal of Australia.* **172:** 592-6.

Matsuno, H., Kozawa, O., Niwa, M., Ueshima, S., Matsuo, O., Collen, D. and Uematsu, T. (1999). Differential role of components of the fibrinolytic system in the formation and removal of thrombus induced by endothelial injury. *Thrombosis & Haemostasis.* **81:** 601-4.

Mead, G. E., Lewis, S. C., Wardlaw, J. M., Dennis, M. S. and Warlow, C. P. (2000a). How well does the Oxfordshire community stroke project classification predict the site and size of the infarct on brain imaging?[comment]. *Journal of Neurology, Neurosurgery & Psychiatry.* **68:** 558-62.

Mead, G. E., Wardlaw, J. M., Dennis, M. S., Lewis, S. C. and Warlow, C. P. (2000b). Relationship between pattern of intracranial artery abnormalities on transcranial doppler and Oxfordshire Community Stroke Project clinical classification of ischemic stroke. *Stroke.* **31:** 714-9.

Meiklejohn, D. J., Vickers, M. A., Morrison, E. R., Dijkhuisen, R., Moore, I., Urbaniak, S. J. and Greaves, M. (2001). In vivo platelet activation in atherothrombotic stroke is not determined by polymorphisms of human platelet glycoprotein IIIa or Ib. *British Journal of Haematology* **112:** 621-31.

Mikkelsson, J., Perola, M., Wartiovaara, U., Peltonen, L., Palotie, A., Penttila, A. and Karhunen, P. J. (2000). Plasminogen activator inhibitor-1 (PAI-1) 4G/5G polymorphism, coronary thrombosis, and myocardial infarction in middle-aged Finnish men who died suddenly. *Thrombosis & Haemostasis.* **84:** 78-82.

Miller, S. A., Dykes, D. D. and Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research.* **16:** 1215.

Millikan, C. and Futrell, N. (1990). The fallacy of the lacune hypothesis. *Stroke.* **21:** 1251-7.

Mohr, J. P., Albers, G. W., Amarenco, P., Babikian, V. L., Biller, J., Brey, R. L., Coull, B., Easton, J. D., Gomez, C. R., Helgason, C. M., Kase, C. S., Pullicino, P. M. and Turpie, A. G. (1997). American Heart Association Prevention Conference. IV. Prevention and Rehabilitation of Stroke. Etiology of stroke. *Stroke.* **28:** 1501-6.

Murray, C. J. and Lopez, A. D. (1997a). Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet.* **349:** 1498-504.

Murray, C. J. and Lopez, A. D. (1997b). Mortality by cause for eight regions of the world: Global Burden of Disease Study. *Lancet.* **349:** 1269-76.

Newby, D. E., Wright, R. A., Labinjoh, C., Ludlam, C. A., Fox, K. A., Boon, N. A. and Webb, D. J. (1999). Endothelial dysfunction, impaired endogenous fibrinolysis, and cigarette smoking: a mechanism for arterial thrombosis and myocardial infarction. *Circulation.* **99:** 1411-5.

Newman, P. J., Derbes, R. S. and Aster, R. H. (1989). The human platelet alloantigens, PlA1 and PlA2, are associated with a leucine33/proline33 amino acid polymorphism in membrane glycoprotein IIIa, and are distinguishable by DNA typing. *Journal of Clinical Investigation.* **83:** 1778-81.

Nguyen, A. (2000). Prothrombin G20210A polymorphism and thrombophilia. *Mayo Clinic Proceedings* **75:** 595-604.

National Heart Foundation (2001). National Centre for Monitoring Cardiovascular Disease.

Nishiuma, S., Kario, K., Yakushijin, K., Maeda, M., Murai, R., Matsuo, T., Ikeda, U., Shimada, K. and Matsuo, M. (1998). Genetic variation in the promoter region of the beta-fibrinogen gene is associated with ischemic stroke in a Japanese population. *Blood Coagulation & Fibrinolysis.* **9:** 373-9.

O'Donnell, C. J., Larson, M. G., Feng, D., Sutherland, P. A., Lindpaintner, K., Myers, R. H., D'Agostino, R. A., Levy, D., Tofler, G. H. and Framingham Heart, S. (2001). Genetic and environmental contributions to platelet aggregation: the Framingham heart study. *Circulation.* **103:** 3051-6.

Okamoto, K., Yamori, Y. and Nagaoka, A. (1974). Establishment of the stroke prone spontaneously hypertensive rat (SHR). *Circulation Research* **33/34:** 1-143 - 1-153.

Pongracz, E., Tordai, A., Csornai, M. and Nagy, Z. (2001). [Platelet glycoprotein IIb/IIIa (LeuPro 33) polymorphism in stroke patients]. *Orvosi Hetilap.* **142:** 781-5 (abstract).

Poort, S. R., Rosendaal, F. R., Reitsma, P. H. and Bertina, R. M. (1996). 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* **88:** 3698-703.

Qizilbash, N., Jones, L., Warlow, C. and Mann, J. (1991). Fibrinogen and lipid concentrations as risk factors for transient ischaemic attacks and minor ischaemic strokes. *BMJ.* **303:** 605-9.

Razdan, K., Hellums, J. D. and Kroll, M. H. (1994). Shear-stress-induced von Willebrand factor binding to platelets causes the activation of tyrosine kinase(s). *Biochemical Journal.* **302:** 681-6.

Reich, D. E. and Goldstein, D. B. (2001). Detecting association in a case-control study while correcting for population stratification. *Genetic Epidemiology.* **20:** 4-16.

Reiner, A. P., Kumar, P. N., Schwartz, S. M., Longstreth, W. T., Jr., Pearce, R. M., Rosendaal, F. R., Psaty, B. M. and Siscovick, D. S. (2000). Genetic variants of platelet glycoprotein receptors and risk of stroke in young women. *Stroke.* **31:** 1628-33.

Resch, K. L., Ernst, E., Matrai, A. and Paulsen, H. F. (1992). Fibrinogen and viscosity as risk factors for subsequent cardiovascular events in stroke survivors. *Annals of Internal Medicine.* **117:** 371-5.

Ridker, P. M. (1994). Plasma Concentration of Endogenous Tissue Plasminogen Activator and the Occurrence of Future Cardiovascular Events. *Journal of Thrombosis & Thrombolysis. 1(1):35-40, 1994.*

Ridker, P. M., Hennekens, C. H. and Miletich, J. P. (1999). G20210A mutation in prothrombin gene and risk of myocardial infarction, stroke, and venous thrombosis in a large cohort of US men. *Circulation* **99:** 999-1004.

Ridker, P. M., Hennekens, C. H., Schmitz, C., Stampfer, M. J. and Lindpaintner, K. (1997). PIA1/A2 polymorphism of platelet glycoprotein IIIa and risks of myocardial infarction, stroke, and venous thrombosis. *Lancet.* **349:** 385-8.

Risch, N. and Teng, J. (1998). The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases I. DNA pooling. *Genome Research.* **8:** 1273-88.

Risch, N. J. (2000). Searching for genetic determinants in the new millennium. *Nature.* **405:** 847-56.

Roest, M., van der Schouw, Y. T., Banga, J. D., Tempelman, M. J., de Groot, P. G., Sixma, J. J. and Grobbee, D. E. (2000). Plasminogen activator inhibitor 4G polymorphism is associated with decreased risk of cerebrovascular mortality in older women. *Circulation.* **101:** 67-70.

Rosenberg, R. D. and Aird, W. C. (1999). Vascular-bed--specific hemostasis and hypercoagulable states. *New England Journal of Medicine.* **340:** 1555-64.

Roy, S. N., Mukhopadhyay, G. and Redman, C. M. (1990). Regulation of fibrinogen assembly. Transfection of Hep G2 cells with B beta cDNA specifically enhances synthesis of the three component chains of fibrinogen. *Journal of Biological Chemistry.* **265:** 6389-93.

Rubattu, S., Volpe, M., Kreutz, R., Ganten, U., Ganten, D. and Lindpaintner, K. (1996). Chromosomal mapping of quantitative trait loci contributing to stroke in a rat model of complex human disease. *Nature Genetics.* **13:** 429-34.

Salonen, J. T., Malin, R., Tuomainen, T. P., Nyyssonen, K., Lakka, T. A. and Lehtimaki, T. (1999). Polymorphism in high density lipoprotein paraoxonase gene and risk of acute myocardial infarction in men: prospective nested case-control study. *BMJ.* **319:** 487-9.

Savage, B., Saldivar, E. and Ruggeri, Z. M. (1996). Initiation of platelet adhesion by arrest onto fibrinogen or translocation on von Willebrand factor. *Cell.* **84:** 289-97.

Scarabin, P. Y., Bara, L., Ricard, S., Poirier, O., Cambou, J. P., Arveiler, D., Luc, G., Evans, A. E., Samama, M. M. and Cambien, F. (1993). Genetic variation at the beta-fibrinogen locus in relation to plasma fibrinogen concentrations and risk of myocardial infarction. The ECTIM Study. *Arteriosclerosis Thrombosis* **13:** 886-91.

Schmidt, H., Schmidt, R., Niederkorn, K., Gradert, A., Schumacher, M., Watzinger, N., Hartung, H. P. and Kostner, G. M. (1998a). Paraoxonase PON1 polymorphism leu-Met54 is associated with carotid atherosclerosis: results of the Austrian Stroke Prevention Study. *Stroke* **29:** 2043-8.

Schmidt, H., Schmidt, R., Niederkorn, K., Horner, S., Becsagh, P., Reinhart, B., Schumacher, M., Weinrauch, V. and Kostner, G. M. (1998b). Beta-fibrinogen gene polymorphism (C148-->T) is associated with carotid atherosclerosis: results of the Austrian Stroke Prevention Study. *Arteriosclerosis, Thrombosis & Vascular Biology* **18:** 487-92.

Schmidt, R., Schmidt, H., Fazekas, F., Kapeller, P., Roob, G., Lechner, A., Kostner, G. M. and Hartung, H. P. (2000). MRI cerebral white matter lesions and paraoxonase PON1 polymorphisms : three-year follow-up of the austrian stroke prevention study. *Arteriosclerosis, Thrombosis & Vascular Biology* **20:** 1811-6.

Schreiber, S. S., Tan, Z., Sun, N., Wang, L. and Zlokovic, B. V. (1998). Immunohistochemical localization of tissue plasminogen activator in vascular endothelium of stroke-prone regions of the rat brain. *Neurosurgery.* **43:** 909-13.

Shah, A. B., Beamer, N. and Coull, B. M. (1985). Enhanced in vivo platelet activation in subtypes of ischemic stroke. *Stroke.* **16:** 643-7.

Shah, P. K., Falk, E., Badimon, J. J., Fernandez-Ortiz, A., Mailhac, A., Villareal-Levy, G., Fallon, J. T., Regnstrom, J. and Fuster, V. (1995). Human monocyte-derived macrophages induce collagen breakdown in fibrous caps of atherosclerotic plaques. Potential role of matrix-degrading metalloproteinases and implications for plaque rupture. *Circulation.* **92:** 1565-9.

Shih, D. M., Gu, L., Xia, Y. R., Navab, M., Li, W. F., Hama, S., Castellani, L. W., Furlong, C. E., Costa, L. G., Fogelman, A. M. and Lusis, A. J. (1998). Mice lacking serum paraoxonase are susceptible to organophosphate toxicity and atherosclerosis. *Nature.* **394:** 284-7.

Simons, L. A., McCallum, J., Friedlander, Y. and Simons, J. (1998). Risk factors for ischemic stroke: Dubbo Study of the elderly. *Stroke.* **29:** 1341-6.

Simsek, S., Bleeker, P. M., van der Schoot, C. E. and von dem Borne, A. E. (1994). Association of a variable number of tandem repeats (VNTR) in glycoprotein Ib alpha and HPA-2 alloantigens. *Thrombosis & Haemostasis.* **72:** 757-61.

Smith, F. B., Rumley, A., Lee, A. J., Leng, G. C., Fowkes, F. G. and Lowe, G. D. (1998). Haemostatic factors and prediction of ischaemic heart disease and stroke in claudicants. *British Journal of Haematology.* **100:** 758-63.

Smith, N., Pathansali, R. and Bath, P. (1997). PlA1/A2 polymorphism of platelet glycoprotein IIIa and risk of cardiovascular disease. *Lancet.* **349:** 1099.

Sonoda, A., Murata, M., Ito, D., Tanahashi, N., Ohta, A., Tada, Y., Takeshita, E., Yoshida, T., Saito, I., Yamamoto, M., Ikeda, Y., Fukuuchi, Y. and Watanabe, K. (2000). Association between platelet glycoprotein Ibalpha genotype and ischemic cerebrovascular disease. *Stroke* **31:** 493-7.

Streifler, J. Y., Rosenberg, N., Chetrit, A., Eskaraev, R., Sela, B. A., Dardik, R., Zivelin, A., Ravid, B., Davidson, J., Seligsohn, U. and Inbal, A. (2001). Cerebrovascular events in patients with significant stenosis of the carotid artery are associated with hyperhomocysteinemia and platelet antigen-1 (Leu33Pro) polymorphism. *Stroke.* **32:** 2753-8.

Szolnoki, Z., Somogyvari, F., Kondacs, A., Szabo, M. and Fodor, L. (2002). Evaluation of the interactions of common genetic mutations in stroke subtypes. *Journal of Neurology.* **249:** 1391-7.

Thomas, A., Lamlum, H., Humphries, S. and Green, F. (1994). Linkage disequilibrium across the fibrinogen locus as shown by five genetic polymorphisms, G/A-455 (HaeIII), C/T-148 (HindIII/AluI), T/G+1689 (AvaII), and BclI (beta-fibrinogen) and TaqI (alpha-fibrinogen), and their detection by PCR. *Human Mutation.* **3:** 79-81.

Thomas, A. E., Green, F. R., Kelleher, C. H., Wilkes, H. C., Brennan, P. J., Meade, T. W. and Humphries, S. E. (1991). Variation in the promoter region of the beta fibrinogen gene is associated with plasma fibrinogen levels in smokers and non-smokers. *Thrombosis & Haemostasis.* **65:** 487-90.

Thrift, A. G., Dewey, H. M., Macdonell, R. A., McNeil, J. J. and Donnan, G. A. (2000). Stroke incidence on the east coast of Australia: the North East Melbourne Stroke Incidence Study (NEMESIS). *Stroke.* **31:** 2087-92.

Thrift, A. G., Dewey, H. M., Macdonell, R. A., McNeil, J. J. and Donnan, G. A. (2001). Incidence of the major stroke subtypes: initial findings from the North East Melbourne stroke incidence study (NEMESIS). *Stroke.* **32:** 1732-8.

Tybjaerg-Hansen, A., Agerholm-Larsen, B., Humphries, S. E., Abildgaard, S., Schnohr, P. and Nordestgaard, B. G. (1997). A common mutation (G-455--> A) in the beta-fibrinogen promoter is an independent predictor of plasma fibrinogen, but not of ischemic heart disease. A study of 9,127 individuals based on the Copenhagen City Heart Study. *Journal of Clinical Investigation* **99:** 3034-9.

Uchiyama, S., Takeuchi, M., Osawa, M., Kobayashi, I., Maruyama, S., Aosaki, M. and Hirosawa, K. (1983). Platelet function tests in thrombotic cerebrovascular disorders. *Stroke.* **14:** 511-7.

Uchiyama, S., Yamazaki, M., Maruyama, S., Handa, M., Ikeda, Y., Fukuyama, M. and Itagaki, I. (1994). Shear-induced platelet aggregation in cerebral ischemia. *Stroke.* **25:** 1547-51.

Underwood, M. J. and De Bono, D. P. (1993). Increased fibrinolytic activity in the intima of atheromatous coronary arteries: protection at a price. *Cardiovascular Research.* **27:** 882-5.

van Goor, M. L., E, G. A. G.-a., Brouwers, G. J., Leebeek, F. W., Koudstaal, P. J. and Dippel, D. W. (2002). PLA1/A2 polymorphism of the platelet glycoprotein receptor IIb/IIIa in young patients with cryptogenic TIA or ischemic stroke. *Thrombosis Research* **108:** 63-5.

Vijayan, K. V., Goldschmidt-Clermont, P. J., Roos, C. and Bray, P. F. (2000). The PlA2 polymorphism of integrin {beta}3 enhances outside-in signaling and adhesive functions. *Journal of Clinical Investigation.* **105:** 793-802.

Vijayan, K. V., Huang, T. C., Liu, Y., Bernardo, A., Dong, J. F., Goldschmidt-Clermont, P. J., Alevriadou, B. R. and Bray, P. F. (2003). Shear stress augments the enhanced adhesive phenotype of cells expressing the Pro33 isoform of integrin beta(3). *FEBS Letters* **540:** 41-6.

Voetsch, B., Benke, K. S., Damasceno, B. P., Siqueira, L. H. and Loscalzo, J. (2002). Paraoxonase 192 Gln-->Arg polymorphism: an independent risk factor for nonfatal arterial ischemic stroke among young adults. *Stroke.* **33:** 1459-64.

Wagner, K. R., Giles, W. H., Johnson, C. J., Ou, C. Y., Bray, P. F., Goldschmidt-Clermont, P. J., Croft, J. B., Brown, V. K., Stern, B. J., Feeser, B. R., Buchholz, D. W., Earley, C. J., Macko, R. F., McCarter, R. J., Sloan, M. A., Stolley, P. D., Wityk, R. J., Wozniak, M. A., Price, T. R. and Kittner, S. J. (1998). Platelet glycoprotein receptor IIIa polymorphism P1A2 and ischemic stroke risk: the Stroke Prevention in Young Women Study. *Stroke.* **29:** 581-5.

Wardlaw, J. M., Sandercock, P. A., Dennis, M. S. and Starr, J. (2003). Is breakdown of the blood-brain barrier responsible for lacunar stroke, leukoaraiosis, and dementia? *Stroke.* **34:** 806-12.

Watson, A. D., Berliner, J. A., Hama, S. Y., La Du, B. N., Faull, K. F., Fogelman, A. M. and Navab, M. (1995). Protective effect of high density lipoprotein associated paraoxonase. Inhibition of the biological activity of minimally oxidized low density lipoprotein. *Journal of Clinical Investigation.* **96:** 2882-91.

Weiss, E. J., Bray, P. F., Tayback, M., Schulman, S. P., Kickler, T. S., Becker, L. C., Weiss, J. L., Gerstenblith, G. and Goldschmidt-Clermont, P. J. (1996). A polymorphism of a platelet glycoprotein receptor as an inherited risk factor for coronary thrombosis. *New England Journal of Medicine.* **334:** 1090-4.

Welin, L., Svardsudd, K., Wilhelmsen, L., Larsson, B. and Tibblin, G. (1987). Analysis of risk factors for stroke in a cohort of men born in 1913. *New England Journal of Medicine.* **317:** 521-6.

Welsh, K. and Bunce, M. (1999). Molecular typing for the MHC with PCR-SSP. *Reviews in Immunogenetics.* **1:** 157-76.

Whisnant, J. P. (1990). Special report from the National Institute of Neurological Disorders and Stroke. Classification of cerebrovascular diseases III. *Stroke.* **21:** 637-76.

Whisnant, J. P. (1997). Modeling of risk factors for ischemic stroke. The Willis Lecture. *Stroke.* **28:** 1840-4.

Whisnant, J. P., Wiebers, D. O., O'Fallon, W. M., Sicks, J. D. and Frye, R. L. (1996). A population-based model of risk factors for ischemic stroke: Rochester, Minnesota. *Neurology.* **47:** 1420-8.

WHO (2002).World Health Organization, Geneva.

Wolf, P. A., D'Agostino, R. B., Kannel, W. B., Bonita, R. and Belanger, A. J. (1988). Cigarette smoking as a risk factor for stroke. The Framingham Study. *JAMA.* **259:** 1025-9.

Wu, K. K., Folsom, A. R., Heiss, G., Davis, C. E., Conlan, M. G. and Barnes, R. (1992). Association of coagulation factors and inhibitors with carotid artery atherosclerosis. Early results of the Atherosclerosis Risk in Communities (ARIC) Study. *Annals of Epidemiology.* **2:** 471-80.

Zhang, C., Li, J., Li, L. and Luo, B. (2001). The study of PAI-1 promotor region gene polymorphism in cerebrovascular disease. *Chung-Hua i Hsueh i Chuan Hsueh Tsa Chih.* **18:** 383-7 (abstract).

Appendix I

The OCSP Classification Criteria (Bamford *et al.* 1991)

Total Anterior Circulation Syndrome (TACS)

A combination of new higher cerebral dysfunction (eg. dysphasia, dyscalculia, visuospatial disorder); homonymous visual field defect; and ipsilateral motor and/or sensory deficit of at least two areas of the face, arm or leg. If consciousness is impaired and formal testing of higher cerebral function or visual fields was not possible, a deficit was assumed.

Partial Anterior Circulation Syndrome (PACS)

Two out of three components of the TACS, with higher cerebral dysfunction alone, or with a motor/sensory deficit more restricted than those classified as a LS.

Lacunar Syndrome (LS)

A pure motor stroke, pure sensory stroke, sensori-motor stroke, or ataxic hemiparesis.

Posterior Circulation Syndrome (PCS)

Any of the following: ipsilateral cranial nerve palsy, with contralateral motor and/or sensory deficit; bilateral motor and/or sensory deficit; disorder of conjugate eye movement; cerebellar dysfunction without ipsilateral long-tract deficit (i.e. ataxic hemiparesis); or isolated homonymous visual field defect.

Appendix II

Study Information Sheet

Polygenic Disease – A Study of Genetic Risk in an Australian Stroke Population
The Adelaide Genetic Stroke Study

Principal Investigator: Dr Jim Jannes
We are grateful for your time to consider participation in this study. We welcome your interest and would like to inform you of our aims.

Background Information and Aims

Stroke is a common disease that affects thousands of Australians each year. Smoking, high blood pressure, diabetes and a high cholesterol level are important factors that increase the risk of stroke. Recent studies have indicated that there may also be a genetic contribution to stroke.

Most strokes occur as a result of a blood clot blocking a major blood vessel within the brain. This process involves many proteins, including platelet glycoproteins, fibrinogen, paraxonase, tissue plasminogen activator and prothrombin. Recently, minor genetic variations within these proteins have been discovered, which result in an increased tendency to form clots within blood vessels.

This study will investigate the risk of stroke associated with these genetically altered proteins. This can be achieved by comparing stroke patients and a group of non-stroke patients for these altered proteins.

Study Procedure

We anticipate that your involvement in this study will take up to 30 minutes. This will involve answering a short questionnaire followed by a blood test (15mls, approximately 3 teaspoons). Blood will be tested for cholesterol, glucose, and for the genetic variations within the glycoprotein, paroxonase, tissue plasminogen activator, prothrombin and fibrinogen genes (where possible, for stroke patients, blood will be taken with other routine blood tests to avoid additional discomfort). An electrical recording of the heart will be taken using an electrocardiograph machine. Four separate blood pressure readings will also be taken. Blood will be drawn using standard precautions, however despite this a very small risk of local skin infection exists. More commonly a small bruise associated with mild discomfort may occur at the injection site. This usually resolves within a few days. If you decide to participate, there will be no financial cost to you.

Confidentiality

All results will be kept in the strictest confidence and will only be available to participants on request. In the event of an abnormal finding (i.e. high blood glucose or cholesterol) your physician or general practitioner will be notified. For stroke patient involvement in this study will not affect their treatment. Please note, that if for any reason you decide that you no longer would like to participate in this study, you may withdraw at any time. Withdrawal will not affect any future treatment you receive at this hospital.

We thank you for your participation and time given to this study which we hope may help us reduce the risk of stroke in the near future. Please feel free to discuss any aspects of concern with the principle investigator, Dr Jim Jannes (phone 82226119). If you wish to speak to a person not involved with this study, you can contact contact Mr Paul Miller, Executive Officer, Ethics Committee, Queen Elizabeth Hospital (phone

Appendix III

Study Consent Form

Polygenic Disease – A Study of Genetic Risk in an Australian Stroke Population
The Adelaide Genetic Stroke Study

1. I, the undersigned
   .....(name) hereby consent to my involvement in the research project titled Polygenic Disease: A Study of Genetic risk In An Australian Stroke Population.

2. I have read the information sheet, and I understand the reasons for this study. The ways in which it will affect me have been explained by the research worker. My questions have been answered to my satisfaction. My consent is given voluntarily.

3. The details of the research project have been explained to me, including:
   - the expected time it will take,
   - the nature of any procedures being performed, and the number of times they will be performed,
   - any discomfort which I may experience.

4. I understand that the purpose of this research project is to improve the quality of medical care, but my involvement may not be of benefit to me.

5. I have been given the opportunity to have a member of my family or a friend present while the project was explained to me.

6. No information about my medical history will be taken from the hospital without the researcher being present. My identity will be kept confidential, and nothing will be published which could possibly reveal my identity.
7. My involvement in the project will not affect my relationship with my medical advisers. I understand that I am free to withdraw from the project at any stage without having to give any reasons, and that if I do withdraw from the project it will not affect my treatment at this hospital in the future.

Signed ..... Date.....
Full Name
(printed).....
Address
.....
Research Worker
.....

Appendix IV

Bivariate Analysis: Identification of Confounders

The table below lists the odds ratio of ischemic stroke following adjustment for each potential confounder. The effect of confounding was evaluated for both heterozygous (ht) vs wild-type (wt) and homozygous (hm) vs wt. Variables that altered the univariate odds ratio (i.e. no confounding influence) for each SNP by more than 10% were considered "important' variables and included in a multivariate analysis.

		Confounder ht vs wt hm vs wt				
SNP.	**Confounder**	**OR**	**OR**	**OR**	**%change in OR**	
PON1 54M/L	None	.	1.07	1.11		
	Smoker	3.24	1.19	1.11	11%	0%
	Systolic BP	0.97	1.01	1.09	-5%	-1%
	Diastolic BP	0.95	0.95	0.95	-12%	-14%
	Hypertension	0.57	1.09	1.14	2%	3%
	Cholesterol	0.56	1.07	1.11	0%	1%
	Hypercholesterol	0.48	1.04	1.10	-3%	0%
	Glucose	1.43	1.05	1.09	-2%	-2%
	Diabetes	2.89	1.16	1.21	8%	10%
	Atrial fib	8.41	1.22	1.22	14%	10%
	Family history	0.98	1.07	1.11	0%	0%
	Ethnic	7.10	1.16	1.15	8%	4%
Gp IIIa PlA1A2	None	.	1.38	0.79		
	Smoker	3.28	1.45	0.74	5%	-6%
	Systolic BP	0.97	1.27	0.65	-8%	-17%
	Diastolic BP	0.95	1.42	0.86	3%	10%
	Hypertension	0.56	1.39	0.70	0%	-11%
	Cholesterol	0.56	1.37	0.95	-1%	20%
	Hypercholesterol	0.49	1.34	0.86	-3%	9%
	Glucose	1.45	1.62	0.59	17%	-25%
	Diabetes	2.87	1.39	0.75	0%	-5%
	Atrial fib	8.22	1.33	0.76	-4%	-4%
	Family history	0.98	1.38	0.79	0%	0%
	Ethnic	8.31	1.41	0.63	2%	-20%

		Confounder	ht vs wt	hm vs wt		
SNP.	**Confounder**	**OR**	**OR**	**OR**	**%change in OR**	
PAI 5G/4G	None	.	0.84	0.83		
	Smoker	3.20	0.82	0.86	-2%	3%
	Systolic BP	0.97	0.91	0.81	9%	-3%
	Diastolic BP	0.95	0.94	0.85	12%	2%
	Hypertension	0.58	0.85	0.85	2%	3%
	Cholesterol	0.56	0.83	0.84	-1%	1%
	Hypercholesterol	0.48	0.80	0.82	-5%	-2%
	Glucose	1.43	0.77	0.86	-8%	3%
	Diabetes	2.85	0.86	0.90	2%	8%
	Atrial fib	8.30	0.97	0.91	16%	9%
	Family history	0.98	0.84	0.83	0%	0%
	Ethnic	6.53	0.89	0.85	6%	2%
Fib 148 C/T	None	.	0.95	1.48		
	Smoker	3.22	0.94	1.59	0%	7%
	Systolic BP	0.97	1.05	1.19	10%	-20%
	Diastolic BP	0.95	1.08	1.14	14%	-23%
	Hypertension	0.57	0.95	1.52	1%	3%
	Cholesterol	0.56	1.00	1.90	5%	28%
	Hypercholesterol	0.48	0.99	1.62	5%	10%
	Glucose	1.44	0.82	1.60	-13%	8%
	Diabetes	2.93	0.85	1.44	-11%	-3%
	Atrial fib	8.44	1.03	1.72	8%	16%
	Family history	0.97	0.95	1.49	0%	1%
	Ethnic	6.86	0.95	1.51	1%	2%
Gp 1b HPA 2a/b	None	.	0.83	.		
	Smoker	3.28	0.87	.	5%	
	Systolic BP	0.97	0.89	.	8%	
	Diastolic BP	0.95	0.88	.	6%	
	Hypertension	0.57	0.86	.	4%	
	Cholesterol	0.56	0.77	.	-7%	
	Hypercholesterol	0.48	0.78	.	-6%	
	Glucose	1.43	0.77	.	-7%	
	Diabetes	2.85	0.81	.	-3%	
	Atrial fib	8.43	0.74	.	-11%	
	Family history	0.98	0.83	.	0%	
	Ethnic	6.60	0.84		2%	

SNP.	Confounder	Confounder ht vs wt OR	ht vs wt OR	hm vs wt OR	%change in OR	
PON1 –107C/T	None	.	0.67	1.35		
	Smoker	3.40	0.66	1.47	-2%	9%
	Systolic BP	0.97	0.66	1.32	-1%	-2%
	Diastolic BP	0.95	0.69	1.26	3%	-7%
	Hypertension	0.57	0.67	1.35	0%	0%
	Cholesterol	0.57	0.72	1.36	7%	0%
	Hypercholesterol	0.48	0.65	1.30	-4%	-4%
	Glucose	1.43	0.68	1.41	2%	5%
	Diabetes	2.78	0.73	1.43	8%	5%
	Atrial fib	8.17	0.68	1.33	1%	-2%
	Family history	0.94	0.67	1.36	0%	0%
	Ethnic	6.36	0.69	1.40	2%	4%
Proth. 202010 G/A	None	.	1.26	.		
	Smoker	3.20	1.30	.	3%	
	Systolic BP	0.97	1.31	.	4%	
	Diastolic BP	0.95	1.35	.	7%	
	Hypertension	0.57	1.27	.	0%	
	Cholesterol	0.56	1.21	.	-4%	
	Hypercholesterol	0.48	1.28	.	1%	
	Glucose	1.43	0.88	.	-30%	
	Diabetes	2.85	1.12	.	-11%	
	Atrial fib	8.45	1.58	.	25%	
	Family history	0.97	1.27	.	1%	
	Ethnic	6.85	1.28		2%	
TPA –7,351C/T	None	.	1.13	1.52		
	Smoker	3.16	1.09	1.44	-4%	-5%
	Systolic BP	0.97	1.13	1.51	0%	0%
	Diastolic BP	0.95	1.15	1.59	2%	5%
	Hypertension	0.58	1.12	1.43	-1%	-6%
	Cholesterol	0.56	1.15	1.53	1%	0%
	Hypercholesterol	0.48	1.15	1.53	2%	1%
	Glucose	1.43	1.23	1.47	9%	-3%
	Diabetes	2.83	1.14	1.44	1%	-6%
	Atrial fib	8.90	1.24	1.90	9%	25%
	Family history	0.97	1.13	1.52	0%	0%
	Ethnic	7.02	1.13	1.55	0%	2%

Publication

The following manuscript has been accepted for publication in 'Stroke' and was in press at the time of submission of this thesis. Publication is anticipated in May 2004.

Full Title

Tissue Plasminogen Activator –7351C/T Enhancer Polymorphism is a Risk Factor for Lacunar Stroke

Authors

First author:

Jannes, Jim BMBS
Department of Medicine,
University of Adelaide
The Queen Elizabeth Hospital
Woodville South, South Australia 5011

Co-Authors (in order):

Hamilton-Bruce, Monica A. PhD
Department of Neurology
The Queen Elizabeth Hospital
Woodville South, South Australia 5011

Pilotto, Louis PhD
Department of General Practice
Flinders Medical Centre
Bedford Park, South Australia 5042

Smith, Brian J. PhD
Department of Respiratory Medicine
The Queen Elizabeth Hospital
Woodville South, South Australia 5011

Mullighan, Charles G. MD
Department of Haematology
The Royal Adelaide Hospital
Adelaide, South Australia 5000

Bardy, Peter G. MBBS
Research and Development / Transplant Services
Australian Red Cross – South Australian Branch
Adelaide, South Australia

Koblar, Simon A. PhD
Correspondence and reprints address:
Department of Medicine,
University of Adelaide
The Queen Elizabeth Hospital
28 Woodville Road
Woodville South
Adelaide, South Australia 5011

Acknowledgements

This study was funded by the Sylvia and Charles Viertel Foundation, The Queen Elizabeth Hospital (Department of Neurology) and Adelaide University. We thank the members of the Adelaide Stroke Genomics Collaboration (Dr M. Robinson – The Queen Elizabeth Hospital, Dr D.Thyagarajan - Flinders Medical Centre, Dr W. Jeffries - Lyell McEwin Health Service and Dr R. Casse – Memorial Hospital). We are also grateful for the voluntary efforts of Sr J. Bradford, Sr F. Buchan and Dr C. Rowe for his encouragement. The authors acknowledge the excellent technical assistance of Mr A. Condina and Mr S. Davis (molecular neurogenetics), Ms B. Kenny (South Australian Department of Human Services, control recruitment), Ms R. Attewell (Covance Pty Ltd, biostatistics), The Australian Red Cross (South Australia) tissue-typing laboratory (DNA preparation) and thank Professor Geoffrey Donnan for reviewing the manuscript.

Full and Cover Title

Tissue Plasminogen Activator –7351C/T Enhancer Polymorphism is a Risk Factor for Lacunar Stroke

Tables

Keywords

genetics, polymorphism, lacunar infarction, tissue plasminogen activator

Word Count – 3935

Abstract

Background and Purpose

Occlusive thrombosis is an important component of both small and large vessel ischemic stroke. Endogenous tissue plasminogen activator (TPA) is the primary mediator of intravascular fibrinolysis and is predominantly expressed by the endothelium of small vessels. The acute release of TPA is influenced by the TPA –7351C/T polymorphism and therefore may play an important role in the pathogenesis of lacunar stroke. In this study we investigated the risk of lacunar and non-lacunar ischemic stroke associated with the TPA –7351C/T polymorphism.

Methods

We conducted a case-control study of 182 cases of ischemic stroke and 301 community controls. Participants were evaluated for known cerebrovascular risk factors and the TPA –7351C/T genotype was established by a polymerase chain reaction (PCR) method. Logistic regression was used to determine the risk of lacunar and non-lacunar ischemic stroke associated with the TPA –7351C/T polymorphism.

Results

The prevalence of the TPA –7351 CC, CT, TT genotype was 46%, 45%, 9% for controls and 41%, 46% and 13% for stroke patients, respectively. After adjustment for known cerebrovascular risk factors, the TT genotype was significantly associated with ischemic stroke (OR 1.9, 95%CI 1.01-3.6). Stratification for stroke subtype showed a significant association between the TT genotype and lacunar stroke, but not non-lacunar stroke (OR 2.7, 95%CI 1.1-6.7).

Conclusions

The TPA –7351C/T polymorphism is an independent risk factor for lacunar stroke. The findings suggest that impaired fibrinolysis may play a role in the pathogenesis of lacunar stroke.

Occlusive thrombosis is an important component of both small and large vessel ischemic stroke. It is usually caused by focal disruption of an atheromatous plaque, exposing sub-endothelial tissues that trigger a local haemostatic response. Thrombosis may also occur under conditions of high shear stress, such as those induced by atherosclerotic stenosis.[1] Irrespective of the cause, the endogenous fibrinolytic defense system plays an important role in counteracting thrombus formation, with the degree of thrombosis governed by the balance of pro- and anti-thrombotic forces.[2]

Endothelium-derived tissue plasminogen activator is the primary mediator of local intravascular fibrinolysis. Knockout mice deficient for TPA have markedly slower rates of reperfusion following thrombotic occlusion of a carotid artery when compared with their wild-type counterparts.[3] In humans, plasma TPA levels are regulated by two mechanisms: (a) an acute release of local TPA that is precipitated by vascular injury and (b) a long-term change in the rate of synthesis of TPA that reflects the chronic activation of the fibrinolytic system.[4] The latter is a marker of atherosclerotic burden, with a high plasma TPA level being an independent predictor of ischemic stroke in both cross-sectional[5-11] and prospective cohort studies.[12, 13] The rapid release of local TPA by the endothelium is considered to be the most important mechanism mediating the dissolution of arterial thrombus.[4] Hence a low capacity for rapid TPA release is likely to predispose to ischemic stroke. Furthermore, as small vessels predominantly express TPA within the brain,[14, 15] factors that influence its release may have particular relevance to lacunar stroke.

In humans, genetic factors play an important role in the variance of endothelial TPA release.[16] Recently, a single nucleotide polymorphism located at position –7351 within the enhancer region of the TPA gene was identified and shown to be strongly correlated with endothelial TPA release rates.[17] The polymorphism was also found to have functional importance as it occurred within an Sp1 binding site, a factor promoting DNA transcription. Possession of a thymidine (T) allele was shown to inhibit Sp1 binding and was associated with less than half the TPA release observed in those homozygous for the cytosine (C) allele.[17] The TPA –7351C/T polymorphism has

subsequently been shown to be clinically relevant, having a strong association with first myocardial infarction.[18]

The risk of ischemic stroke associated with the TPA –7351C/T polymorphism has not been reported. We therefore performed a case-control study to determine the risk of ischemic stroke in relation to either lacunar or non-lacunar events associated with this genetic determinant of intravascular fibrinolysis, with adjustment for known risk factors.

Subjects and Methods

Patients admitted with acute ischemic stroke to one of five major hospitals within metropolitan Adelaide, South Australia, were invited to participate in the study. The diagnosis of ischemic stroke was made by a neurologist (JJ) in accordance with the World Health Organization definition.[19] All patients underwent brain computerized tomography or magnetic resonance imaging, allowing exclusion of intracerebral and subarachnoid haemorrhage. Ischemic stroke was sub-typed into four categories using the Oxfordshire Community Stroke Project (OCSP) classification system:[20] (a) Total anterior circulation syndrome (TACS); (b) Partial anterior circulation syndrome (PACS); (c) Posterior circulation syndrome (PCS) and (d) Lacunar syndrome (LS). The control group consisted of non-hospitalized subjects who resided predominantly in metropolitan Adelaide and did not have a personal history of cerebrovascular disease. Controls were selected by means of random sampling of the South Australian electronic telephone directory and group matched with batches of patients for age (within five-year strata) and gender.

Following informed, written consent, patients and controls were interviewed and demographic information was recorded. If the patient was unable to communicate verbally, information was recorded from the next of kin. Cerebrovascular risk factors including hypertension, hypercholesterolemia and diabetes were recorded if there was a self-reported history or the individual was receiving medical treatment for the condition(s) at the time of investigation. An electrocardiogram was undertaken at the time of the interview to determine the presence of atrial fibrillation. Patients were

coded with AF if there was objective evidence on ECG or there was a reported history of paroxysmal AF. Subjects were classified as smokers if they smoked cigarettes or tobacco on a regular basis at the time of the interview or within the last five years. A history of stroke in a first degree relative was also recorded. A venous blood sample was taken from all participants for genetic analysis. The study was approved by the clinical research ethics committees of participating hospitals.

Genotype Determination

Genomic DNA was isolated from 6ml of whole blood by a standard method.[21] A 405 base pair DNA fragment involving the –7351C/T polymorphism of the TPA gene was amplified using the sequence-specific primer PCR method.[22] The sequence of the two reverse allele-specific primers was 5'-ATGGCTGTGTCTGGGGCG-3', 5'-ATGGCTGTGTCTGGGGCA-3', and that of the forward consensus primer 5'-ATTGGCGCAAACTCCTCA-3' (Genbank Accession Number: Z48484). The PCR products were separated on a 2% agarose gel and the allelic frequency was determined. DNA was de-identified and analyzed in coded form, blinding laboratory staff and data abstractors to case/control status. Genotyping was also performed by two independent investigators. They were in full agreement.

Statistical Analysis

Statistical analysis was performed using STATA statistical software (Version 7.0, College Station, Texas, USA). Based on a population TPA –7351 TT genotype frequency of 10%,[18] we estimated that a sample size of 150 patients and 150 controls would be sufficient to determine a minimum odds ratio of 2.5 (80% power, 95% confidence). The strength of association between TPA –7351C/T genotype and ischemic stroke was estimated by calculating the odds ratio (OR) and 95% confidence intervals (CI). Odds ratios were also calculated to determine the relative strength of association of known risk factors. Unconditional logistic regression analysis was used to examine the modifying effect of known risk factors (age, gender, ethnicity, family history of stroke, hypertension, diabetes, smoking, hypercholesterolemia, and atrial fibrillation) on the risk of stroke associated with the TPA –7351C/T genotype. Although cases and controls were matched for age and gender, the matching was not

on an individual basis and these variables were included as potential confounders. Variables that altered the univariate odds ratio by 10% or greater were then incorporated into an unconditional logistic regression model, to determine of the risk of ischemic stroke associated with the TPA –7351C/T polymorphism. This model also examined for the effect of interaction between the TPA –7351C/T genotype and each of the known risk factors. A univariate analysis stratified for ischemic stroke subtype was performed. A two tailed p-value of <0.05 was considered significant.

Results

Two hundred and one consecutive patients with acute ischemic stroke were invited to participate in the study. Of these, 182 (90.5%) agreed to participate. One hundred and thirty-seven patients (75%) presented with first ischemic stroke. The demographic characteristics and prevalence of cerebrovascular risk factors for cases and 301 controls are shown in Table 1. No significant differences were observed between the two groups in terms of demographic variables. Of the known risk factors examined, atrial fibrillation was associated with the highest risk of ischemic stroke, with 23% of cases versus 3% of controls affected (OR 8.5, 95% CI 4.1 - 17.4). A history of smoking within the last five years (OR 3.1, 95% CI 1.9-5.2) and diabetes (OR 2.7, 1.6 - 4.4) were also found to be associated significantly with ischemic stroke. No association was observed between ischemic stroke and a history of stroke in a first degree relative, hypertension or hypercholesterolemia (Table 1), or for the use of antihypertensive, antiplatelet and lipid lowering medication (results not shown).

Two patients died after enrolment in the study and prior to venous blood sampling, thus genetic analysis could not be performed in these cases. Table 2 shows the prevalence of the TPA –7351C/T polymorphism in the remaining 180 ischemic stroke cases and 301 controls. The allelic frequencies were in Hardy-Weinberg equilibrium. Among the control group, 46% were homozygous for the TPA –7351 C allele (CC), 45% were heterozygous (CT) and 9% were homozygous for the T allele (TT). The genotype distribution in the ischemic stroke cohort was 41% (CC), 46% (CT) and 13% (TT), respectively.). In a univariate analysis, the risk of ischemic stroke was not significantly associated with the TT genotype (OR 1.5, 95%CI 0.8-2.8). Atrial

fibrillation was the only variable that significantly altered the univariate risk of ischemic stroke associated with the TT genotype and following adjustment for this, a significant association between this genotype and ischemic stroke was observed. The distribution of the TPA –7351C/T genotype among patients stratified for stroke subtype and controls is shown in Table 3. Forty-four ischemic stroke patients (24.5%) were classified with lacunar stroke. In this sub-group, the TPA –7351 TT genotype was significantly associated with lacunar stroke (OR 2.6, 95%CI 1.1-6.4, p=0.039). The corresponding attributable risk associated with the TPA –7351 TT genotype was 13% (2-17%). Removal of the non-caucasoid participants from the analysis did not alter the significance of this finding. In contrast, the TPA –7351C/T polymorphism was not associated with an increased risk of non-lacunar stroke. No interaction between the TPA TT genotype and smoking, hypertension, hypercholesterolemia, diabetes, family history for stroke, atrial fibrillation or a combination of the above was found (results not shown).

Discussion

This is the first study to identify an association between the TPA –7351C/T polymorphism and ischemic stroke, after adjustment for most known cerebrovascular risk factors. In a subgroup analysis, the TPA –7351TT genotype was significantly associated with lacunar, but not non-lacunar stroke. The prevalence of the TPA –7351TT genotype amongst cases was low and consequently the risk of lacunar stroke attributable to the TPA –7351TT genotype was 13%. This implicates other factors in the pathogenesis of lacunar stroke and is in keeping with the polygenic model of ischemic stroke, in which the predisposition arises following the accumulation and interaction of multiple genetic loci, each of which possess a weak phenotypic effect.[23]

Two distinct small vessel pathologies underlie most lacunar stroke: lipohyalinosis and micro-atheroma.[24] Both lead to slow luminal narrowing causing hemodynamic compromise that may initially present with the Capsular Warning Syndrome.[25] Subsequent lacunar infarction occurs following non-thrombotic obliteration of diseased small vessels or from occlusive thrombus.[24] The finding of a nearly 3-fold increase in risk of lacunar stroke associated with homozygosity for the T allele of the

TPA –7351C/T polymorphism suggests that fibrinolytic factors may play an important role in the preservation of small vessel patency. This is supported by studies of normal primate and rat brains, that show TPA expression is restricted to the endothelial cells of pre-capillary arterioles and post-capillary venules, with no expression by the endothelium of large arteries.[14, 15] This differential expression of TPA suggests that under normal circumstances, vascular patency of small vessels relies on an intact fibrinolytic system and that these vessels may be vulnerable to thrombotic occlusion by factors that impair fibrinolytic potential. Functional studies have also shown that impaired fibrinolytic activity may be particularly important in the pathogenesis of lacunar stroke. Fibrinolytic activity, as measured by plasma fibrinogen degradation products, varies considerably between lacunar, cardioembolic and atherothrombotic ischemic stroke, with the lowest levels being observed in lacunar stroke.[26] In another study, impairment of fibrinolytic potential, as determined by the euglobulin lysis time, was observed more often in patients with lacunar stroke than those with cortical infarction due to large vessel occlusion.[27]

Our findings are consistent with those of a recent study showing an increased risk of ischemic stroke associated with the D allele of the TPA Alu-repeat insertion/deletion polymorphism.[28] This polymorphism is in tight linkage disequilibrium with the TPA –7351C/T polymorphism with a linkage disequilibrium coefficient of 0.71.[17] Functional significance, however, is greatest for the TPA –7351C/T polymorphism which accounts for 18% of the phenotypic variance in TPA release rates, a finding nearly double that associated with the TPA Alu-repeat insertion/deletion polymorphism (10%).[29] It is thus plausible that the positive association between the TPA Alu-repeat insertion/deletion polymorphism and ischemic stroke occurs through linkage disequilibrium with the functional TPA –7351C/T polymorphism.

The potential limitations of our study warrant consideration. The sample size was small, resulting in wide confidence intervals that approached unity. Therefore, the possibility of a type 1 error cannot be excluded and a study of greater power is required to confirm the strength of association. Various sources of bias may have also arisen. Bias due to incomplete ascertainment of cases may have occurred as the study

sample consisted entirely of hospitalised stroke survivors who were able to give informed consent, and did not include patients with severe stroke or those who did not present to hospital. Although this study controlled for many of the known cerebrovascular risk factors, other potential risk factors such as alcohol intake and markers of inflammation were not recorded and remain a potential source of confounding bias. Furthermore, misclassification of diabetes, hypertension and hypercholesterolemia may have occurred as diagnosis was based on self-reported history alone. The risk factor profile (excluding hypertension) for stroke patients was, however, consistent with that previously reported.[30] The limitations associated with the OCSP classification system could have also led to misclassification of stroke subtype. For example, the OCSP cannot differentiate between lacunar infarction (small vessel disease) and striato-capsular infarction without cortical signs (large vessel disease). However, the OCSP classification system has been validated, with concordance between clinical subtype and brain neuro-imaging reported in 75% of cases.[31] In the present study, clinical subtype was made in conjunction with neuro-imaging, thus further limiting this error. This study did not measure for the effects of population stratification, which may have led to a spurious association. The distribution of the TPA –7351C/T genotype amongst controls, however, was similar to that reported in another Caucasoid population from Sweden (CC-50.8%, CT-39.2%, TT-10.0%).[18]

In conclusion, the results of this study suggest that impaired fibrinolysis plays a role in pathogenesis of lacunar stroke and supports the hypothesis that the TPA –7351C/T polymorphism represents an inherited risk factor in a Caucasoid population. Confirmation by a larger study is required, which may then provide a better means to predict the risk of lacunar stroke.

Table 1: Demographic Characteristics and Prevalence of Risk Factors for Patients with Ischemic Stroke and Controls

	Controls n (%)	Cases n (%)	Odds Ratio (95% CI)	p value
Age* (years)	73.4 ± 11.6	73.6 ± 12		0.8
Gender				
Females	134 (45)	80 (44)		
Males	167 (55)	102 (56)	1.0 (0.7 - 1.5)	0.9
Ethnic Origin				
Caucasoid	299 (99.7)	178 (98)		
Non-Caucasoid	1 (0.3)	4 (2)	6.7 (0.8 - 60.6)	0.09
Family History				
No	201 (67)	122 (67)		
Yes	100 (33)	60 (33)	1.0 (0.7 – 1.5)	0.95
Hypertension				
No	164 (54)	87 (48)		
Yes	137 (46)	85 (52)	1.3 (0.9 – 1.9)	0.2
Diabetes Mellitus				
No	268 (89)	137 (75)		
Yes	33 (11)	45 (25)	2.7 (1.6 – 4.4)	<0.0001
Smoking				
No	271 (90)	135 (74)		
Yes	30 (10)	47 (26)	3.1 (1.9 – 5.2)	<0.0001
Hypercholesterolemia				
No	182 (60)	117 (64)		
Yes	119 (40)	65 (36)	0.9 (0.6 – 1.2)	0.4
Atrial Fibrillation				
No	291 (97)	141 (77)		
Yes	10 (3)	41 (23)	8.5 (4.1 – 17.4)	<0.0001
Total	301	182		

**Age expressed as a mean ± 1 standard deviation*

Table 2.

Prevalence, odds ratios (OR), 95% confidence intervals (CI) and p values for the TPA –7351 C/T polymorphism among patients with ischemic stroke and controls

			Univariate		Multivariate*	
Genotype	**Control n(%)**	**Cases n(%)**	**OR(95% CI)**	**p**	**OR(95% CI)**	**p**
CC	137 (46)	74 (41)	1.0		1.0	
CT	136 (45)	83 (46)	1.1 (0.8-1.7)	0.5	1.2 (0.8-1.9)	0.3
TT	28 (9)	23 (13)	1.5 (0.8-2.8)	0.2	1.9 (1.01-3.6)	0.049
Total	301 (100)	180 (100)				

**Risk factors considered include age, gender, ethnicity, family history, hypertension, diabetes, smoking, hypercholesterolemia and atrial fibrillation*

Table 3.

Prevalence, odds ratios (OR), 95% confidence intervals (CI) and p values for the TPA – 7351 C/T polymorphism among patients with ischemic stroke stratified for stroke subtype and controls

	Control	**Lacunar Stroke**			**Non-Lacunar Stroke***		
Genotype	**n (%)**	**n (%)**	**OR (95% CI)**	**p**	**n (%)**	**OR (95% CI)**	**p**
CC	137 (46)	17 (39)	1.0		57 (42)	1.0	
CT	136 (45)	18 (41)	1.1 (0.5 – 2.1)	0.9	65 (48)	1.2 (0.8 – 1.8)	0.5
TT	28 (9)	9 (20)	2.6 (1.1 – 6.4)	0.039	14 (10)	1.2 (0.6 – 2.5)	0.6
Total	301 (100)	44 (100)			136(100)		

**Stroke patients presenting with either TACS, PACS and PCS*

References

1. Konstantopoulos K, Grotta JC, Sills C, Wu KK, Hellums JD. Shear-induced platelet aggregation in normal subjects and stroke patients. *Thrombosis & Haemostasis.* 1995; 74:1329-34.
2. Rosenberg RD, Aird WC. Vascular-bed--specific hemostasis and hypercoagulable states. *New England Journal of Medicine.* 1999; 340:1555-64.
3. Matsuno H, Kozawa O, Niwa M, Ueshima S, Matsuo O, Collen D, Uematsu T. Differential role of components of the fibrinolytic system in the formation and removal of thrombus induced by endothelial injury. *Thrombosis & Haemostasis.* 1999; 81:601-4.
4. Kooistra T, Schrauwen Y, Arts J, Emeis JJ. Regulation of endothelial cell t-PA synthesis and release. *International Journal of Haematology.* 1994; 59:233-55.
5. Johansson L, Jansson JH, Boman K, Nilsson TK, Stegmayr B, Hallmans G. Tissue plasminogen activator, plasminogen activator inhibitor-1, and tissue plasminogen activator/plasminogen activator inhibitor-1 complex as risk factors for the development of a first stroke. *Stroke.* 2000; 31:26-32.
6. Kain K, Catto AJ, Carter AM, Young J, Bamford J, Bavington J, Grant PJ. Decreased fibrinolytic potential in South Asian women with ischaemic cerebrovascular disease. *British Journal of Haematology.* 2001; 114:155-61.
7. Margaglione M, Di Minno G, Grandone E, Vecchione G, Celentano E, Cappucci G, Grilli M, Simone P, Panico S, Mancini M. Abnormally high circulation levels of tissue plasminogen activator and plasminogen activator inhibitor-1 in patients with a history of ischemic stroke. *Arteriosclerosis & Thrombosis.* 1994; 14:1741-5.
8. Lindgren A, Lindoff C, Norrving B, Astedt B, Johansson BB. Tissue plasminogen activator and plasminogen activator inhibitor-1 in stroke patients. *Stroke.* 1996; 27:1066-71.
9. Kristensen B, Malm J, Nilsson TK, Hultdin J, Carlberg B, Dahlen G, Olsson T. Hyperhomocysteinemia and hypofibrinolysis in young adults with ischemic stroke. *Stroke.* 1999; 30:974-80.

10. Kristensen B, Malm J, Nilsson TK, Hultdin J, Carlberg B, Olsson T. Increased fibrinogen levels and acquired hypofibrinolysis in young adults with ischemic stroke. *Stroke.* 1998; 29:2261-7.

11. Macko RF, Kittner SJ, Epstein A, Cox DK, Wozniak MA, Wityk RJ, Stern BJ, Sloan MA, Sherwin R, Price TR, McCarter RJ, Johnson CJ, Earley CJ, Buchholz DW, Stolley PD. Elevated tissue plasminogen activator antigen and stroke risk: The Stroke Prevention In Young Women Study. *Stroke.* 1999; 30:7-11.

12. Ridker PM. Plasma Concentration of Endogenous Tissue Plasminogen Activator and the Occurrence of Future Cardiovascular Events. *Journal of Thrombosis & Thrombolysis.* 1994; 1(1): 35-40

13. Smith FB, Rumley A, Lee AJ, Leng GC, Fowkes FG, Lowe GD. Haemostatic factors and prediction of ischaemic heart disease and stroke in claudicants. *British Journal of Haematology.* 1998; 100:758-63.

14. Levin EG, del Zoppo GJ. Localization of tissue plasminogen activator in the endothelium of a limited number of vessels. *American Journal of Pathology.* 1994; 144:855-61.

15. Schreiber SS, Tan Z, Sun N, Wang L, Zlokovic BV. Immunohistochemical localization of tissue plasminogen activator in vascular endothelium of stroke-prone regions of the rat brain. *Neurosurgery.* 1998; 43:909-13.

16. Jern C, Ladenvall P, Wall U, Jern S. Gene polymorphism of t-PA is associated with forearm vascular release rate of t-PA. *Arteriosclerosis, Thrombosis & Vascular Biology.* 1999; 19:454-9.

17. Ladenvall P, Wall U, Jern S, Jern C. Identification of eight novel single-nucleotide polymorphisms at human tissue-type plasminogen activator (t-PA) locus: association with vascular t-PA release in vivo. *Thrombosis & Haemostasis.* 2000; 84:150-5.

18. Ladenvall P, Johansson L, Jansson JH, Jern S, Nilsson TK, Tjarnlund A, Jern C, Boman K. Tissue-type plasminogen activator -7,351C/T enhancer polymorphism is associated with a first myocardial infarction. *Thrombosis & Haemostasis.* 2002; 87:105-9.

19. Hatono S. Experience from a multicentre stroke register: a preliminary report. *Bull WHO* 1976; 54:541-553.

20. Bamford J, Sandercock P, Dennis M, Burn J, Warlow C. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet.* 1991; 337:1521-6.

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

22. Bunce M, O'Neill CM, Barnardo MC, Krausa P, Browning MJ, Morris PJ, Welsh KI. Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP). *Tissue Antigens.* 1995; 46:355-67.

23. Rubattu S, Volpe M, Kreutz R, Ganten U, Ganten D, Lindpaintner K. Chromosomal mapping of quantitative trait loci contributing to stroke in a rat model of complex human disease.[comment]. *Nature Genetics.* 1996; 13:429-34.

24. Fisher C. Capsular Infarcts-the underlying vascular lesions. *Arch Neurol* 1979; 36:65-73.

25. Donnan GA, O'Malley HM, Quang L, Hurley S, Bladin PF. The capsular warning syndrome: pathogenesis and clinical features.[comment]. *Neurology.* 1993; 43:957-62.

26. Kataoka S, Hirose G, Hori A, Shirakawa T, Saigan T. Activation of thrombosis and fibrinolysis following brain infarction. *Journal of the Neurological Sciences.* 2000; 181:82-8.

27. Kilpatrick TJ, Matkovic Z, Davis SM, McGrath CM, Dauer RJ. Hematologic abnormalities occur in both cortical and lacunar infarction. *Stroke.* 1993; 24:1945-50.

28. Austin H, Chimowitz MI, Hill HA, Chaturvedi S, Wechsler LR, Wityk RJ, Walz E, Wilterdink JL, Coull B, Sila CA, Mitsias P, Evatt B, Hooper WC, Genetics, Stroke in the Young Study G. Cryptogenic stroke in relation to genetic variation in clotting factors and other genetic polymorphisms among young men and women. *Stroke.* 2002; 33:2762-8.

29. Ladenvall P, Wall U, Jern S, Jern C. Eight single-nucleotide polymorphisms (SNPs) at the human tissue-type plasminogen activator (t-PA) locus.[comment]. *Journal of Human Genetics.* 2001; 46:737-8.

30. Davis PH, Dambrosia JM, Schoenberg BS, Schoenberg DG, Pritchard DA, Lilienfeld AM, Whisnant JP. Risk factors for ischemic stroke: a prospective study in Rochester, Minnesota. *Annals of Neurology.* 1987; 22:319-27.

31. Mead GE, Lewis SC, Wardlaw JM, Dennis MS, Warlow CP. How well does the Oxfordshire community stroke project classification predict the site and size of the infarct on brain imaging?[comment]. *Journal of Neurology, Neurosurgery & Psychiatry.* 2000; 68:558-62.