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

Thesis Abstract .................................................................................................................. 5  
Acknowledgements ........................................................................................................... 7  
Declaration ....................................................................................................................... 8  
Conference Presentations ................................................................................................. 9  
List of Figures ................................................................................................................... 10  
List of Tables ................................................................................................................... 11  
Abbreviations .................................................................................................................. 14

## Chapter 1: Introduction

1.1 The Global Burden of Stroke ....................................................................................... 16  
1.2 Stroke: an Australian Perspective ............................................................................. 17  
1.3 Classification of Stroke ............................................................................................ 19  
1.3.1 Stroke Subtypes .................................................................................................... 19  
1.3.2 Pathophysiological Classification of Ischemic Stroke ......................................... 20  
1.4 Thesis Overview ....................................................................................................... 23

## Chapter 2: Genetics and Ischemic Stroke

2.1 The Attributable Risk of Known Risk Factors ............................................................ 25  
2.2 Evidence of a Genetic Predisposition to Stroke ....................................................... 27  
2.2.1 Twin Studies ....................................................................................................... 27  
2.2.2 Family-based Studies .......................................................................................... 28  
2.2.3 Animal Studies .................................................................................................... 30  
2.3 The Genetic Heterogeneity of Ischemic Stroke ......................................................... 31  
2.4 Candidate Genes for Ischemic Stroke ..................................................................... 33  
2.4.1 Paraoxonase (PON1) .......................................................................................... 34  
2.4.2 Glycoprotein Ib .................................................................................................. 36  
2.4.3 Glycoprotein IIb/IIIa ......................................................................................... 38  
2.4.4 Fibrinogen ......................................................................................................... 41  
2.4.5 Prothrombin ...................................................................................................... 44  
2.4.6 Tissue Plasminogen Activator (TPA) .................................................................. 45  
2.4.7 Plasminogen Activator Inhibitor – 1 (PAI-1) ...................................................... 47  
2.4.8 Conclusions ....................................................................................................... 49
Chapter 3: Thesis Aims and Rationale

3.1 Thesis Aims ................................................................. 50
3.2 Selection of Study Design ........................................... 51
3.3 Selection of Candidate SNP’s ....................................... 54
3.4 Sample Size Estimation .............................................. 55

Chapter 4: Research Methods

4.1 Clinical Methods .......................................................... 59
4.1.1 Participating Hospitals ............................................. 59
4.1.2 Recruitment of Ischemic Stroke Cases ......................... 60
4.1.3 Questionnaire Assessment of Ischemic Stroke Cases ...... 61
4.1.4 Clinical Assessment of Ischemic Stroke Cases ............ 62
4.1.5 Case Procedures ...................................................... 64
4.1.6 Recruitment of Community Controls .......................... 64
4.1.7 Questionnaire Assessment of Community Controls ...... 65
4.1.8 Clinical Evaluation of Community Controls ................ 66
4.1.9 Control Procedures ................................................. 66
4.2 Laboratory Methods .................................................... 66
4.2.1 Glucose and Cholesterol Determination ....................... 66
4.2.2 DNA Extraction ....................................................... 66
4.2.3 Genotyping Methods ................................................ 66
4.2.4 Oligonucleotide Primer Design ................................ 67
4.2.5 PCR-SSP Method ..................................................... 70
4.2.6 Genotype Determination .......................................... 72
4.2.7 Quality Assurance Measures .................................... 73
4.3 Statistical Methods ..................................................... 74
4.3.1 Descriptive Statistics .............................................. 74
4.3.2 Univariate Analysis .................................................. 75
4.3.3 Multivariate Analysis ............................................. 76
4.3.4 Interactions ............................................................ 77

Chapter 5: Results

5.1 Study Population .......................................................... 78
5.2 Univariate Analysis: Stroke Risk Factors ....................... 80
5.3 Univariate Analysis: Medication .................................... 83
5.4 Univariate Analysis: SNP’s .......................................... 84
5.5 Bivariate Analysis: Identification of Confounders ............ 86
5.6 Multivariate Analysis: SNP’s ....................................... 87
5.7 Subgroup Analysis: Lacunar Stroke ................................ 92
5.8 Subgroup Analysis: Cardioembolic Stroke ..................... 96
5.9 Gene-Risk Factor Interactions ...................................... 99
5.10 Gene-Gene Interactions ............................................. 101
5.11 Results Summary ..................................................... 102
Chapter 6: Discussion

6.1 Introduction ........................................................................................................ 104
6.2 TPA -7351 C/T SNP and Ischemic Stroke ......................................................... 104
6.3 PAI-1 5G/4G SNP and Ischemic Stroke .............................................................. 109
6.4 PON1 SNP’s and Ischemic Stroke ..................................................................... 110
6.5 β Fibrinogen -148C/T SNP and Ischemic Stroke ............................................... 113
6.6 Prothrombin 20210G/A SNP and Ischemic Stroke ........................................... 115
6.7 Platelet Glycoprotein SNP’s and Ischemic Stroke ............................................. 117
6.8 Study Limitations ............................................................................................. 119

Chapter 7: Future Directions

7.1 Selection of Appropriate Study Design ............................................................ 125
7.2 The Importance of Stroke Sub-typing ............................................................... 127
7.3 Adjustment for Population Stratification .......................................................... 130
7.4 Alternate Disease Phenotypes .......................................................................... 131
7.5 Final Considerations ......................................................................................... 132

Bibliography ............................................................................................................ 135
Appendix I: The OCSP Classification Criteria ....................................................... 154
Appendix II: Study Information Sheet ................................................................. 155
Appendix III: Study Consent Form ....................................................................... 157
Appendix IV: Bivariate Analysis: Identification of Confounders ............................ 158
Publication .............................................................................................................. 161
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:...................
Conference Presentations

Poster Presentation:

Poster Presentation:

Platform Presentation:
Recipient of the “New Investigator Award for Scientific Research”

Platform Presentation:
Recipient of the “Best Higher Degree Research Award” and the “Allan Kerr Grant Most Outstanding Postgraduate Learner Award”.

Platform Presentation:
List of Figures

Figure 2.1
Attributable Risks of a Combination of Risk Factors for First Ischemic Stroke. The Rochester Minnesota Study: ................................................................. 26

Figure 2.2
Threshold Model for the Inheritance of Common Ischemic Stroke ............................................. 32

Figure 2.3
Genetic Candidates for Common Ischemic Stroke .......................................................................... 33

Figure 2.4
Structure of the Glycoprotein Ib/IX/V Complex ............................................................................. 37

Figure 2.5
Structure of the Glycoprotein IIb/IIIa Complex .............................................................................. 39

Figure 2.6
Structure of the β fibrinogen gene .................................................................................................... 42

Figure 2.7
The Human Fibrinolytic Pathway .................................................................................................... 46

Figure 2.8
Structure of the TPA Gene ............................................................................................................. 47

Figure 4.1
Genotype Determination by Visual Inspection: Glycoprotein Ib T/M, Glycoprotein IIaPIA1/A2, PAI 5G/4G and Fibrinogen 148 C/T Genotype In One Subject. .................. 72

Figure 4.2
Chromatogram Showing DNA Sequence of TPA –7351 C/T SNP .............................................. 74

Figure 7.1
Comparison of Linkage and Association Analysis for Detecting Genetic Effects ................. 126
List of Tables

Table 2.1
Monogenic Disorders Causing Ischemic Stroke...................................................... 24

Table 2.2
Odds Ratios, Confidence Intervals and Probability Values for Risk Factors for First
Ischemic Stroke: The Rochester, Minnesota Study.................................................. 26

Table 2.3
Odds Ratios and Confidence Intervals for Proband Stroke by Familial History:
The Family Heart Study............................................................................................... 30

Table 3.1
Sample size estimations for selected single nucleotide SNP’s based on genotype
frequency data in a Caucasian population .............................................................. 56

Table 3.2
Sample size estimations for selected single nucleotide SNP’s based on previous
positive case-control studies of ischemic stroke..................................................... 57

Table 4.1
Oligonucleotide Primer Sequences, Genebank Accession Numbers and PCR Product
Sizes ............................................................................................................................ 68

Table 4.2
Primer Reaction Mix for each SNP ........................................................................... 70

Table 5.1
Gender and Age Characteristics of Consenters and Non-Consenters........................ 79

Table 5.2
Demographic Characteristics Between Controls and Cases...................................... 80

Table 5.3
Cerebrovascular Risk Factor Characteristics Between Controls and Cases.............. 81

Table 5.4
Hypertension Between Controls and Cases. ............................................................. 82

Table 5.5
Hypercholesterolemia Between Controls and Cases.................................................. 83
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.6</td>
<td>Medication Use Between Controls and Cases</td>
<td>84</td>
</tr>
<tr>
<td>5.7</td>
<td>Univariate Analysis of SNP’s (by genotype) between Controls and Cases</td>
<td>85</td>
</tr>
<tr>
<td>5.8</td>
<td>Univariate Analysis of SNP’s (by allele) between Controls and Cases</td>
<td>86</td>
</tr>
<tr>
<td>5.9</td>
<td>Multivariate Analysis of SNP’s (by genotype) between Controls and Cases Adjusted for ‘Important’ Confounders</td>
<td>88</td>
</tr>
<tr>
<td>5.10</td>
<td>Multivariate Analysis of SNP’s (by allele) between Controls and Cases Adjusted for ‘Important’ Confounders</td>
<td>89</td>
</tr>
<tr>
<td>5.11</td>
<td>Multivariate Analysis of SNP’s (by genotype) Between Controls and Cases Adjusted for all Potential Confounders</td>
<td>90</td>
</tr>
<tr>
<td>5.12</td>
<td>Multivariate Analysis of SNP’s (by allele) between Controls and Cases Adjusted for all Potential Confounders</td>
<td>91</td>
</tr>
<tr>
<td>5.13</td>
<td>Lacunar Stroke: Univariate Analysis of SNP’s Between Controls and Cases</td>
<td>93</td>
</tr>
<tr>
<td>5.14</td>
<td>Attributable Risk of Lacunar Stroke Associated With TPA –7351TT and PON1–107CC Genotypes</td>
<td>94</td>
</tr>
<tr>
<td>5.15</td>
<td>Non-Lacunar Stroke: Univariate Analysis of SNP’s Between Controls and Cases</td>
<td>95</td>
</tr>
<tr>
<td>5.16</td>
<td>Cardioembolic (CE) Stroke: Univariate Analysis of SNP’s Between Controls and Cases</td>
<td>97</td>
</tr>
<tr>
<td>5.17</td>
<td>Non-Cardioembolic Stroke: Univariate Analysis of SNP’s Between Controls and Cases</td>
<td>98</td>
</tr>
<tr>
<td>5.18</td>
<td>Probability Values for Gene-Risk Factor Interactions</td>
<td>100</td>
</tr>
<tr>
<td>5.19</td>
<td>Probability Values for Gene-Gene Interactions</td>
<td>102</td>
</tr>
</tbody>
</table>
Table 6.1
Sample Size Estimation Based on Adjusted OR’s Determined for Each SNP 120

Table 6.2
Variation of PON1 -107 T/C and TPA-7351 C/T Genotype Distributions Amongst Caucasian Populations 123

Table 7.1
Relationship Between Age of Stroke and Positive Family History of Stroke < 65 Years for Small and Large Vessel Disease 128

Table 7.2
Estimated Sample Size Requirement for Studies Using Specific Stroke Subtypes and Age Groups 129

Table 7.3
Recommendations for the Model Genetic Association Study 133
Abbreviations

Adenosine
Adenosine Diphosphate
Computerised Tomography
Cytosine
Deoxyribonucleic Acid
Diastolic Blood Pressure
Disability Adjusted Life Year
Glycoprotein
Guanine
High Density Lipoprotein
Human Platelet Alloantigen
Lacunar Syndrome
Leucine
Low Density Lipoprotein
Magnetic Resonance Imaging
Messenger Ribodeoxynucleic Acid
Metaloproteases
Methionine
National Heart Foundation
North East Melbourne Stroke Incidence Study
Oxfordshire Community Stroke Project
Paraoxonase
Partial Anterior Circulation Syndrome
Patent Foramen Ovale
Perth Community Stroke Study
Plasminogen Activator Inhibitor
Polymerase Chain Reaction
Population Research and Outcome Studies
Posterior Circulation Syndrome

A
ADP
CT
C
DNA
DBP
DALY
Gp
G
HDL
HPA
LS
L
LDL
MRI
MRNA
MMP
M
NHF
NEMESIS
OCSP
PON1
PACS
PFO
PCSS
PAI
PCR
PROS
PCS
Ribonucleic Acid
Sequence Specific Primer Polymerase Chain Reaction
Sibling Transmission Disequilibrium Test
Single Nucleotide Polymorphism
Spontaneously Hypertensive Rat
Stroke Prone Spontaneously Hypertensive Rat
Systolic Blood Pressure
The Trial of ORG 10172 in Acute Stroke Treatment
Threonine
Thymidine
Tissue Plasminogen Activator
Total Anterior Circulation Syndrome
Transcription factor IID
Transient Ischemic Attack
Transmission Disequilibrium Test
World Health Organization

RNA
SSP-PCR
S-TDT
SNP
SHR
SP-SHR
SBP
TOAST
Thr
T
TPA
TACS
TFIID
TIA
TDT
WHO