LUND UNIVERSITY

Hereditary risk factors for stroke in humans - Association studies with emphasis on familial and genotypic factors

Lövkvist, Håkan

2012

Link to publication

Citation for published version (APA):
Lövkvist, H. (2012). *Hereditary risk factors for stroke in humans - Association studies with emphasis on familial and genotypic factors*. [Doctoral Thesis (compilation), Neurology, Lund]. Neurology, Dept of Clinical Science, Lund.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 17. Dec. 2025

Hereditary risk factors for stroke in humans

Association studies with emphasis on familial and genotypic factors

LUND
UNIVERSITY

by

Håkan Lövkvist

DOCTORAL DISSERTATION

With due permission of the Faculty of Medicine at Lund University to be publicly defended
on September 14, 2012 at 9:00 am, in the Segerfalk Lecture Hall,
Wallenberg Neuroscience Center, Sölvegatan 17, Lund,
Skåne University Hospital, Lund, Sweden

Faculty opponent:
Associate Research Professor
Hanne Christensen
Department of Neurology
University of Copenhagen
Bispebjerg Hospital

Organization

LUND UNIVERSITY
Department of Clinical Sciences, Lund

Neurology
Faculty of Medicine
Skåne University Hospital
221 85 LUND

Document name

DOCTORAL DISSERTATION

Date of issue
2012-09-14

Author(s)
Håkan Lövkvist

Sponsoring organization

Title and subtitle

Hereditary risk factors for stroke in humans – Association studies with emphasis on familial and genotypic factors.

Abstract

Background
Stroke is a serious vascular disorder that comprises intracerebral hemorrhages, subarachnoid hemorrhages and ischemic stroke (IS). The etiology of stroke, including hereditary components induced by cellular mechanisms, is therefore a research field of vital importance. The aim of this thesis was to assess possible genetic impact on stroke risk. We thus evaluated family history of stroke-related individual traits as well as allelic variations in selected candidate genes.

Methods
Association studies, primarily of Lund Stroke Register data, were fundamentals in the analyses in order to find possible genetic association with stroke risk. Hereditary risk factors, based on family-history data and candidate gene studies, were considered. Statistical methods for assessing the data, including basic chi-square tests of two-by-two tables as well as various logistic regression approaches and meta-analysis applications using the DerSimonian-Laird method, were used. TOAST subtypes of ischemic stroke were considered when available. Also, 25 coronary artery disease (CAD) susceptible SNPs from various genetic loci were joined together in risk scores and tested against IS risk by logistic regression.

Results
Paper I: The prevalence of stroke or TIA among first-degree relatives may affect the proband's stroke risk (odds ratio, OR=1.74; 95% confidence interval, CI: 1.36-2.22), especially when considering mothers and offsprings (OR=2.04; 95% CI: 1.30-3.20). No such association was seen at all between spouses. Papers II-III: SNP rs12188950 was significantly associated with IS in Paper II (OR=0.72; 95% CI: 0.58-0.91; N=1324), but not in Paper III (OR=0.93; 95% CI: 0.83-1.05; N=4692) where a different sample representing an analogous Swedish population was used. Paper IV: SNP rs4977574 on chromosome 9p21 was related to IS risk (OR=1.12; 95% CI: 1.04-1.20) and large vessel disease risk (OR=1.36; 95% CI: 1.13-1.64). This genetic effect of chromosome 9p21 on IS was however already known. None of the other 24 SNPs or compiled risk scores were significant.

Conclusions and discussion
Significant transmission of stroke from parents to offsprings but not between spouses suggests a genetic inheritance component. But, for SNPs representing some particular carefully selected genes, the findings regarding possible impact on IS were not that clear: Ambiguous results were obtained, many significance tests were also negative. Moreover, we could not generally find significant associations between SNPs susceptible for CAD and IS risk.

Key words
allelic association, candidate genes, case-control, family history, ischemic stroke, polymorphisms.

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language

ISSN and key title
1652-8220

ISBN
978-91-87189-26-5

Recipient's notes

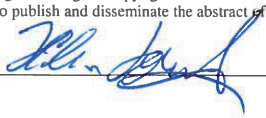
Number of pages
138

Price

Security classification

Distribution by (name and address)

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature 

Date 2012-08-08

Hereditary risk factors for stroke in humans

Association studies with emphasis on familial and genotypic factors

LUND
UNIVERSITY

Håkan Lövkvist

© 2012 Håkan Lövkvist

Lund University, Faculty of Medicine Doctoral Dissertation Series 2012:63

ISSN 1652-8220
ISBN 978-91-87189-26-5

Printed in Sweden by Media-Tryck, Lund University
Lund 2012

To my mother

Table of Contents

"The essence of the beautiful is unity in variety."

Felix Mendelssohn,
German composer, conductor, etc.
Born in 1809.
Died of a stroke in 1847.

Abstract

Background

Stroke is a serious vascular disorder that comprises intracerebral hemorrhages, subarachnoid hemorrhages and ischemic stroke (IS). The etiology of stroke, including hereditary components induced by cellular mechanisms, is therefore a research field of vital importance. The aim of this thesis was to assess possible genetic impact on stroke risk. We thus evaluated family history of stroke-related individual traits as well as allelic variations in selected candidate genes.

Methods

Association studies, primarily of Lund Stroke Register data, were fundamentals in the analyses in order to find possible genetic association with stroke risk. Hereditary risk factors, based on family-history data and candidate gene studies, were considered. Statistical methods for assessing the data, including basic chi-square tests of two-by-two tables as well as various logistic regression approaches and meta-analysis applications using the DerSimonian-Laird method, were used. TOAST subtypes of ischemic stroke were considered when available. Also, 25 coronary artery disease (CAD) susceptible SNPs from various genetic loci were joined together in risk scores and tested against IS risk by logistic regression.

Results

Paper I: The prevalence of stroke or TIA among first-degree relatives may affect the proband's stroke risk (odds ratio, OR=1.74; 95% confidence interval, CI: 1.36-2.22), especially when considering mothers and offsprings (OR=2.04; 95% CI: 1.30-3.20). No such association was seen at all between spouses. **Papers II-III**: SNP rs12188950 was significantly associated with IS in Paper II (OR=0.72; 95% CI: 0.58-0.91; N=1324), but not in Paper III (OR=0.93; 95% CI: 0.83-1.05; N=4692) where a different sample representing an analogous Swedish population was used. **Paper IV**: SNP rs4977574 on chromosome 9p21 was related to IS risk (OR=1.12; 95% CI: 1.04-1.20) and large vessel disease risk (OR=1.36; 95% CI: 1.13-1.64). This genetic effect of chromosome 9p21 on IS was however already known. None of the other 24 SNPs or compiled risk scores were significant.

Conclusions and discussion

Significant transmission of stroke from parents to offsprings but not between spouses suggests a genetic inheritance component. But, for SNPs representing some particular carefully selected genes, the findings regarding possible impact on IS were not that clear: Ambiguous results were obtained, many significance tests were also negative. Moreover, we could not generally find significant associations between SNPs susceptible for CAD and IS risk.

List of papers

I
Lindgren A, **Lövkvist H**, Hallström B, Höglund P, Jönsson AC, Kristoffersson U, Luthman H, Petersen B, Norrving B.
Prevalence of stroke and vascular risk factors among first-degree relatives of stroke patients and control subjects.
Cerebrovasc Dis. 2005;20:381-387.

II
Lövkvist H, Smith J G, Luthman H, Höglund P, Norrving B, Kristoffersson U, Jönsson AC, Lindgren A.
Ischaemic stroke in hypertensive patients is associated with variations in the *PDE4D* genome region.
Eur J Hum Genet. 2008;16:1117-1125.

III
Lövkvist H, Olsson S, Höglund P, Melander O, Jern C, Sjögren M, Engström G, Smith J G, Hedblad B, Andsberg G, Delavaran H, Jood K, Kristoffersson U, Luthman H, Norrving B, Lindgren A.
A large-sample assessment of possible association between ischaemic stroke and rs12188950 in the *PDE4D* gene.
Eur J Hum Genet. 2012;20:783-789.

IV
Lövkvist H, Sjögren M, Höglund P, Engström G, Jern C, Olsson S, Smith J G, Hedblad B, Andsberg G, Delavaran H, Jood K, Kristoffersson U, Norrving B, Melander O, Lindgren A.
Are genetic variations reported by the CARDIoGRAM study also associated with ischemic stroke?
Manuscript.

Abbreviations

ALOX5AP	Arachidonate 5-Lipoxygenase Activating Protein.
CAD	Coronary Artery Disease.
CARDIoGRAM	Coronary Artery Disease Genome-Wide Replication and Meta-analysis (a transatlantic consortium for large-scale multi-center studies of genetic data).
CE	Cardioembolism (mainly a defined TOAST subclass of ischemic stroke).
CHD	Coronary Heart Disease.
CI	1. Confidence Interval. 2. Cerebral Infarction.
GWAS	Genome-Wide Association Study.
HWD	Hardy-Weinberg Disequilibrium.
HWE	Hardy-Weinberg Equilibrium.
ICH	Intracerebral Hemorrhage.
IS	Ischemic Stroke.
LD	Linkage Disequilibrium.
LSR	Lund Stroke Register.
LVD	Large Vessel Disease (mainly a defined TOAST subclass of ischemic stroke).
MAF	Minor Allele Frequency.
MDC	The Malmö Diet and Cancer Study.
MHC2TA	Major Histocompatibility Complex class II Transactivator.
MI	Myocardial Infarction.
OR	Odds Ratio.
PDE4D	Phosphodiesterase 4D.
RAF	Risk Allele Frequency.

RERI	Relative Excess Risk due to Interaction.
RR	Relative Risk.
SAH	Subarachnoid Hemorrhage.
SAHLSIS	The Sahlgrenska Academy Study on Ischemic Stroke.
SAS	Statistical Analysis System (a statistical software).
SNP	Single Nucleotide Polymorphism.
SPSS	Statistical Package for the Social Sciences (a statistical software).
SVD	Small Vessel Disease (mainly a defined TOAST subclass of ischemic stroke).
TIA	Transient Ischemic Attack.
TOAST	Trial of Org 10172 in Acute Stroke (a nomenclature for classifying ischemic stroke into subclasses).
VSMC	Vascular Smooth Muscle Cells.

Populärvetenskaplig sammanfattning

Stroke, eller slaganfall, är en sammanfattande benämning på sjukdomar som drabbar blodförsörjningen i hjärnan. Varje år insjuknar cirka 30000 personer i Sverige i stroke, av dessa drabbas cirka 15% av blödningar medan resterande 85% får infarkter, dvs blodpropp i hjärnan.

Stroke upptar fler vårddagar i slutenvården än någon annan somatisk sjukdom. Medianåldern för insjuknande är ungefär 75 år, 20% av alla insjuknanden sker före 65 års ålder. Risken för en enskild att drabbas av stroke har minskat under 2000-talets första decennium, men samtidigt räknar man med att antalet strokepatienter totalt sett kommer att öka under de närmaste decennierna till följd av att vi generellt sett lever längre.

Riskfaktorer för stroke

Det finns ett flertal väldokumenterade faktorer som kan öka risken för en stroke. Som exempel kan nämnas hypertoni (högt blodtryck), diabetes, hjärt- och kärlsjukdom, rökning och förhöjda kolesterolvärden. En del av dessa riskfaktorer, som har ett nära samband med cirkulationsrubbningar och sjukdomstillstånd som till exempel hjärtinfarkt, har visat sig vara nära förknippade med biologiska processer i människokroppen som gått i arv i släkten enligt studier som gjorts.

På senare år har man i allt högre utsträckning uppmärksammat denna nedärvning, hereditet, inom strokeforskningen. Man har bland annat dels sett på familjehistorik bakom insjuknande i stroke, dels undersökt om vissa bestämda gener kopplade till sjukdomar i kroppens blodkärl också kan påverka risken att insjukna i stroke. Avhandlingen innehåller båda dessa typer av studier, som ofta benämns familjehistorikstudier respektive kandidatgen-studier.

Vad är en genetisk markör, och vad är en ”SNP”?

En genetisk markör visar en observerbar nedärvd egenskap hos individen. Låt oss ta ett exempel: Vi misstänker att personer med blå ögon bär på en viss nedärvd egenskap som ökar deras risk för att drabbas av en viss sjukdom. Vi vill undersöka om en sådan risk föreligger genom att göra en observationsstudie. Detta gör vi genom att jämföra sjukdoms-

förekomsten bland en stor grupp individer med blå ögon med en referensgrupp som har en annan ögonfärg. Variabeln (kännetecknet) ögonfärg är här en genetisk markör.

I detta avhandlingsarbete handlar begreppet genetiska markörer om *polymorfier* som återfinns i vissa *gener*.

En gen kan sägas bära på en bestämd egenskap, t ex för ett visst enzym som antingen kan förstärka eller förebygga en viss typ av sjukdomar och ohälsotillstånd. En gen finns representerad på ett visst identifierbart ställe (region) i en kromosom, som kan beskrivas som en lång DNA-tråd. I denna region kan man återfinna en form av genetiska markörer som brukar benämnas polymorfier, eller *Single Nucleotide Polymorphisms*, förkortat SNP (uttalas ofta ”snipp”).

Alleler finns i två varianter, man brukar tala om *referens-alleler* och *alternativ-alleler* (ibland används uttrycken *major alleles* och *minor alleles*). Om V är en referens-allel och v är en alternativ-allel för en viss SNP, så kan en människa bära på tre alternativa allelrepresentationer för just den SNP:en eftersom vi får en allel från modern och en från fadern. En bestämd individ kan således representeras av antingen

VV – dvs vara *homozygot* för referens-allelen,

Vv – dvs vara *heterozygot* (Vv och vV är likvärdiga), eller

vv – dvs vara *homozygot* för alternativ-allelen.

Fördelningen av dessa tre kategorier av homo- och heterozygota individer jämförs mellan en grupp patienter (fall) och en grupp personer som inte har insjuknat i stroke (kontroller). Med lämpliga skattningsmetoder kan man under vissa givna förutsättningar testa om det går att med ett givet krav på statistisk säkerhet fastställa en skillnad mellan dessa fall och kontroller för att därmed kunna dra några bestämda slutsatser kring en viss sjukdomsförekomst och dess eventuella förklaring i genetiska termer.

Avhandlingens delarbeten

Delarbete I är en familjehistorikstudie som omfattar analys av strokepatienters och friska kontrollpersoners förstagradssläktingar (dvs föräldrar, barn, syskon och halvsyskon). Även sambor och äkta makar ingår. Totalt 925 patienter och 286 kontrollpersoner hade ombetts att ange uppgifter om eventuell förekomst av stroke eller TIA (transitorisk ischemisk attack, som är en övergående blodproppsbildning i hjärnan) bland släktingar. Av dessa lämnade 606 patienter och 261 kontrollpersoner sådana uppgifter om 4972 förstagradssläktingar och 738 sambor eller äkta makar.

Patientgruppens förstagradssläktingar visade sig ha högre förekomst av stroke eller TIA samt högt blodtryck än vad som var fallet med motsvarande släktingar till kontrollpersonerna. Det verkade också kunna finnas en antydan till högre grad av ärftlighet av stroke på mödernet än på fädernet. Mellan äkta makar kunde vi däremot inte bekräfta något sådant samband. Detta stärker vår tro på att de samband vi fann är renodlat hereditära, utan någon påtaglig inblandning av socialt arv och gemensamma miljöfaktorer inom familjen.

Delarbete II är en kandidatgenstudie. Med detta menas att ett antal på förhand utvalda SNP:ar undersöks. Vi analyserade sålunda nio SNP:ar som representerar tre olika

gener. Vi ville se om dessa hade något statistiskt samband med förekomsten av hjärninfarkter (ischemisk stroke). De tre generna har beteckningarna *PDE4D*, *ALOX5AP* och *MHC2TA*, och de påverkar vissa biologiska processer i kroppen som kan påverka risken för sjukdom i blodkärlen som i sin tur ökar risken för en stroke. Vi fokuserade särskilt på en bestämd SNP som har benämningen rs12188950 (betecknas ibland även SNP45) och som finns representerad i genen *PDE4D*. Denna SNP var även föremål för en metaanalys av 13 vetenskapliga artiklar. En metaanalys är en studie av andra studier, där man smälter samman resultaten. Poängen med denna sammanslagning av tidigare resultat är att man får ett betydligt större material. Det ger en ökad statistisk styrka och därmed också bättre underlag för fastställande av även ganska svaga statistiska samband. En viktig förutsättning är dock att de inkluderade studierna i metaanalysen är något sånär lika beträffande definition av sjukdomsbild, deltagarnas demografiska sammansättning mm. Totalt inkluderade vi 932 patienter och 396 kontroller i vår egen studie, medan metaanalysen omfattade 6221 patienter och 6750 kontrollpersoner.

Vi fann att personer som är bärare av alternativ-alleler för rs12188950 kan vara skyddade mot stroke orsakad av hjärninfarkt. Denna skyddande effekt ser ut att vara särskilt påtaglig bland personer som lider av högt blodtryck. Övriga åtta SNP:ar visade opåvisbart svaga eller obefintliga samband med risken att drabbas av hjärninfarkt. Metaanalysen visade att resultaten beträffande SNP rs12188950 varierade i betydande grad från studie till studie (och därmed troligen också från befolkningsgrupp till befolkningsgrupp).

Delarbete III tar vid där delarbete II slutar. Vi såg närmare på SNP rs12188950 genom att även inkludera patienter och kontrollpersoner från Malmö kost-cancer-studie och Sahlgrenska sjukhusets ischemiska stroke-studie i och kring Göteborg. Dessutom exkluderade vi de patienter och kontroller från LSR som var med i delarbete II. Vårt eget material omfattade därmed totalt 2599 patienter med hjärninfarkt och 2093 ”friska” kontrollpersoner. Vi gjorde även här en metaanalys där vårt uppdaterade material nu omfattade 17 artiklar innehållande 20 urskiljbara demografiska populationsgrupper för analys.

Resultaten i delarbete III såg annorlunda ut om man jämför med delarbete II. Vi kunde inte statistiskt påvisa att rs12188950 kan ha en effekt på risken för hjärninfarkt. Vi kunde inte heller se något sådant tydligt samband när vi analyserade de 1727 patienter och 1056 kontrollpersoner i vårt material som led av högt blodtryck. Vi kunde se en svag antydan till beskyddande effekt när 382 patienter och 321 kontroller under 55 års ålder på motsvarande sätt undersöktes, men det sambandet är för svagt för att man ska kunna dra någon generell slutsats.

Den utökade metaanalysen i delarbete III ledde inte fram till några andra slutsatser än de som framkom i motsvarande analys som presenterades i delarbete II. En lite annorlunda strukturering av datamaterialet, med delpopulationer snarare än publicerade studier i fokus, gav läsaren möjlighet att illustrativt bilda sig en uppfattning om eventuella skillnader som kan bero på kön, nationalitet eller etnicitet. Några tendenser till skillnad mellan dessa undergrupper som skulle kunna vara intressanta att ta tag i och undersöka närmare kunde dock inte skönjas.

Delarbete IV innehåller en observationsstudie som, i likhet med delarbete III, omfattas av patienter med stroke pga hjärninfarkt och ”friska” kontroller från Lunds strokeregister, Malmö kost-cancer-studie och Sahlgrenska sjukhusets ischemiska strokestudie i Göteborg med omnejd. I denna kandidatgenstudie såg vi på 25 SNP:ar från olika genetiska regioner. Gemensamt för de 25 SNP:arna är att de i tidigare studier som gjorts visat sig ha ett tydligt samband med hjärt- och kärlsjukdomar. Vi var intresserade av att undersöka om dessa SNP:ar även kunde ha någon effekt på risken att insjukna i *hjärn*infarkt.

Trots att vi hade ett relativt stort material att tillgå, med ett bruttourval på 3986 patienter och 2459 kontrollpersoner, kunde vi inte finna några större tecken på ett sådant samband. Endast ett genetiskt område i kromosom 9, nämligen 9p21 som vi redan har sett på i en tidigare studie, visade sig vara påvisbart associerat med förekomsten av hjärninfarkt genom SNP rs4977574. En viss typ av stroke, s.k. storkärlssjukdom som har ett nära släktskap med hjärt- och kärlsjukdomar som exempelvis åderförkalkning och hjärtinfarkt, visade sig ha ett ännu starkare samband med denna enda genetiska markör i den genetiska regionen 9p21.

En konstruerad riskfaktor som byggde på en sammanvägning av de 25 SNP:arna, en s.k. risk score, visade sig inte ha något påvisbart samband med varken hjärninfarkt i stort eller ovan nämnda storkärlssjukdom.

Slutsatser

Sammanfattningsvis har våra statistiska analyser lett fram till att vi tämligen klart kan konstatera att förekomsten av stroke bland annat påverkas av genetiska faktorer. Det finns påfallande tydliga indikationer på att barn till föräldrar som haft stroke eller TIA själva löper högre risk att få en stroke. Nedärvning från moderns sida ser ut att vara särskilt tydlig. Även vissa SNP:ar, t ex rs4977574 och i viss mån även rs12188950, visar att det finns en genetisk komponent som spelar en roll för risken att insjukna i hjärninfarkt, den typ av stroke som vi har analyserat i våra kandidatgenstudier.

Det finns sannolikt en del ytterligare samband som ännu inte går att statistiskt verifiera. Dessa skulle kanske kunna påvisas om vi hade tillgång till ett större material. Särskilt intressant blir det när hjärninfarkt delas in i undergrupper. Ett exempel är kardiell emboli, som innebär att små klumpar av levrat blod lossnar från hjärtats förmak och transporteras upp till hjärnan där de kan sätta sig som blodproppar och därmed strypa blodtillförseln med vävnadsdöd som följd. Här skulle studier på gener som ligger bakom exempelvis förmaksflimmer kunna vara av betydelse då förmaksflimmer är en känd riskfaktor för just kardiell emboli.

Introduction

Stroke is the third most common cause of death in the Western World today. Every year about 30000 persons are suffering from stroke in Sweden, and about one fifth of those are below 65 years of age (**Toivanen 2011, Warlow 2003**). The consequences of stroke in terms of e.g. production loss, sickness absence from work, negative impact on family relationships and social life are immense, and in Swedish domestic expenditure terms one may talk about losses of about 10 billion SEK or more (**Daniel 2009, Ghatnekar 2004**). While increasing stroke incidence rates were observed between 1989 and 2000 (**Medin 2004**), a decline in age-standardized stroke incidence rates is suggested for the beginning years of the 21st century (**Hallström 2008**).

The concept of stroke

Stroke can be defined as rapid loss of brain function (or functions) caused by disturbance in the focal blood supply to the brain (**WHO 1988**). There are two main types of stroke: hemorrhages (bleedings) and ischemic strokes (**Fagius 2006**). Ischemic strokes are sometimes also denoted cerebral infarctions. This definition is however inadequate because cerebral infarctions also include *silent* cerebral infarctions while ischemic strokes do not.

A hemorrhage is a leakage of blood caused by a rupture of a blood vessel. Such a rupture may have been preceded by a cerebral aneurysm, often located in the circle of Willis and its branches (**Figure 1**). This is common in the case of *subarachnoid hemorrhage* (SAH), in which the blood leakage will spread within the subarachnoid space (the cerebrospinal fluid-filled space between the pia mater membrane and the arachnoid membrane, see **Figure 2**). A bleeding due to blood vessel rupture may also occur within the brain tissue itself. That kind of hemorrhage is often denoted *intracerebral hemorrhage* (ICH) and it is categorized as stroke if occurring spontaneously.

An *ischemic stroke* (IS) is a sudden loss of function due to loss of blood supply to an area of the brain that controls that function. Usually, IS is directly caused by a partial or complete blockage of an artery that supplies the brain. IS can be categorized into different subgroups due to different etiologic structures. There are thus four important such main types of IS that are defined by a well-established international nomenclature, the Trial of Org 10172 in Acute Stroke (TOAST) (**Adams 1993, Grau 2001**):

Figure 1: The anatomy of the circle of Willis and its branches (Source: Wikimedia commons public domain 2007).

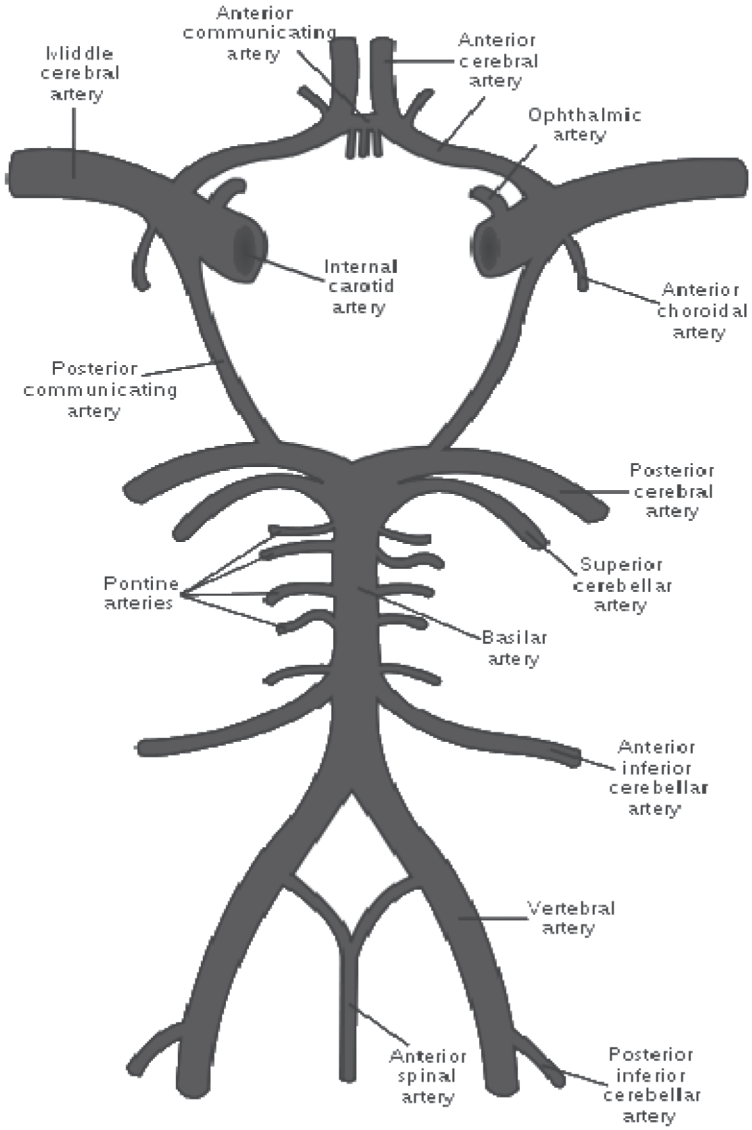
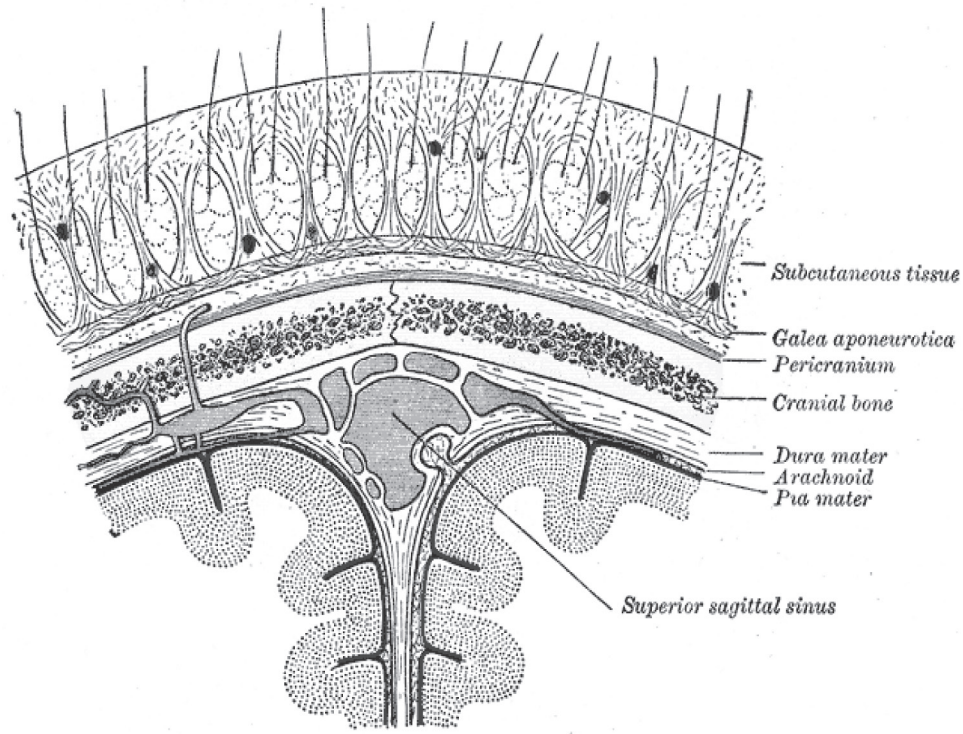


Figure 2: A section of the human scalp showing the membranes (dura mater, arachnoid, pia mater) between the cranium bone and the cortex, covered by pia mater. (Source: Gray's Anatomy 1918).



– Cardioembolism

– Large vessel disease

– Small vessel disease

– Cryptogenic IS

Cardioembolism (CE), also denoted embolic stroke, refers to the blockage of an artery by an arterial embolus, a travelling particle or debris in the arterial bloodstream originating from the heart. An embolus is most frequently a thrombus, and in about 50% of all CEs a thrombus-induced blockage caused by atrial fibrillation is the primary reason of the stroke. CE may also be caused by large vessel dysfunctions, valvular heart diseases, prosthetic valves and tumors in the heart (**Hirsh 2005**).

Large vessel disease (LVD), or large artery atherosclerosis, is a disease usually caused by atherosclerotic plaques that involves the common and internal carotid arteries, the vertebral artery and the circle of Willis (**Bevan 2005, Kuo 1992**). Other diseases that may

cause thrombus formation in the large vessels include e.g. dissection, fibromuscular dysplasia and rare inflammatory diseases of the arterial vessel wall.

Small vessel disease (SVD) refers to a group of pathologic processes that affect small arteries, arterioles, venules and capillaries of the brain (**Pantoni 2010**). Age-related and hypertension-related SVDs and cerebral amyloid angiopathy are the most common forms (**Khan 2007, Ross 1993**). SVD may cause lesions located in the subcortical cerebral structures. Lacunar infarcts, white matter lesions, large hemorrhages, and microbleeds are examples of such lesions.

Cryptogenic stroke can be defined as IS without identified structural etiology. One theory suggests a possibly identifiable etiological subtype that can be explained by the occurrence of emboli caused by blood shunting from the right to the left atrium of the heart through *patent foramen ovale* (**Tobis 2005, Lamy 2002**).

Transient Ischemic Attack

A transient ischemic attack (TIA) is a milder, reversible form of cerebral ischemia. TIA is also in itself considered to be a strong risk factor for more fatal forms of cerebrovascular diseases. Also, although TIA is considered as a reversible occurrence of artery occlusion that persists between a few minutes and 24 hours, it should be regarded as a serious predictor of IS and an indicator for considering adequate prophylactic treatment.

Hereditary aspects of stroke

It is evident that stroke onsets (hemorrhages as well as ischemic strokes) are manifestations of various vascular disorders. Consequently, the most influential risk factors for stroke (apart from age, smoking habits and previous TIAs) are hypertension, diabetes mellitus, hypercholesterolemia, myocardial infarction and atrial fibrillations (**Flossmann 2004, Wolf 1998, Al Mamun 2004**). These vascular risk factors may possibly also explain certain etiologic structures behind stroke. This opens for assessments on e.g. the particular main TOAST subtypes of IS described in the section above.

Two kinds of hereditary studies are carried out in this thesis: A family history study (Paper I) and a candidate gene study (Papers II-IV). Both of these approaches consider the vascular disorders mentioned above as parts of the etiology behind the occurrence of stroke.

Familial inheritance

Previous studies have suggested that familial inheritance of the vascular risk factors for stroke plays an essential part for familial inheritance of stroke risk itself (**Jerrard-Dunne 2003, Marenberg 1994, Meschia 2004**). In some rare cases a well-defined stroke com-

ponent is present (**Tournier-Lasserve 1991**). It is also notable that other kinds of diseases than stroke, e.g. particular cancers that are parallel to stroke by the means of inheritance in families (especially when occurring in younger ages) have been assessed and documented previously by the use of equivalent methodological approaches (**Knudson 1985**).

Candidate genes for stroke

Previous genome-wide association studies (GWASs) have provided essential information about cellular mechanisms that may influence inflammatory processes with impact on vascular disorders that are associated with some particular genetic regions on the human chromosomes (e.g. **Gretarsdottir 2003, Helgadottir 2004, Schunkert 2011**). A fundament in our candidate gene studies was to use these previous findings in order to select genetic markers appropriate for analysis. We thus selected SNPs that possibly might affect IS risk by using knowledge from these findings. We then carried out allelic association tests to either confirm or reject our hypotheses. We mainly concentrated on examining SNPs within the following genetic regions:

PDE4D, an enzyme that may reduce the proliferation of vascular smooth muscle cells (**Conti 2003, Pan 1994**). This might in turn affect the occurrence of hypertension that is a potential risk factor for IS. An Icelandic study, published in 2003 and frequently referred to, has suggested that the *PDE4D* region on chromosome 5q12 might be associated with stroke related to atherosclerosis (**Gretarsdottir 2003**). These findings have later been replicated or utilized for further analyses in other studies, also including meta-analyses (e.g. **Bevan 2005, Bevan 2008, Brophy 2006, Kostulas 2007, Kuhlenbäumer 2006, Meschia 2005, Nakayama 2006, Nilsson-Ardnor 2005, Quarta 2009, Rosand 2006, Salaheen 2005, Song 2006, Staton 2006, van Rijn 2005, Woo 2006, Yoon 2011, Zee 2006**).

ALOX5AP, a protein that encodes the 5-Lipoxygenase-Activating Protein (FLAP) enzyme, which regulates production of leukotriene inflammatory mediators, shown to be of relevance for occurrence of atherogenesis and plaque rupture (**Helgadottir 2004, Kostulas 2007, Löhmussaar 2005, Meschia 2005**).

MHC2TA, that encodes a protein involved in regulation of expression of an antigen associated with increased susceptibility to myocardial infarction. It is reported as a genetic marker for diseases with inflammatory etiology (including rheumatoid arthritis, multiple sclerosis and myocardial infarction), and might therefore be a considerable vascular risk factor for IS (**Ghaderi 2006, Lindholm 2006, Swanberg 2005**).

CDKN2A and ***CDKN2B***, that are known as key regulators of the cell cycle. Their possible influence on increased risk of heart disease may be mediated through reduced regrowth of arterial intimal cells, a phenomenon implicated in the development of atherosclerosis. (**Bishop 2002, Cluett 2009, Harismendy 2011, Helgadottir 2007, Ishii**

1999, Smith 2009, Visel 2010). These genetic loci are found within the 9p21 chromosome. Within the same genome region, it has also been reported that ***ANRIL*** is expressed in cells and tissues affected by atherosclerosis (**Broadbent 2008, Holdt 2010**).

Demographic aspects of stroke and genetics

It is generally known that demographic and socioeconomic structures in a population have a non-ignorable influence on stroke risk and stroke mortality (**The Swedish National Board of Health and Welfare 2012**). Age, gender, family income and ethnicity are factors that may affect stroke incidence. Consequently, these factors have been considered in many previous reports on genetic impact on stroke (**Brophy 2006, Domingues-Montanari 2010, Li 2008, Meschia 2005, Song 2006, Welin 1987**). A report that compared data from 18 populations in 10 countries by utilizing data from the WHO MONICA (World Health Organization Monitoring Trends and Determinants in Cardiovascular Disease) Project confirms the assumption that differences regarding stroke risk due to nationality and demographic subgroupings risk may exist (**Thorvaldsen 1995**).

Stroke and age

Age is a strong risk factor for stroke. The risk of being hit by a first-time stroke increases exponentially from about 30 per 100000 individuals each year at 30-39 years of age, to about 2000-3000 per 100 000 each year at ages above 85 (**Fagius 2006**). Vascular risk factors, such as hypertension, diabetes mellitus, smoking and obesity, may partially confer an increased risk for cardiovascular diseases and atherosclerosis, and thus also stroke (primarily IS), among individuals aged 50-75 years (**Lloyd-Jones 2006**).

In recent years researchers have also focused on the etiology of stroke among younger stroke patients. Some risk factors, e.g. the use of oral contraceptives, or the presence of unusual (or rare) congenital malformations of the heart such as patent foramen ovale or atrial septal aneurysms, may increase the risk of stroke even among younger persons (**Song 2006, Lamy 2002, Cerrato 2004**). Cryptogenic strokes and CEs, as well as SAHs that are more common among younger adults (**vanGijn 2007**), are motivating subjects of study regarding younger stroke patients.

Stroke and gender

Possible differences between men and women regarding the risk of being suffering from stroke may to some extent be associated with life-style indicators, such as smoking habits, alcohol intake and type of occupation. Findings from a pre-clinical study suggested that estrogen may milder stroke outcomes during vascular occlusions among pre-menopausal female rats (**Alkayed 1998**). Another publication, addressing to the Framingham Study, suggested that atrial fibrillation might cause a higher risk for IS

among women than among men when considering an age-standardized coronary heart disease subpopulation (**Wolf 1991**).

There are also other aspects of stroke and gender. A previous family history study suggested that maternal history of stroke might affect stroke risk, especially among women (**Tentschert 2003**). This finding was considered in Paper I.

Stroke and ethnicity

Studies on the U.S. population have suggested an augmented risk for stroke (IS as well as SAH and ICH) among the African American population compared to Americans with European ancestors (**Gorelick 1998, National Stroke Association 2012**). In Sweden, selected groups of immigrants might be associated with stroke-related vascular disorders, such as coronary artery diseases and coronary heart diseases (**Borné 2012, Gadd 2005**). This is also reported in a study linked to the STROMA register project (**Khan 2004**). The findings from the studies on Afro Americans in the U.S. and immigrants in Sweden have some similarities: Socioeconomic status, including education, occupation, income and life-stile, known to affect the rates of known risk factors such as hypertension, diabetes mellitus and heart failures, seems to be a contributing factor.

The analyses in brief

In Paper I we analyzed the genetic (and environmental) inheritance of stroke within families. Questionnaires allowing for self-report regarding possible stroke-history within the family is a rather rough diagnostic tool. The patients and control subjects usually know what a stroke is, but in most cases they cannot give detailed information about family history of hemorrhages or IS with or without a specific subtype. Nevertheless, family history-based assessments of genetic influence of stroke is valuable as an aid for allocating optimal subsets of data when performing further studies based on candidate genes (**Jerrard-Dunne 2003**).

In Papers II and III we used SNP data from genes encoding the phosphodiesterase 4D gene (*PDE4D*). In Paper II we also put emphasis on the 5-lipoxygenase activating protein (*ALOX5AP* on chromosome 13q12-13), and the major histocompatibility complex class II transactivator (*MHC2TA* on chromosome 16). In Paper IV we focused on a wide spectrum of genes with one similarity: their impact on CAD according to previous studies

Keeping the ambiguous results regarding one particular SNP in the *PDE4D* gene in mind (namely rs12188950, also denoted SNP45), we performed a meta-analysis for that single SNP on a Caucasian population, including 13 studies, in Paper II. That meta-analysis was updated in Paper III, where 17 published studies were involved. Our own studies were included in these meta-analyses.

"We do not play the piano with our fingers but with our mind."

Glenn Gould,
Canadian pianist.
Born in 1932.
Died of a stroke in 1982.

Aims

This thesis can be regarded as a cross-sectional work that includes four main fields of science:

– A genetic field comprising a family history approach, i.e. assessments intended to find possible association of increased stroke risk between close, biologically related members of a family.

– A genetic field comprising the detection of appropriate candidate genes, known from previous studies to be involved in cellular mechanisms that might affect known vascular risk factors of stroke.

– A pathogenetic field, including neurologic and cardiologic etiology in individuals with or without stroke.

– A statistical field, including application of quantitative methods for assessing data and knowledge of how to interpret the results of these assessments.

When taking these scientific disciplines into account, the specific aims could be stated as follows:

– To examine a possible relation of stroke and vascular risk factors for stroke between family members (Paper I)

– To examine a possible environmental effect of inheritance besides the purely genetic effect when analyzing the effect of family history on stroke (Paper I).

– To explore possible gender-dependent factors influencing the inheritance of stroke (Paper I).

– To replicate previous findings that provided conflicting results by using large sample sizes (including the utilization of meta-analyses) to see if the previously thus assessed polymorphism may be associated with IS (Papers II and III).

– To examine polymorphisms from a specific genetic region joined together to a haplotype to find possible association with the risk of (ischemic) stroke (Paper II).

– To examine polymorphisms from completely different genetic loci, but with a common impact on a specific histopathologic mechanism and thereby susceptibility for coronary artery disease, to find possible association also with IS (Paper IV).

– To cope with particular statistical problems, such as Hardy-Weinberg disequilibrium, that might appear when handling population data providing genetic information (Papers II-IV, see also the *Further statistical considerations* section below).

Methods

Study population

Lund Stroke Register

The four studies that form this thesis are all principally based on analyses of data from Lund Stroke Register (LSR). LSR is a population-based data base aimed for case-control studies, and it includes all patients with first-ever stroke since 2001 from the local catchment area of the Skåne University Hospital located in Lund (**Hallström 2007**). Each patient was initially matched by one or more control subjects, randomly selected by stratification variables age and sex from Befolkningsregistret, the Swedish Population Register (e.g. **Statistics Sweden 2012**). The control subjects were residents of the same geographical area as the patients when selected.

The patients were consecutively selected from March 1, 2001. The inclusion of control subjects has been done in 2001 and 2006. The number of participants recruited from LSR in the studies referring to Papers I-IV were:

For Paper I (inclusion between March 1, 2001 and May 31, 2003): 606 stroke patients and 261 control subjects, providing questionnaire data regarding self-reported stroke, TIA and possible vascular risk factors for stroke from a total of 4972 first-degree relatives and 738 spouses.

For Paper II (inclusion between March 1, 2001 and September 30, 2004): 932 IS patients and 396 control subjects.

For Paper III (inclusion between March 1, 2001 and August 31, 2007, participants included in Paper II were excluded in this study): 920 IS patients and 566 control subjects. In addition this multi-center study included participants from The Malmö Diet and Cancer Study (MDC) and The Sahlgrenska Academy Study on IS (SAHLSIS). Se section below.

For Paper IV (inclusion between March 1, 2001 and February 28, 2009): 2250 IS patients and 930 control. In addition this multi-center study included participants from MDC and SAHLSIS. Se section below.

Two other included study populations – MDC and SAHLSIS

In Papers III and IV, the two study populations MDC and SAHLSIS were included in addition to LSR.

MDC is a prospective cohort study that was initiated in the early 1990's (**Dahlberg 2011, Smith 2009**). A cohort of 28449 individuals from the local catchment area of former Malmö University Hospital (since 2009 included in Skåne University Hospital) were included between 1991 and 1996 for baseline examinations. First-ever IS patients were recruited from that cohort until 2006. Control subjects without stroke were selected from the same cohort and matched to the cases for sex and year of birth. Participant characteristics, including vascular risk factors for stroke (e.g. hypertension, diabetes mellitus and smoking), were collected at baseline.

SAHLSIS is a case-control study of IS that comprises selected geographic areas in Western Sweden due to inclusion of patients from four stroke units. First-ever stroke patients were recruited to SAHLSIS between 1998 and 2008 (**Jood 2005, Olsson 2011.1, Olsson 2011.2, Olsson 2011.3**). Control subjects consist of individuals without cardiovascular disease that were randomly recruited from the same geographic area as the patients. The recruitment was carried out through a population-based health survey. Control subjects were also included from the Swedish Population Register. The age of patients and control subjects ranged between 18 and 70 years at inclusion.

Variables

Stroke

Any person that was diagnosed with first-ever stroke by hospitalization or visiting the primary health care within the local catchment area of Skåne University Hospital's subdivision in Lund (the former Lund University Hospital) was defined as a stroke patient. For the multi-center studies (also including the MDC and SAHLSIS populations) the overall IS definitions were similar.

Stroke was diagnosed according to the WHO definition (**WHO 1988**). It is notable that subdural hematomas and TIAs are not defined as stroke.

Ischemic stroke and ischemic stroke subtypes

IS was diagnosed through computer tomography or magnetic resonance scan, or autopsy findings. In addition, a classification of IS subtypes was performed by using the nomenclature of TOAST (**Adams 1993**). Three distinguishable TOAST subtypes were thus determined, namely LVD, SVD and CE. These TOAST subtypes are more profoundly described in the Introduction section.

LSR included in February 28, 2009 a total of 2177 examined patients, of whom 162 patients were diagnosed with LVD, 479 with SVD and 685 with CE. Remaining non-missing patients were either subjects with cryptogenic stroke etiology, or defined with more than one subgroup or unclearly defined (denoted as cursory examination). SAHLSIS, but not MDC, also included TOAST-subtyped data (for SAHLSIS study data 111 LVD, 165 SVD and 151 CE patients were included).

Phenotypes

As mentioned in the Introduction section, stroke is associated with various vascular risk factors. The most essential of these risk factors are

- Age
- Hypertension
- Diabetes mellitus
- Heart diseases
- Smoking habits, especially current smoking
- Previous TIAs
- Heart disease

For the LSR population as well as the MDC and SAHLSIS populations, the variable *age* was defined at the date of first-time stroke onset for patients. For control subjects of LSR and SAHLSIS, age was defined at the time-point of inclusion.

Variable definitions of age and vascular stroke risk factors for MDC control subjects were however different because these subjects were included in a longitudinal cohort study where the patients suffered a stroke during the period (from 1991 to 2006) of surveillance. Thus, control subjects were individually matched to stroke patients according to age and sex, and the age of each control subject was consequently defined at the date of the matching stroke patient's stroke onset.

Hypertension was defined as either medical treatment for hypertension, or blood pressure 160/90 mm Hg or higher at the time of discharge or after at least 1 week of hospitalization (measurements among SAHLSIS cases were performed at 3 months follow-up) (**Jood 2005**).

Diabetes mellitus was defined as dietary or medical treatment for diabetes mellitus or measurements at discharge or at least 1 week of hospitalization as follows (measurements among SAHLSIS cases were performed at 3 months follow-up) (**Jood 2005**):

Blood glucose >6.1 mmol/l at two occasions

Plasma glucose >7.0 mmol/l at two occasions, or alternatively, >11.0 mmol/l at one occasion with concurrent symptoms suggestive of diabetes mellitus.

Smoking habits was defined through self-report. Three alternatives were given: Current smoking, previous smoking and never-smoking. A dichotomous variable for *current smoking* (used in Papers I-IV) was defined as "current smoking versus previous smoking and never-smoking".

Previous TIAs was defined as sudden onset of pareses, sensory loss, visual loss, speech disorder and similar, with duration shorter than stroke, usually minutes or hours and not more than 24 hours.

For LSR, *heart disease* was defined as diagnosis of angina pectoris, myocardial infarction, congestive heart failure, atrial fibrillation/flutter, medical treatment for heart disease or history of cardiac surgery. The *heart disease* variable was not included from MDC and SAHLSIS. Heart disease was included as an intermediary variable in Papers I and II.

For MDC, the phenotypic risk factors for stroke were defined at the inclusion of each participant (i.e. fully six years in average before stroke onset for the MDC stroke patient group). The subsequent time-lag may cause some loss of validity for these vascular stroke risk indicators.

Indicators for family history-based inheritance (LSR)

To the family history assessments (Paper I) we have defined the inheritance indicators in the subsequent way:

Patients and control subjects from LSR were considered as probands. By completing a questionnaire these probands have provided information about the occurrence of stroke (including TIAs), diabetes mellitus, hypertension and smoking habits among their relatives (i.e. the probands' grandparents, parents, siblings, spouses and children).

Only first-degree relatives were relevant for the family-history study in Paper I. A first-degree relative can either be a parent, a sibling (also a half-sibling) or a biologic child of the proband. Half-siblings were accepted as first-degree relatives. The information from the first-degree relatives were self-reported by the respective proband in the study or, if not possible, by his/her next of kin.

Table 1. Availability of outcome variables and explanatory variables (risk factors for stroke) among probands and first-degree relatives of probands in the Paper I family-history study.

	Available for probands	Available for first-degree relatives*) of proband
Outcome variable:		
– Ischemic strokes (IS)	Yes	No
– Intracerebral hemorrhages (ICH)	Yes	No
– Subarachnoid hemorrhages (SAH)	Yes	No
– All types of stroke	Yes	Yes (stroke + TIA)
Risk factors considered:		
– Previous TIA	Yes	Partially (stroke + TIA)
– Hypertension	Yes	Yes
– Diabetes mellitus	Yes	Yes
– Smoking	Yes	Yes, partially**)

*) First-degree relatives *and* spouses. Note, that all information is based on anamnesis from the proband or his/her next-of-kin.

**)The variable smoking includes non-smokers, previous smokers and current smokers for the probands at stroke onset, but only a general "yes/no"-reply for the first-degree relatives.

Table 1 gives an overview about the variables obtained from the probands, and from the questionnaire replies regarding the first-degree relatives and siblings of the probands. For natural reasons we were not able to collect information about stroke subtypes among first-degree relatives, all stroke types and additionally TIA were defined as one indivisible cluster of strokes. Another constraint may be the validity of these data, as all information is based on anamnesis from the proband.

Indicators for genotype-based inheritance

Assessments of the associations between IS and particular genetic markers were performed in Papers II, III and IV. These genetic markers, also denoted genotypes, were defined as single nucleotide polymorphisms (SNPs) within specific regions of particular selected chromosomes.

In Paper II five SNPs coding for the *PDE4D* gene within chromosome 5 were assessed. These five SNPs were also subjects for a very plain haplotype analysis with the intention to detect a possible linkage disequilibrium structure within the *PDE4D* gene (**Broad Institute 2008**). Paper II additionally included the analyses of three SNPs within *ALOX5AP* on chromosome 13 and one SNP within the *MHC2TA* region in chromosome 16.

In Paper III one particular SNP, SNP rs12188950 in the 5' end of chromosome 5, also denoted SNP45 and encoding for *PDE4D*, was analyzed thoroughly.

Paper IV comprises 25 SNPs related to different loci. The main purpose was to find a predictor of IS risk by using the information from previously reported genetic association with coronary heart disease in recent GWASs (**Erdmann 2009, Hamrefors 2011, Kathiresan 2009, Newton-Cheh 2009, Preuss 2010, Reilly 2010, Visel 2010, Samani 2007, Schunkert 2010**). The Coronary Artery Disease Genome-Wide Replication and Meta-Analysis Study (CARDIoGRAM) played a central part. We thus focused on compilation of risk scores having this previous findings regarding coronary heart disease risk in mind (**Ripatti 2010, Schunkert 2011**). Consequently, the following SNPs were considered when these risk scores were compiled:

rs646776	on chromosome	1p13,
rs17114036	on chromosome	1p32.2,
rs11206510	on chromosome	1p32.3,
rs17465637	on chromosome	1q41,
rs6725887	on chromosome	2q33.1,
rs9818870	on chromosome	3q22.3,
rs17609940	on chromosome	6p21.31,
rs12190287	on chromosome	6q23.2,
rs12526453	on chromosome	6p24.1,
rs3798220	on chromosome	6q25.3,

rs11556924	on chromosome	7q32.2,
rs4977574	on chromosome	9p21.3,
rs579459	on chromosome	9q34.2,
rs1746048	on chromosome	10q11.21,
rs12413409	on chromosome	10q24.32,
rs964184	on chromosome	11q23.3,
rs653178	on chromosome	12q24,
rs4773144	on chromosome	13q34,
rs2895811	on chromosome	14q32.2,
rs3825807	on chromosome	15q25.1,
rs216172	on chromosome	17p13.3,
rs12936587	on chromosome	17p11.2,
rs46522	on chromosome	17q21.32,
rs1122608	on chromosome	19p13.2, and
rs9982601	on chromosome	21q22.1.

Natural logarithms of the odds ratios (ORs) from the previous GWASs (**Schunkert 2011**) were used as weights when compiling the risk scores. A plain, alternative compilation method that simply added the number of risk alleles representing the 25 included SNPs was also utilized.

Statistical methods

Three kinds of statistical analysis, each referring to a specific purpose, were performed: 1.) To examine if the data was appropriate for the desired analysis. 2.) To perform the desired analysis to find possible significant associations that are consistent with the hypothesis formulated. 3.) To confirm that the desired analyses were interpreted correctly.

The first category of statistical analysis comprised univariate statistics of mostly background data, such as Mann-Whitney's U-test of two independent samples, Wilcoxon's test of paired samples or Fisher's exact test of two-by-two tables. The purpose was to briefly examine cases and control subjects regarding possible differences in age, sex and some dichotomous vascular risk factor indicators (e.g. hypertension, diabetes mellitus or smoking/never-smoking).

The second category included methods of univariate and multivariable analysis, mainly in order to confirm possible findings regarding association between stroke risk (i.e. the outcome variable) and an assumed hereditary risk factor for stroke (i.e. family history of stroke/TIA or genotyped genetic markers). ORs from computations of two-by-two tables as well as simple logistic regression and multiple logistic regression

analysis were utilized. Multiple logistic regression approaches were mostly controlling for vascular risk factors (e.g. hypertension, diabetes mellitus, smoking).

In Paper II we practiced a method to compute relative excess risk due to interaction (RERI). The purpose was to test for possible difference between two ORs, discriminated by two groups: hypertensive versus not hypertensive participants (**Hosmer 1992**). We also performed correlation analyses using a standard program for analysis of haplotype structures to find a possible cluster of SNPs forming a haplotype block (**Broad institute 2008**).

The third category comprised statistical methods intended to evaluate the quality and credibility of the assessed data. In Papers II-IV we performed Hardy-Weinberg equilibrium (HWE) tests to examine if the allelic frequencies deviated from a state of being constant from generation to generation. Such a deviation might have affected the interpretation of our results. See also the *Further Statistical Considerations* section below. In Papers II and III we carried out heterogeneity tests for possible ambiguousness regarding the findings in the meta-analyses.

Meta-analysis

In Papers II and III we performed random effects DerSimonian-Laird meta-analyses of selected previous publications (**DerSimonian 1986**). The main purpose was to find a precise overall OR. The meta-analyses (and the forest plots linked to these, see for example **Figure 3** under the Findings section) were also aimed to visualize possible geographical, ethnical and gender-related differences regarding IS susceptibility of SNP rs12188950 in the *PDE4D* gene.

Findings

Findings in Paper I

We noticed a significant augmented risk for stroke among probands with a family history of stroke or TIA. Conversely, first-degree relatives to a group of 582 patients showed a family-history prevalence of stroke or TIA of 12.3%. For the first-degree relatives of 236 control subjects this rate was 7.5%. Also, OR=1.74; 95% CI: 1.36-2.22 when considering stroke/TIA among first-degree relatives as a risk factor for stroke among the probands themselves. The number of first-degree relatives attached to these patient and control subject probands was 2439 and 1241, respectively.

When considering 2188 first-degree relatives to patients with stroke subtype cerebral infarctions only, this association remained significant (OR=1.76; 95% CI: 1.38-2.26; estimates are adjusted for age and gender of proband).

Another notification is the detection of a possible association between stroke risk and family-history of hypertension (OR=1.33; 95% CI: 1.10-1.60). Heart disease, diabetes mellitus and current/previous smoking among first-degree relatives were also tested against stroke risk among probands, but were found non-significant.

The association between stroke among the probands and previous stroke or TIA onsets among their respective spouses (414 stroke patients and 186 control subjects with non-missing values were examined) was non-significant (OR=1.20; 95% CI: 0.70-2.06; p-value=0.497). This notation strengthens the assumption that stroke risk is not generated by social impact within families.

It is noticeable that prevalence of stroke or TIA among mothers might increase the stroke risk for the proband (OR=2.04; 95% CI: 1.30-3.20; p-value=0.002). The observed effect size estimate of this transmission from father to proband is weaker and non-significant (OR=1.47; 95% CI: 0.95-2.29; p-value=0.085). When considering probands below 75 years of age, this distinction between mother-proband and father-proband may be less apparent (with ORs of 2.08 and 1.91, respectively – and p-values of 0.015 and 0.027).

Findings in Paper II

We examined nine SNPs in three genetic regions (*PDE4D* – 5 SNPs; *ALOX5AP* – 3 SNPs; and *MHC2TA* – 1 SNP) and we found two of these to be significantly associated with IS risk: SNP rs12188950 (also denoted SNP45) in the *PDE4D* gene, with OR=0.72; 95% CI: 0.58-0.91; p-value=0.0055, and rs3887175 (or SNP39) also in the *PDE4D* gene, with OR=0.81; 95% CI: 0.65-1.00; p-value=0.0460. However, when taking Bonferroni correction for mass significance (considering multiple comparison of nine distinguishable tests) into account, only SNP rs12188950 remained significant (with a p-value of 0.050). For SNP rs12188950 we examined 929 non-missing patients and 395 non-missing control subjects.

When performing similar analyses, considering patients and control subjects with hypertension, we were able to detect an association with IS that might be stronger: 575 patients and 133 control subjects showed OR=0.52; 95% CI: 0.37-0.73, p-value=0.0001 for allelic variants in rs12188950, while rs3887175 provided OR=0.57; 95% CI: 0.41-0.79; p-value=0.0007 (565 patients and 131 control subjects with hypertension were included when assessing rs3887175).

Furthermore, one additional SNP, rs17222814 (also denoted SG13S25) in the *ALOX5AP* gene, was significantly associated with IS (OR=1.82; 95% CI: 1.21-2.74; p-value=0.0039) when assessing 353 non-hypertensive patients and 261 non-hypertensive control subjects.

We thus found that variants in minor alleles (T) of SNP rs12188950 might have a protecting effect against IS, especially among subjects with hypertension. Within this hypertension subgroup the minor alleles (T) of rs3887175 might also have an inhibitory effect on IS. Additionally, our findings from this study suggest that variants of minor alleles (A) of rs17222814 might have a synergic impact on IS. However, our dividing into two subgroups of participants of the study (hypertensives and non-hypertensives) can lead to non-significance regarding the latter SNP test because it may be necessary to consider Bonferroni correction considering 18 comparative tests.

An examination of the association between the five SNPs within the *PDE4D* gene showed a linkage disequilibrium (R^2=0.63) between rs12188950 and rs3887175. A haplotype block including four SNPs (rs12188950, rs3887175, rs26956 and rs27653, but not rs2910829) was detected.

The over-all OR obtained from the random effects (DerSimonian-Laird) meta-analysis of 13 studies including the apparent study (OR=0.95; CI: 0.86-1.05; p-value from test for homogeneity=0.0419), did not confirm the suggested protective effect against IS for the rs12188950 SNP according to our study, presented in Paper II.

Findings in Paper III

We examined one SNP, rs12188950 within the *PDE4D* gene. By performing an assessment where data from not only LSR, but also MDC and SAHLSIS was included (in all 2599 IS patients and 2093 control subjects), we intended to confirm the previous findings from 929 IS patients and 395 control subjects with non-missing values regarding rs12188950 polymorphism data in Paper II. We excluded those previously assessed subjects when performing the analyses in this apparent study.

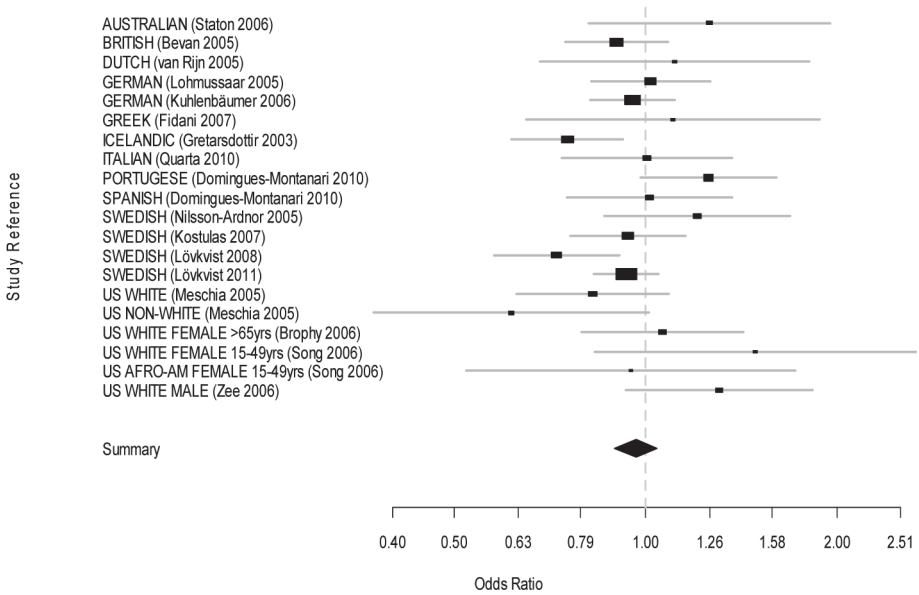
Our findings from this large-sized multi-center study differed from the previous study presented in Paper II: A univariate analysis provided OR=0.93; 95% CI: 0.83-1.05; p-value=0.251, to be compared with the results shown from the similar assessments presented in Paper II, OR=0.72; 95% CI: 0.58-0.91; p-value=0.0055. When assessing patients and controls with hypertension the difference became even more noteworthy: OR=0.95; 95% CI: 0.81-1.11; p-value=0.521 (1727 patients, 1056 controls), to be compared with OR=0.52; 95% CI: 0.37-0.73, p-value=0.0001 as presented in Paper II.

When considering a subgroup of participants below 55 years of age we found a scarcely significant tendency towards a protective effect of the rs12188950 minor allele on IS (OR=0.73; 95% CI: 0.53-1.00; p-value=0.057, CIs approximately estimated by Woolf´s approach, p-value computed by Fisher's exact test) (**Woolf 1955**).

Paper III also comprised a weighted multiple logistic regression approach using the three dichotomous vascular risk factor indicators hypertension, diabetes mellitus and current smoking (versus never-smoking and previous smoking) as regression covariates. The weight variable was compiled in order to attain a standardization of the three included populations (LSR, MDC and SAHLSIS) according to their age and sex distribution. Weighted multiple logistic regression did not change the above findings dramatically: for the three populations joined together we obtained OR=0.91; 95% CI: 0.80-1.04; p-value=0.159.

We updated the previous random-effects meta-analysis from Paper II and changed its structure to achieve more interpretable results. From the 20 populations found in the 17 included studies (including results from Papers II and III) we now obtained a persisting non-significant overall-estimate of OR=0.96; 95% CI: 0.89-1.04; a test for heterogeneity providing a p-value=0.0416. A total of 10500 IS patients and 10102 control subjects were included in the meta-analysis. Populations labeled due to nationality and (for a few populations:) sex and race are shown in **Figure 3**.

Figure 3. Random effects (DerSimonian-Laird) meta-analysis of 17 published studies (including Papers II and III), comprising 20 distinguishable study populations to assess the association between the minor allele (T) of rs12188950 and IS risk.



Overall OR: 0.96 (95% CI: 0.89-1.04), test for homogeneity (chi-square distributed with 19 degrees of freedom): χ^2 =30.9 (p-value=0.0416).

Findings in Paper IV

We considered previous findings from the CARDIoGRAM study and performed a candidate gene study to examine if genetic markers susceptible for coronary heart disease also might affect IS risk. We thus assessed 25 SNPs and found variations in one SNP to be significant: rs4977574 on chromosome 9p21.3, slightly linked with the *ANRIL* gene region (**NCBI 2012.1**). The risk allele (G) showed an OR=1.12; 95% CI: 1.04-1.20; p-value=0.002 when examining the entire multi-center data joined together (pooled data included 3916 IS patients and 2419 control subjects that comprised non-missing data for this particular SNP). This single genetic marker remained significant even after Bonferroni-adjustment for 25 comparisons.

When assessing risk scores compiled by plain addition of risk alleles from the 25 included SNPs we did not find any significant association with IS at all. We examined all participants in the study as well as participants younger than 55 years of age. We also assessed men and women separately.

When analyzing LSR, MDC and SAHLSIS separately we found a tendency towards significant association for the MDC data (p-value=0.049). This association appeared

to be stronger when the weighted approach for compiling risk scores was utilized (p=0.008; this result is still significant when performing a Bonferroni-adjustment for up to six comparisons). This is shown in **Table 2**. As a complement to these simple logistic regression analyses, we also performed multiple logistic regression assessments controlling for the vascular risk factors diabetes mellitus, hypertension and current smoking.

Table 2. Simple logistic regression of possible association between allelic risk score (unweighted and weighted, respectively) and IS risk.

			Unweighted RS		Weighted RS	
	$N_{controls}$	N_{cases}	Coeff.	P-value	Coeff.	P-value
IS, populations:						
LSR	930	2208	-0.016	0.224	-0.051	0.664
MDC	834	880	0.032	**0.049**	0.395	**0.008**
SAHLSIS	661	814	-0.004	0.814	0.087	0.576
Total	2425	3902	-0.002	0.807	0.086	0.267
IS, subgroups:						
<55 years of age	350	470	0.017	0.486	0.180	0.405
Women	1049	1745	-0.004	0.746	0.101	0.388
Men	1376	2157	-0.001	0.959	0.071	0.491

In Paper IV, we also focused on three TOAST subtypes providing data from LSR and SAHLSIS, but not from MDC. SNP rs4977574 on chromosome 9p21.3 was significant for patients with large-vessel disease (LVD), providing an OR=1.36; 95% CI: 1.13-1.64; p-value=0.001. No other SNP showed any significance for neither the LVD subgroup nor the SVD or CE group. Assessments using unweighted and weighted risk scores did not confer any association between coronary artery disease-related genetic traits and IS risk. A remarkable finding is that CE risk showed a negative association to the unweighted risk score with p-value 0.046. When applied on the weighted risk score, and when performing multiple logistic regression controlling for diabetes mellitus, hypertension and current smoking, this association was not significant.

Summary of findings

It is obvious that stroke, at least partly, can be explained by hereditary components. We have shown that stroke might be genetically inherited from first-degree relatives (parents, siblings) to proband. We could not find a corresponding association between spouses and probands. We have also found indications suggesting that some SNPs within some particular genes affect IS risk. However, our significant findings within that subject of study are not necessarily unambiguous. Our dissimilar findings from the studies presented in Papers II and III suggest that more studies need to be performed

to discover etiologic structures, different susceptibility of particular genetic markers within different demographic subgroups, and other mechanisms that might affect genetic impact on stroke.

Discussion

The findings of family history of stroke

From the studies on familial stroke inheritance in Paper I, we found that there might be a transmission of stroke risk (and hypertension) within families. This familial association seems to be particularly evident when considering transmission of stroke from mother to offspring. However, such a transmission between spouses could not be detected in our study. We have some remarks about these findings.

Transmission to offspring

We found a tendency towards a stronger inheritance of stroke risk from mother to offspring compared to corresponding inheritance from father to offspring. Such a transmission has been reported previously in e.g. an Austrian study that suggested that parental, especially maternal history of stroke is associated with prevalence of hypertension among the related probands (**Tentschert 2003**).

Maternal inheritance can be transmitted through either autosomal chromosomes, the X chromosome, or through monogenic mitochondrial mutations. Chromosomes in the nucleus of cells are carriers of vast quantities of genetic information that makes them more important (i.e. more relevant for exploration for finding more possibly relevant information) than monogenetic mitochondrial mutations. All analyses on candidate genes in this thesis are performed on autosomal chromosomes.

Nevertheless, possible maternal inheritance of stroke risk may in some few cases be explained by a certain mitochondrial mutation that is transmitted from mother to offspring (**Schaefer 2004**). A specific mutation of mitochondrial DNA, A3243G, is known to be associated with e.g. mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes (MELAS) (**Dichgans 2007, Mariotti 1995, Verny 2008**). MELAS may be manifested as occipital brain infarcts which are primarily diagnosed among younger patients (18-45 years of age) (**Martínez-Fernández 2001**). It has to be mentioned that stroke and stroke-like episodes are two quite different diagnoses that should not be confused. Other known manifestations of monogenic mitochondrial disorders are diabetes mellitus and cardiomyopathy (**Mariotti 1995, Murphy 2008**), that might be involved as intermediary risk factors for stroke.

Transmission between spouses

Any association between spouses regarding stroke was not found significant in Paper I. This is expected from a strictly genetic point of view, simply because any form of biological inheritance is absent. Nevertheless, transmission of stroke within a family may also be caused by non-genetic risk factors that are shared within that family. Examples are smoking habits, alcohol consumption, dietary habits, physical activity, type of residence (own house or apartment, located in city environment or country-side), exposure to local environmental pollution etc (**Björk 2008, Giskes 2005, Oudin 2010, You 1999**).

Additional aspects

The analyses of familial aggregation of stroke have limitations. Only 66 percent of the 925 patients returned a completed questionnaire about the self-reported family history. However, the responses regarding mothers, fathers and spouses of all included, eligible probands was good, with a coverage of 95%, 93% and 85%, respectively. One should also be aware of that collected data, based upon anamnesis from patients about stroke and its risk factors among the members of these patients' families, might have a weaker validity. The results shown in Paper I should therefore be interpreted carefully.

The *PDE4D* gene area findings

The two studies on the *PDE4D* genetic region (Papers II and III) provided ambiguous results regarding SNP rs12188950. The three populations included in the latter study (Paper III), LSR, MDC and SAHLSIS, were to some extent dissimilar regarding the participants' mean age, the design of the case-control study and the definitions of the vascular risk factors for stroke (e.g. the time for registering information about hypertension, diabetes mellitus etc). This might partially explain this ambiguousness between the results in the two papers.

However, the estimated OR for variants of rs12188950 on IS did also differ when comparing the LSR subpopulation results in Paper III with the corresponding estimates of the LSR subpopulation in Paper II. This was surprising because the underlying population is in practice identical for the two samples: LSR data for Paper II was collected between March 2001 and September 2004. LSR data for Paper III was correspondingly collected until August 2007. However, participants in the study linked with Paper II were not included in the study reported in Paper III.

An inhibitory effect against IS of the minor allele of rs12188950 was significant when presented in Paper II, but not in Paper III. For hypertensive participants this result was even more ambiguous: The results of our assessments in Paper II showed a strong protective effect of variants in rs12188950 against IS while the corresponding results in Paper III did not.

A change in the hypertension definitions (to systolic/diastolic blood-pressure: 140/90 mmHg) in the year of 2003 cannot explain that discrepancy, because we consequently use the former definition (160/90 mmHg) in both studies. Nevertheless, 56.4% of the control subjects in the Paper III-sample were hypertensive, but only 33.6%, or 133 individuals, from the Paper II-sample. With an allele frequency of 21% (57/266 alleles) it is a subject of discussion whether this is an outcome occurring by coincidence or not.

A brief biological approach

Cyclic adenosine monophosphate (cAMP) is a second messenger, i.e. it can be regarded as a specific molecule that relays signals from receptors of the cell surface to target molecules in the cell. This will result in some kind of change in the activity of that cell, e.g. a hormone activity (**NIH 2011**). *PDE4D* is an important regulator of intracellular levels of cAMP. The original publication of the Icelandic GWAS from 2003 reported that the gene on chromosome 5q12, encoding *PDE4D*, may be involved in the pathogenesis of stroke through such a regulating effect on cAMP (**Gretarsdottir 2003**). It has been reported that this may affect e.g. vascular smooth muscle cells, endothelial cells, T-lymphocytes and macrophages, and thus have an impact on the pathogenesis of atherosclerosis (**Baillie 2001, Gretarsdottir 2003, Jin 2002, Landells 2001, Liu 1999, Liu 2000**). A more recent study have confirmed the involvement of *PDE4D* through cAMP in various heart disorders including cardiac arrhythmias in mice (**Lehnart 2005**).

What can we learn from conflicting findings in the meta-analyses?

Statistical analyses of observational studies, aimed to show possible association between a selected genetic marker and the risk of having some manifestation of a particular disease, in this case IS, can sometimes reveal results that may be vague or even conflicting.

Considerable effort has been put on SNP rs12188950 within the *PDE4D* gene, and previous studies have documented ambiguous results: In Chinese, Taiwanese and Japanese populations this SNP was found to be monomorphic and thus not useful for analysis (e.g. **Lin 2007, Matsushita 2009, Sun 2009, Xu 2010**). In the Icelandic original study from 2003, the rs12188950 SNP showed an inhibitory effect against IS (**Gretarsdottir 2003**). One British study, used in the two meta-analyses, showed a vague, non-significant tendency towards the same direction as the Icelandic study (**Bevan 2005**). But studies on e.g. Central and South European populations did not explicitly confirm these findings (e.g. **Löhmussaar 2005, Kuhlenbäumer 2006, Fidani 2007, Quarta 2010, Domingues-Montanari 2010**). Studies on U.S. populations (African-Americans as well as participants with European ancestors) did not show any clear tendency neither regarding possible association between rs12188950 and IS risk (**Brophy 2006, Meschia 2005, Song 2006, Zee 2006**).

Ethnical and racial differences may, when occurring, reflect cultural and socioeconomic factors as well as genetically inherited physical traits in some populations. This subject should be handled respectfully (**Nature Genetics Editorial 2000**).

There are many facts that might explain the ambiguousness regarding the meta-analysis results. Three such facts to consider are:

1. Conflicting methods of data collection

2. Conflicting definitions of stroke

3. Demographic differences between populations due to participating subjects' nationality, ethnicity/race, sex, age etc.

1. Conflicting methods of data collection: Most of the studies included in the two meta-analyses in Papers II and III were population-based case-control studies where the patients were recruited consecutively, and the control subjects were drawn by some random selection procedure from any administrative register. In some studies a nested case-control design was created, i.e. each patient was assigned one or more control subjects with similar demographic attributes, such as age and sex (e.g. **Brophy 2006, Nilsson-Ardnor 2005, Zee 2006**). One study was designed as a controlled case-control trial using a randomization procedure when selecting cases and control subjects (**Zee 2006**). Another study included only elderly women with hip fractures (**Brophy 2006**), a third study focused on women aged 15-49 years (**Song 2006**).

2. Conflicting definitions of stroke and/or risk factors for stroke: In most of the included studies the definition of IS was explicitly described to be in accordance with the World Health Organization criteria.

3. Demographic differences: The most important demographic variable, also an important risk factor for stroke, is age. In three articles, participating in the meta-analysis presented in Paper III, we found the control subjects to be approximately five years younger than the IS patients (**Meschia 2005, Kuhlenbäumer 2006, Kostulas 2007**). In another article the patients were about five years younger than the control subjects (**Domingues-Montanari 2010**). This may affect the estimated crude ORs that we used in our meta-analyses.

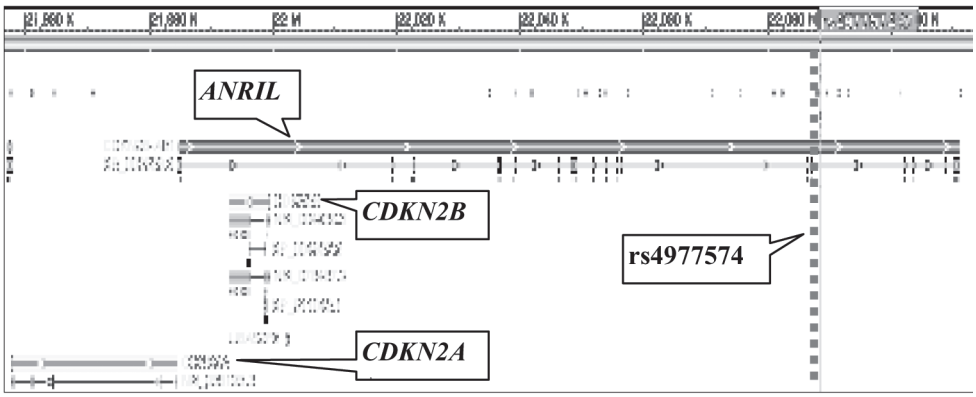
Another aspect regarding demographic differences is the difference in e.g. mean age between two studies. Such between-groups variations are detectable by the use of the chi-square test for heterogeneity.

The 9p21.3 chromosomal region findings

When performing the assessments for the study presented in Paper IV, we found SNP rs4977574 within the 9p21.3 chromosomal region and located upstreams from the *CDKN2A*, *CDKN2B* and *ANRIL* regions (**Figure 4**) to be significantly associated with

IS. This was particularly obvious when considering the LVD subtype according to TOAST. These findings were not unexpected.

Figure 4. Sequential section of chromosome 9p21 (range: 21960 Kbp-22120 Kbp), showing the position of SNP rs4977574 (at 22088 Kbp) within the *ANRIL* region (also denoted the *CDKN2B-AS1* region, ranged 21984 Kbp-22111 Kbp), and upstreams from the *CDKN2A* and *CDKN2B* regions. Kbp=kilo base pairs (**Kristoffersson 2003, Strachan 2004**).



The impact of 9p21-located SNPs on coronary heart diseases and IS has been reported previously. A GWAS from 2007 reported that data included from the Wellcome Trust Case Control Consortium (WTCCC) study and two other studies, identified allelic variants (e.g. SNP rs1333049, fairly correlated with rs4977574 used in Paper IV) within chromosome 9p21 as associated with coronary artery disease (**Samani 2007, WTCCC 2007**). An Icelandic study found especially SNP rs1333040 and two other SNPs on chromosome 9p21 to be associated with the risk of myocardial infarction (**Helgadottir 2007**). Also, a report from 2009 suggested that variants of SNP rs10757278 on chromosome 9p21 were associated with coronary heart disease, but could not find such an impact on stroke generally (**Lemmens 2009**). Another report found a number of selected SNPs (e.g. rs1537378, but not rs1333049) within the 9p21 chromosome to be significantly associated with atherosclerosis (**Gschwendtner 2009**).

The two coding sequences of genes for two cyclin-dependent kinase inhibitors, i.e. the *CDKN2A* and he *CDKN2B* regions within 9p21, play an important role in the regulation of the cell cycle and may be involved in the pathogenesis of atherosclerosis (**Lowe 2003, Hannon 1994, Kalinina 2004, Samani 2007**). The *ANRIL* region may also, as a large antisense non-coding RNA (ncRNA) be associated with atherosclerosis through its role in epigenetic mechanisms (**NCBI 2012.2, Costa 2008, Broadbent 2008, Lemmens 2009**).

Non-significant results

Genetic representation of two genetic regions could not be reported to be associated with IS. Three SNPs within the *ALOX5AP* gene on chromosome 13 and one SNP within the *MHC2TA* gene on chromosome 16 were not found significant in Paper II when tested against IS risk. Also, the multi-locus risk scores that were compiled and analyzed in Paper IV, did not confer an association between genetic markers, susceptible for coronary artery disease, and IS risk.

ALOX5AP on chromosome 16

A polymorphism within the *ALOX5AP* gene, encoding 5-lipoxygenase activating protein (FLAP), has been reported to be associated with increased risk of myocardial infarction in Iceland (**Helgadottir 2004**). Other studies report a possible association with myocardial infarction as well as ischemic and hemorrhagic stroke risk (e.g. **Dong Hwan Kim 2011, Dichgans 2007, Linsel-Nitschke 2008, Lõhmussaar 2005, Meschia 2005**).

FLAP is a protein that is involved in the production of leukotrienes, which play a part regarding contraction of smooth muscles lining the trachea (**Nelson 2008**). Consequently, FLAP is involved in some types of inflammatory processes including asthma. FLAP has also in recent years been linked to risk of e.g. myocardial infarction, restenosis and stroke (**Evans 2008**).

MHC2TA on chromosome 13

The major histocompatibility complex class II transactivator (*MHC2TA*) on chromosome 16, was not reported to be associated with stroke when we planned and performed the analyses for Paper II. *MHC2TA* has previously been reported to be associated with diseases of inflammatory etiology, including myocardial infarction (**Swanberg 2005, Lindholm 2006**).

To our knowledge, the possible associations between IS and polymorphisms in the *MHC2TA* gene have not been examined since our publication of Paper II. Nevertheless, it should be noted that a recently published Spanish study has assessed possible impact of SNP rs3087456 (i.e. the same polymorphism as we tested in Paper II) within the *MHC2TA* gene on cardiovascular diseases among patients with rheumatoid arthritis (**Garcia-Bermudez 2012**): No significant association was found in that study.

Comments on the multi locus risk score study

We used a set of effect-size estimates (ORs) from GWASs within CARDIoGRAM (**Schunkert 2011**), and furthermore a reasonably large successfully genotyped sample from a three-center study involving 3902 patients and 2425 control subjects from the LRS, MDC and SAHLSIS populations. Despite this starting point, we could not confirm our hypothesis that SNPs susceptible for coronary artery disease may also influence

IS risk, by using risk scores based on 25 SNPs from various chromosomal loci (Paper IV).

A recent publication, based on a large-sized multi-center study including over 18000 participants and involving assessments of 11 SNPs, could not identify any ischemic-stroke associated risk score neither (**Cheng 2012**). This negative "finding" also included results from analyses of separate TOAST subtypes, including LVD.

Concluding remarks

Genetic influence on stroke risk is a complex issue with no simple answer. Our effort with the four papers that formed this thesis was to give a contribution for making this complex issue a little more more comprehensible. The intention was to (1) formulate hypotheses; (2) find and evaluate methods for analyses in order to accept or reject our hypothesis impartially; and (3) present and discuss our findings, with focus on improvements in further assessments.

Some notable significant associations regarding genetics and stroke, but also negative or even ambiguous findings, were detected. The ambition was to reveal and discuss all these findings and try to give suggestions about how to go further with the analyses and find more optimal solutions in the future.

Further statistical considerations

Allelic association analysis – two approaches

Computation of ORs in allelic association analysis can be performed with either an allele-based (1 unit=*1* allele) or an individual-based (1 unit =1 individual=*2* alleles) approach. Both of these approaches were used in this thesis.

Allele-based compilation of OR

The most basic method for computation of OR with relevance for allelic association analysis can be described as follows: we consider a sample of N_p patients and N_c control subjects, represented by M_p and M_c alleles respectively (**Table 3**). Consequently, $M_p = 2N_p$ and $M_c = 2N_c$ because each individual is represented by two alleles. We also denote minor alleles with ***v*** and major alleles with ***V*** for further use in the sections below.

Table 3. Notations for cell values in a two-by-two table.

	Exposed for risk (e.g. minor alleles **v**)	Not exposed for risk (e.g. major alleles **V**)	Total
Patients	A	B	M_p
Control subjects	C	D	M_c
Total	M_v	M_V	M

The odds ratio computation is simple:

$$OR = (A/B)/(C/D)$$

We can easily compute the approximate standard error of the logarithmized OR (**Woolf 1955**):

$$S.E.(\ln(OR)) = \sqrt{(1/A + 1/B + 1/C + 1/D)}$$

Consequently, the 95% confidence limits are computed as

$OR_{\text{lower 95\% CI limit}} = e^{(\ln(OR)-1.96*S.E.(\ln(OR))}$, and
$OR_{\text{upper 95\% CI limit}} = e^{(\ln(OR)+1.96*S.E.(\ln(OR))}$, respectively (**Morris 1988**).

A great advantage with the allele-based approach is its convenience. It is easy to use, and the results from the analyses are also easy to interpret as long as we do not consider possible dependence between the two alleles within each individual.

Individual-based compilation of OR

When performing individual-based compilation of the OR estimate, we treat the genotypic information as traits of an individual, i.e. as phenotypic information. Thus, an individual-based genotypic covariate may be defined as:

0 if the individual is a major allele homozygote (i.e. represented by two major alleles, that may be denoted **VV**),

1 if the individual is a heterozygote (i.e. represented by major allele and one minor allele, that may be denoted **Vv**, which is equivalent to **vV**), and

2 if the individual is a minor allele homozygote (i.e. represented by two minor alleles, that may be denoted **vv**).

The genotypic covariate will thus indicate the number of minor alleles for each individual it represents.

An advantage with the individual-based approach is its usefulness. The genotypic covariate can easily be included as an independent variable in e.g. a logistic regression model. Multiple logistic regression modeling controlling for intermediary or otherwise confounding phenotypic covariates (e.g. age, sex, BMI, indicators hypertension, diabetes mellitus etc) is thereby available.

However, there is also an interpretation problem linked with the individual-based approach: the thus achieved OR estimates will not always be consistent with the OR estimates that would be obtained by using the allele-based compilation approach, since the paired alleles within each individual might be mutually dependent. Such dependence between alleles within each individual can be referred to as Hardy-Weinberg disequilibrium (HWD).

Hardy-Weinberg equilibrium and disequilibrium

HWE expresses a state of independence between paired alleles within each individual. Such a state of independence will implicate an expected distribution of homozygotes and heterozygotes as described in **Table 4.** When the distribution of homozygotes and

heterozygotes confirm exactly to HWE (which in reality is an almost hypothetical situation), we may say that it is in a state of *strict* Hardy-Weinberg equilibrium (sHWE) (**Wellek 2004**).

Table 4. The distribution of major allele homozygotes, heterozygotes and minor allele homozygotes when data is in sHWE.

Genotype:	0 (***VV***)	1 (***Vv***)	2 (***vv***)	Total
Patients	$(1- p_p)^2$	$2*p_p*(1- p_p)$	p_p^2	1
Control subjects	$(1- p_c)^2$	$2*p_c*(1- p_c)$	p_c^2	1

Footnote: p_p=the minor allele frequency for patients, i.e. $p_p=A/ M_p$ when using the notations in **Table 3** above. Equivalently, for control subjects $p_c=A/ M_c$.

Possible departure from HWE can be detected by a simple chi-square test. If $N_{(VV)c}$=the number of major allele homozygotes; $N_{(Vv)c}$=the number of heterozygotes; and $N_{(vv)c}$=the number of minor allele homozygotes, the test function may be written as

$$\chi^2_{\text{(2 degrees of freedom)}} = (N_{(VV)c} - E(N_{(VV)c}))^2/E(N_{(VV)c}) + (N_{(Vv)c} - E(N_{(Vv)c}))^2/E(N_{(Vv)c}) + (N_{(vv)c} - E(N_{(vv)c}))^2/E(N_{(vv)c})$$

where

$E(N_{(VV)c}) = N_c*(1- p_c)^2$ = the expected number of major homozygotes when HWE is present,

$E(N_{(Vv)c}) = N_c*2* p_c *(1- p_c)$ = the expected number of heterozygotes when HWE is present,

and

$E(N_{(vv)c}) = N_c*p_c^2$ = the expected number of minor homozygotes when HWE is present.

The formula above is intended for the control subjects group, which is normally recommended for case-control data because the case (patient) group's breeding history might have been affected by the outcome we want to examine.

When HWD is apparent we will obtain different effect sizes (or ORs) depending on if we have used an allele-based or an individual-based approach for our computations. In such situations the allele-based approach may possibly not be an adequate tool for allelic association analysis. It has therefore been suggested that Armitage's trend test, applied on an individual-based approach using the 0/1/2-encoded genotypic covariate introduced above, should be considered instead (**Sasieni 1997**).

But, OR estimates based on allele counts will also be compromised when extending the statistical model to include influence of additional phenotypic variables. Logistic regression is an alternative to Armitage's trend test when assessing case-control data (**Balding 2006**). Logistic regression has an interesting advantage over Armitage's trend test as a method for assessing additive genotypic models encoded numerically with 0/1/2: The method allows for including auxiliary phenotypes as either confounding or intermediary covariates, or as non-linearly interacting factors (**Balding 2006, Lunetta 2008**).

A deviation from HWE is often used as a marker for genotyping error. However, such equilibrium departure may have several other causes including natural selection, genetic drift, population subdivision or random variation (**LaChance 2009**). Not only systematic departures from HWE, but also more trivial and statistically non-significant deviations from sHWE might affect the OR estimates when performing allelic association analysis.

A practical example from LSR (Paper II)

Let us reconsider the findings regarding possible allelic effect of variants in SNP rs12188950 in the *PDE4D* region on IS risk. We have a sample including 929 IS patients and 395 control subjects after exclusion of non-valid data. Subsequently, 245 alleles of 1858 in the patients group and 137 alleles of 790 in the control subjects group are minor alleles (**Table 5**). Minor and major alleles are denoted T and C, respectively (T for nitrogenous DNA-molecule base thymine, C for cytosine). **Table 5** also shows the number of individuals that are major allele homozygotes (encoded 0; allele-pair CC), heterozygotes (1; CT) and minor allele homozygotes (2; TT).

From the original analyses of these data in Paper II, using the allelic-based approach, we obtained an OR=0.724; 95% CI: 0.576-0.909; p-value=0.0054. When performing a simple logistic regression using the 0/1/2-encoded genotype as a numeric covariate we obtained an OR=0.728 (95% CI: 0.581-0.913; p-value=0.0061).

Table 5. The distribution of major allele homozygotes, heterozygotes and minor allele homozygotes. Data from Paper II, concerning SNP rs12188950.

	Allele distribution			Homozygote and heterozygote distribution			
	C	T	Number of alleles	0 (CC)	1 (CT)	2 (TT)	Number of individuals
Patients	1613	245	1858	702	209	18	929
Control subjects	653	137	790	271	111	13	395

A simulation study using bootstrapping

A further analysis of the SNP rs12188950 data in Paper II was carried out:

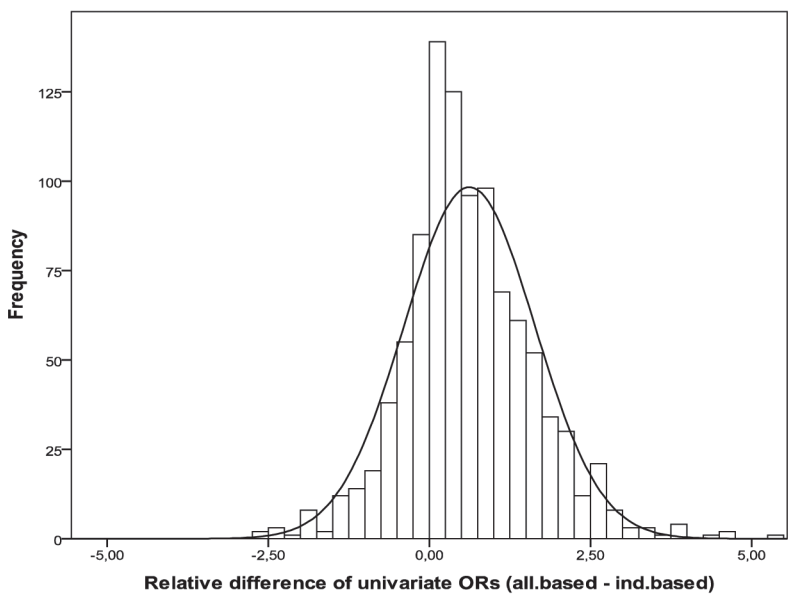
By using resampling without replacement, a method often referred to as bootstrapping, we created 1000 simulated samples with sample sizes similar to the "original sample" (929 IS patients and 395 control subjects) (**Efron & Tibshirani 1994**). The resampling procedure was performed by using a random process. In each of these bootstrapped samples we then compared the ORs obtained by simple logistic regression on individual-based data with the ORs obtained by the allelic-based approach using allele-counts from two-by-two tables. We were focusing on the relative difference between the two methods, expressed as

$$\text{rel.diff.} = 100\% * (OR_{(\text{logistic regression approach})} - OR_{(\text{allele-based approach})}) / OR_{(\text{allele-based approach})}) \ .$$

The result is illustrated in **Figure 5**. We found two samples with a relative difference of 5% or more, while 47 bootstrapped samples showed such a difference of 2.5% or more. From these simulations we conclude that coincidentally occurring departures from sHWE may result in conflicting results between estimated ORs obtained from allelic-based compilation and estimated ORs obtained from logistic regression analysis using a 0/1/2-encoded genotypic covariate in a simple logistic regression model.

For the original sample used in Paper II we obtained $OR_{(\text{allele-based approach})}=0.7240$ (when truncating to four decimals) and $OR_{(\text{logistic regression approach})}=0.7283$. This resulted in a relative difference of about 0.6%.

Figure 5. Simulation of 1000 samples from the rs12188950 data in Paper II. Histogram showing the relative difference, i.e. $(OR_{(\text{individual-based, logistic regression})} - OR_{(\text{allele-based})}) / OR_{(\text{allele-based})}$.



Conclusions and further aspects

Properties of genotypes in HWE, and interpretational problems

Departure from HWE may cause conflicting results when comparing allele-based ORs with corresponding ORs achieved from analyses using an individual-based approach. This might cause interpretational problems, even if the departure from sHWE is randomly occurring and rather modest. Let us again illustrate with an example, considering SNP rs12188950 in LSR, as described and analyzed in Paper II:

For the IS patient group, we have $N_{(CC)p}$=the number of major allele homozygotes; $N_{(CT)p}$=the number of heterozygotes; and $N_{(TT)p}$=the number of minor allele homozygotes. For the control subjects, we have $N_{(CC)c}$=the number of major allele homozygotes; $N_{(CT)c}$=the number of heterozygotes; and $N_{(TT)c}$=the number of minor allele homozygotes. We compile

$$
\begin{aligned}
OR_{(heterozygotes)} &= (N_{(CT)p} / N_{(CT)c})/(N_{(CC)p} / N_{(CC)c}) = (209/111)/(702/271) = 0.7269\\
OR_{(minor\text{-}homozygotes)} &= (N_{(TT)p} / N_{(TT)c})/(N_{(CC)p} / N_{(CC)c}) = (18/13) / (702/271) = 0.5345\\
OR_{(serological)} &= ((N_{(CT)p} + N_{(TT)p})/ (N_{(CT)c}+ N_{(TT)c})) / (N_{(CC)p} / N_{(CC)c})\\
&= (227/124) / (702/271)\\
&= 0.7067.
\end{aligned}
$$

It is true that $OR_{(minor\text{-}homozygotes)} = OR^2_{(heterozygotes)}$ when sHWE is present (**Sasieni 1997**). It is also true that $OR_{(heterozygotes)} = OR_{(allele\text{-}based\ approach)}$. Since $OR_{(allele\text{-}based\ approach)}$= 0.7240 and $OR^2_{(heterozygotes)} = 0.7269^2 = 0.5283$, one can notice a small discrepancy from the equivalences that would conform with that rule.

Another notation is that the serological odds ratio, $OR_{(serological)}$, always is taking a value between $OR_{(minor\text{-}homozygotes)}$ and $OR_{(heterozygotes)}$ (**Sasieni 1997**). The serological OR is also denoted the dominant minor allele approach while it is an indicator for any minor allele representation by an individual.

Comments regarding bootstrapping

The bootstrapping procedure (illustrated in **Figure 5**) was based on a sample, not a population. The sample (i.e. the original SNP rs12188950 sample from Paper II) was indeed regarded as an infinite population from which 1000 samples were independently drawn with replacement. This explains why the relative difference is not 0 (in fact, it is approaching the relative difference obtained from the original sample, i.e. 0.728-0.724=0.004).

One might also regard it this way: In a first step we selected 929 IS patients and 395 control subjects from an infinite population approaching a state of sHWE. In a second step we used that sample as a fictitious infinite "population" when we bootstrapped the 1000 simulated samples. The non-systematic departure from sHWE that occurred at the sample selection in the first step has now become a systematic error in the second step (i.e. when performing the bootstrapping procedure).

An alternative to this bootstrapping approach would be a simulation procedure based on a purely artificial population in strict HWE. Such manipulation of data might however violate the present, true data.

Further aspects of Bootstrapping procedures are addressed to the literature (e.g. **Efron 1982, Martin 2007, Schenker 1985**).

Is it possible to adjust for discrepancies between allelic-based and individual-based OR estimates ?

By performing weighted logistic regression, one can adjust for the discrepancies between allele-based and individual-based estimates of ORs. Hence, let p_p and p_c denote minor allele frequency for patients and control subjects, respectively. Furthermore, the observed number of (1) major allele homozygotes is $N_{(VV)p}$ and $N_{(VV)c}$; (2) minor allele homozygotes is $N_{(vv)p}$ and $N_{(vv)c}$; and (3) heterozygotes is $N_{(Vv)p}$ and $N_{(Vv)c}$.

For patients, the regression weights for every individual can be compiled by using the formula

W_i = expected ($N_{(VV)p}$)/ $N_{(VV)p}$ if genotype=0
W_i = expected ($N_{(Vv)p}$)/ $N_{(Vv)p}$ if genotype=1
W_i = expected ($N_{(vv)p}$)/ $N_{(vv)p}$ if genotype=2

(the weights will thus "reconstruct" a state of sHWE)

The expected frequencies are computed as

expected ($N_{(VV)p}$) = $N_p*(1-p_p)^2$
expected ($N_{(Vv)p}$) = $N_p*2*p_p*(1-p_p)^2$
e xpected ($N_{(vv)p}$) = $N_p*p_p^2$

For control subjects, the weights are compiled by the use of corresponding frequencies.

Let us return to the estimation of ORs for rs12188950 according to the data analyzed in Paper II. Three methods can be compared:

1. *Traditional allele-based OR calculation:*
   OR** = (245/1613) / (137/653) = **0.724 (95% CI: 0.576-0.909; P=0.0054)

2. *Simple logistic regression on individual-based data:*
   OR = 0.728 (95% CI: 0.581-913; P=0.0061)

3. *HWE-adjusting weighted simple logistic regression on individual-based data:*
   OR = 0.724 (95% CI: 0.576-0.910 ; P=0.0056)

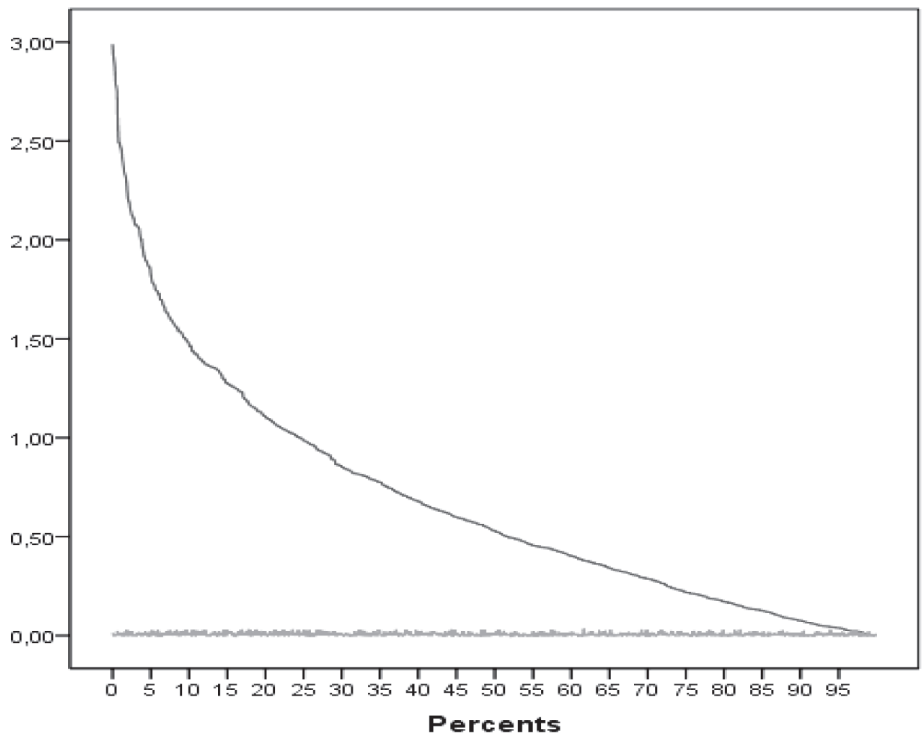
The above mentioned methods can be considered when returning to the bootstrapped simulation of 1000 samples described in the previous section. The allele-based approach for computing OR and the individual-based HWE-adjusting weighted logistic regression approach provided approximately similar estimates, with relative difference constantly touching zero as shown by a horizontal line in **Figure 6**, while the unweighted logistic regression provided relative differences shown by the decreasing curve.

It is however noteworthy that the HWE-adjusting weighted logistic regression method will fail if any cell is empty. This will be the case when modest sample sizes and small minor allele frequencies are considered.

It is thus possible to counterbalance distortions in the OR estimates caused by departure from sHWE. However, a subsequent issue will rice: will this add anything meaningful to the analysis?

That issue may be regarded as a matter of how to interpret the results.

Figure 6. Simulation of 1000 samples from the rs12188950 data in Paper II. Relative difference, i.e. $(OR_{(individual\text{-}based,\ logistic\ regression)}\text{-}OR_{(allele\text{-}based)})/\ OR_{(allele\text{-}based)}$ is shown on the Y-axis, percentiles of the simulations in sorted order on the X-axis.



Allelic-based ORs are unaffected by any departure from sHWE, whereas individual-based OR estimates from (unweighted) simple logistic regression are not. When carrying out an analysis of a case-control study (e.g. IS patients versus a reference group of control subjects) the primary focus is on individuals. It is the individual that suffers a

stroke, not the single allele. It is also the individual that is the carrier of phenotypic risk factors, such as hypertension, diabetes mellitus and heart diseases.

Nevertheless, the allelic variation (and therein, the single allele) is representing a possible cause of the augmented (or reversely, inhibited) stroke risk that is analyzed.

“A legend is an old man with a cane known for what he used to do. I’m still doing it.”

Miles Davis,
American trumpeter.
Born in 1926.
Died of a stroke in 1991.

Conclusions

– We examined family-history data to find possible indications that stroke may be transmitted within families. We found that these indications are true, especially when considering transmission of stroke to probands aged 75 years or younger. We also found that a family history of hypertension might be a risk factor for stroke (Paper I).

– We examined spouses to find out if stroke might be transmitted between non-biologically related close family members. We could not find any indication on this. (Paper I).

– We aimed to explore possible gender-dependent factors behind the inheritance of stroke. We found an association between mother and offspring regarding stroke that was considerably strong (Paper I).

– We expanded previous findings regarding a region within the PDE4D gene on chromosome 5, but we also found that the particular genetic marker rs12188950 (SNP45) was providing ambiguous results when analyzed against IS risk. This was explicitly shown by a meta-analysis (Paper II, meta-analysis updated and enhanced in Paper III).

– We performed a haplotype analysis on the five included SNPs within the PDE4D gene. We found that two of these SNPs, rs12188950 and rs3887175, were correlated with a linkage disequilibrium (R^2) of 0.63. Also, a haplotype block comprising four of the SNPs examined, was noted (which was in analogy with a previous, Icelandic study that our report partially referred to) (Paper II).

– We examined polymorphisms from different loci with a common histopathologic impact. Our aim was to find possible association between coronary artery disease-susceptible SNPs and IS risk. We compiled multi-locus risk scores based on these previous findings from GWAS-assessments and we subsequently performed assessments on the data. However, we did not find any significant association with IS, except for one particular genetic marker: SNP rs4977574 in chromosome 9p21.3 was significantly affecting IS risk. This finding was particularly evident when examining patients with IS subtype LVD, while no such association could be obtained when assessing IS patients subtyped with SVD or CE (Paper IV).

– We used a simple chi-square test to detect possible departure from Hardy-Weinberg equilibrium and we could not find such significant departure for any of the SNPs in any of the included studies (Papers II-IV). However, even coincidently occurring departure from strict HWE may be a subject for discussion: Such "trivial" departure may result in conflicting OR estimates when comparing results from allelic-based statistical analyses with corresponding results from logistic regression analysis on individual-based genotypic data. These conflicting OR estimates may be considered as an interpretational issue. We approached that problem in the *Further statistical considerations* section.

Acknowledgements

I would like to express my sincere gratitude to all those who have contributed to this thesis. Clinical research is a team-work where each helpful and encouraging co-worker is an eminent contributor.

First of all, and with sincere pleasure, I would like to thank my principal supervisor **Arne Lindgren** for his helpfulness and infallible determination to help me to get through with my thesis. Arne, we spent a lot of evenings discussing different neurologic topics in your office, and we did not always agree about how to cope with a specific issue or how to put words on findings aimed for publication or for oral presentation at a conference. But you were always there to aid and guide me, and despite that you were very often as busy as a bee, you did never hesitate to give me a lot of your expensive time to advice me and let me take part of your experience. Whenever I came to a dead stop in any of the projects, you unselfishly offered me your lunch break to help me if no other time was available in your calendar. I am very thankful for what you did for me through the years, Arne, even before I was formally adopted for post graduate studies!

To my second supervisor **Peter Höglund** I will give my deepest regards for his valuable devices and his willingness to help and guide me. Peter, you encouraged me already in 2003 to get involved in the Lund Stroke Register project and start a very good cooperation with Arne and his research team, first as a data manager and statistical adviser, and by the time passing also as an active researcher and post-graduate student. I have had many profound and positive discussions with you about statistical methods as well as ethical issues and neurological subjects from your perspectives throughout the years. I am very thankful for all the occasions when you, Peter, let me take part of your wisdom and experience through a cup of coffee in your office!

Moreover, I will profoundly thank my collaborators through the years within the Lund Stroke register:

Bo Norrving, whose solid knowledge about the scientific field of neurology was particularly valuable. Bo contributed as co-writer in all four papers enclosed with this thesis. Thank you, Bosse, for your constructive and encouraging comments and advices about e.g. writing and submitting manuscripts!

Ulf Kristoffersson, the clinical geneticist in the research team that contributed as co-writer in all four papers enclosed with this thesis.

Holger Luthman, whose solid knowledge about genetics has been a very important asset for the research group and for me.

Olle Melander, who generously let me implement and document a considerable part of a prestigious collaboration project between Lund, Malmö and Gothenburg that, inter alia, resulted in Papers III and IV. At the same time, I will mention the other co-writers of the papers linked to this thesis, namely **Gunnar Andsberg**, **Gunnar Engström**, **Björn Hallström**, **Bo Hedblad**, **Christina Jern**, **Katarina Jood**, **Sandra Olsson**, **Björn Petersen**, **Marketa Sjögren** and **Jan Gustav Smith**.

I have had many constructive discussions regarding Lund Stroke Register with my equal-in-merit colleagues **Hossein Delavaran** and **Ann-Cathrin Jönsson**. I have also received valuable informal support from other people in various positions at Lund Stroke Register, and by some students linked to the project. I will thank **Björn Hansen**, **Kristina Hermansdotter**, **Björn Larsson**, **Ingrid Lindgren**, **Gunilla Nilsson**, **Maria Norrving**, **Helene Starby** and all others more or less involved in LSR.

My sincere regards also to all my colleagues at the R&D-center Skåne (FoU-center Skåne), among others including:

Ulf Malmqvist, the head of the center, for offering me a favorable opportunity to get through with my thesis. Ulf, we made a gentlemen's agreement, and I am both proud and thankful to say that the two of us met that agreement.

Jonas Björk, who helped me make it possible to combine my effort to get through with my thesis with the daily work that had to be done. Jonas, you let me hold a number of seminars within the local professional statistician area, and that was important for me.

My closest statistician co-workers **Nuray Güner**, **Helene Jacobsson**, **Fredrik Nilsson** and **Susann Ullén**, and the secretary of the R&D-center, **Kerstin Tollstam**.

In addition, as a non-practitioner, I will express my sincerest appreciation to all those collaborators that may appear as "invisible records" to me: the stroke patients and the control subjects included in Lund Stroke Register. You gave us consent to use data representing a very intimate part of your personal integrity. We are very thankful for that confidence and we are doing our very best to administer it well.

…

Jag vill dessutom tacka familj, släkt och vänner. Min bror **Per Lövkvist** med fru och två vuxna barn och deras respektive måste nämnas i det sammanhanget. Pers dotter **Mia Ingmarsson** har för övrigt bidragit till denna avhandling genom att under en kort period arbeta som sommarvikarie i anslutning till Lund Stroke Register. Bland mina utomstående vänner bör de som jag sjunger gospel med i Matteuskyrkan i Malmö på onsdagskvällar ges särskild "cred".

Min mor **Alice Lövkvist**, som idag är 96 år, intar en särskild plats i mitt hjärta. Jag har alltid haft ditt stöd mamma, även i tider av motgångar såväl som när jag befann mig långt hemifrån och byggde upp min karriär. Du hade ett hem som jag kunde komma till, och du tog alltid emot mig med öppna armar, bland annat under de år som jag bodde i Norge. Utan din omtänksamhet och självuppoffrande modersroll hade jag inte kommit dit jag är idag.

…

This work was supported by grants from the Swedish Research Council, the Swedish State, the Swedish Heart and Lung Foundation, the Yngve Land Foundation for Neurological Research, the Crafoord Foundation, Region Skåne, the Freemasons Lodge of Instruction EOS in Lund, King Gustav V and Queen Victoria's Foundation, Lund University, the Department of Neurology Lund, the Swedish Stroke Association, the Tore Nilsson Foundation, the Lars Hierta Foundation, the Segerfalk foundation and the Emelle Foundation. Lund University and the Sahlgrenska Academy are members of the International Stroke Genetics Consortium.

“I do read music, but I prefer playing from the heart.”

Clarence Clemons,
American saxophonist.
Born in 1942.
Died of a stroke in 2011.

References

Adams HP, Jr., Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE, 3rd. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. Toast. Trial of org 10172 in acute stroke treatment. Stroke. 1993;24:35-41

Al Mamun AA, Peeters A, Barendregt J, Willekens F, Nusselder W, Bonneux L. Smoking decreases the duration of life lived with and without cardiovascular disease: A life course analysis of the Framingham Heart Study. Eur Heart J. 2004;25:409-415

Alkayed NJ, Harukuni I, Kimes AS, London ED, Traystman RJ, Hurn PD. Gender-linked brain injury in experimental stroke. Stroke. 1998;29:159-165; discussion 166

Baillie G, MacKenzie SJ, Houslay MD. Phorbol 12-myristate 13-acetate triggers the protein kinase a-mediated phosphorylation and activation of the *PDE4D5* cAMP phosphodiesterase in human aortic smooth muscle cells through a route involving extracellular signal regulated kinase (ERK). Mol Pharmacol. 2001;60:1100-1111

Balding DJ. A tutorial on statistical methods for population association studies. Nat Rev Genet. 2006;7:781-791

Bevan S, Dichgans M, Gschwendtner A, Kuhlenbaeumer G, Ringelstein EB, Markus HS. Variation in the *PDE4D* gene and ischemic stroke risk - a systematic review and meta-analysis on 5200 cases and 6600 controls. Stroke. 2008;39:1966-1971

Bevan S, Porteous L, Sitzer M, Markus HS. Phosphodiesterase 4d gene, ischemic stroke, and asymptomatic carotid atherosclerosis. Stroke. 2005;36:949-953

Bishop DT, Demenais F, Goldstein AM, Bergman W, Bishop JN, Bressac-de Paillerets B, Chompret A, Ghiorzo P, Gruis N, Hansson J, Harland M, Hayward N, Holland EA, Mann GJ, Mantelli M, Nancarrow D, Platz A, Tucker MA. Geographical variation in the penetrance of *CDKN2A* mutations for melanoma. J Natl Cancer Inst. 2002;94:894-903

Björk J, Albin M, Grahn P, Jacobsson H, Ardo J, Wadbro J, Östergren PO. Recreational values of the natural environment in relation to neighbourhood satisfaction, physical activity, obesity and wellbeing. J Epidemiol Community Health. 2008;62:e2

Borne Y, Engstrom G, Essen B, Hedblad B. Immigrant status and increased risk of heart failure: The role of hypertension and life-style risk factors. BMC Cardiovasc Disord. 2012;12:20

Broadbent HM, Peden JF, Lorkowski S, Goel A, Ongen H, Green F, Clarke R, Collins R, Franzosi MG, Tognoni G, Seedorf U, Rust S, Eriksson P, Hamsten A, Farrall M, Watkins H. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the *ANRIL* locus on chromosome 9p. Hum Mol Genet. 2008;17:806-814

Broad Institute 2008. The Haploview 4.2 software, available from http://broadinstitute.org/haploview/haploview, accessed July 25, 2012

Brophy VH, Ro SK, Rhees BK, Lui LY, Lee JM, Umblas N, Bentley LG, Li J, Cheng S, Browner WS, Erlich HA. Association of phosphodiesterase 4d polymorphisms with ischemic stroke in a US population stratified by hypertension status. Stroke. 2006;37:1385-1390

Cerrato P, Grasso M, Imperiale D, Priano L, Baima C, Giraudo M, Rizzuto A, Azzaro C, Lentini A, Bergamasco B. Stroke in young patients: Etiopathogenesis and risk factors in different age classes. Cerebrovasc Dis. 2004;18:154-159

Cheng YC, Anderson CD, Bione S, Keene K, Maguire JM, Nalls M, Rasheed A, Zeginigg M, Attia J, Baker R, Barlera S, Biffi A, Bookman E, Brott TG, Brown RD, Jr., Chen F, Chen WM, Ciusani E, Cole JW, Cortellini L, Danesh J, Doheny K, Ferrucci L, Grazia Franzosi M, Frossard P, Furie KL, Golledge J, Hankey GJ, Hernandez D, Holliday EG, Hsu FC, Jannes J, Kamal A, Khan MS, Kittner SJ, Koblar SA, Lewis M, Lincz L, Lisa A, Matarin M, Moscato P, Mychaleckyj JC, Parati EA, Parolo S, Pugh E, Rost NS, Schallert M, Schmidt H, Scott RJ, Sturm JW, Yadav S, Zaidi M, Boncoraglio GB, Levi CR, Meschia JF, Rosand J, Sale M, Saleheen D, Schmidt R, Sharma P, Worrall B, Mitchell BD. Are myocardial infarction-associated single-nucleotide polymorphisms associated with ischemic stroke? Stroke. 2012;43:980-986

Cluett C, McDermott MM, Guralnik J, Ferrucci L, Bandinelli S, Miljkovic I, Zmuda JM, Li R, Tranah G, Harris T, Rice N, Henley W, Frayling TM, Murray A, Melzer D. The 9p21 myocardial infarction risk allele increases risk of peripheral artery disease in older people. Circ Cardiovasc Genet. 2009;2:347-353

Conti M, Richter W, Mehats C, Livera G, Park JY, Jin C. Cyclic amp-specific *PDE4* phosphodiesterases as critical components of cyclic amp signaling. J Biol Chem. 2003;278:5493-5496

Costa FF. Non-coding RNAs, epigenetics and complexity. Gene. 2008;410:9-17

Dahlberg J, Smith G, Norrving B, Nilsson P, Hedblad B, Engström G, Lövkvist H, Carlson J, Lindgren A, Melander O. Genetic variants in serum and glucocortocoid regulated kinase 1, a regulator of the epithelial sodium channel, are associated with ischaemic stroke. J Hypertens. 2011;29:884-889

Daniel K, Wolfe CD, Busch MA, McKevitt C. What are the social consequences of stroke for working-aged adults? A systematic review. Stroke. 2009;40:e431-440

DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7:177-188

Dichgans M. Genetics of ischaemic stroke. Lancet Neurol. 2007;6:149-161

Domingues-Montanari S, Fernandez-Cadenas I, del Rio-Espinola A, Corbeto N, Krug T, Manso H, Gouveia L, Sobral J, Mendioroz M, Fernandez-Morales J, Alvarez-Sabin J, Ribo M, Rubiera M, Obach V, Marti-Fabregas J, Freijo M, Serena J, Ferro JM, Vicente AM, Oliveira SA, Montaner J. Association of a genetic variant in the *ALOX5AP* with higher risk of ischemic stroke: A case-control, meta-analysis and functional study. Cerebrovascular Diseases. 2010;29:528-537

Efron B. The jackknife, the bootstrap and other resampling plans. Philadelphia, Penn.: Society for Industrial and Applied Mathematics; 1982.

Efron B, Tibshirani R. An introduction to the bootstrap. New York: Chapman & Hall; 1994.

Erdmann J, Grosshennig A, Braund PS, Konig IR, Hengstenberg C, Hall AS, Linsel-Nitschke P, Kathiresan S, Wright B, Tregouet DA, Cambien F, Bruse P, Aherrahrou Z, Wagner AK, Stark K, Schwartz SM, Salomaa V, Elosua R, Melander O, Voight BF, O'Donnell

CJ, Peltonen L, Siscovick DS, Altshuler D, Merlini PA, Peyvandi F, Bernardinelli L, Ardissino D, Schillert A, Blankenberg S, Zeller T, Wild P, Schwarz DF, Tiret L, Perret C, Schreiber S, El Mokhtari NE, Schafer A, Marz W, Renner W, Bugert P, Kluter H, Schrezenmeir J, Rubin D, Ball SG, Balmforth AJ, Wichmann HE, Meitinger T, Fischer M, Meisinger C, Baumert J, Peters A, Ouwehand WH, Deloukas P, Thompson JR, Ziegler A, Samani NJ, Schunkert H. New susceptibility locus for coronary artery disease on chromosome 3q22.3. Nat Genet. 2009;41:280-282

Evans JF, Ferguson AD, Mosley RT, Hutchinson JH. What's all the flap about?: 5-lipoxygenase-activating protein inhibitors for inflammatory diseases. Trends Pharmacol Sci. 2008;29:72-78

Fagius J, Aquilonius S-M. Neurologi, 4:e reviderade uppl. Stockholm: Liber; 2006.

Fidani L, Clarimon J, Goulas A, Hatzitolios AI, Evans W, Tsirogianni E, Hardy J, Kotsis A. Association of phosphodiesterase 4D gene G0 haplotype and ischaemic stroke in a greek population. European Journal of Neurology. 2007;14:745-749

Flossmann E, Schulz UG, Rothwell PM. Systematic review of methods and results of studies of the genetic epidemiology of ischemic stroke. Stroke. 2004;35:212-227

Gadd M, Johansson SE, Sundquist J, Wandell P. The trend of cardiovascular disease in immigrants in sweden. Eur J Epidemiol. 2005;20:755-760

Garcia-Bermudez M, Gonzalez-Juanatey C, Lopez-Mejias R, Rodriguez-Rodriguez L, Perez-Esteban S, Castaneda S, Urcelay E, Miranda-Filloy JA, Gomez-Vaquero C, Fernandez-Gutierrez B, Balsa A, Gonzalez-Alvaro I, Blanco R, Llorca J, Martin J, Gonzalez-Gay MA. Influence of MHCIITA rs3087456 and rs4774 polymorphisms in the susceptibility to cardiovascular disease of patients with rheumatoid arthritis. Clin Exp Rheumatol. 2012;30:51-57

Ghaderi M, Gambelunghe G, Tortoioli C, Brozzetti A, Jatta K, Gharizadeh B, De Bellis A, Pecori Giraldi F, Terzolo M, Betterle C, Falorni A. *MHC2TA* single nucleotide polymorphism and genetic risk for autoimmune adrenal insufficiency. J Clin Endocrinol Metab. 2006;91:4107-4111

Ghatnekar O, Persson U, Glader EL, Terent A. Cost of stroke in sweden: An incidence estimate. Int J Technol Assess Health Care. 2004;20:375-380

Giskes K, Kunst AE, Benach J, Borrell C, Costa G, Dahl E, Dalstra JA, Federico B, Helmert U, Judge K, Lahelma E, Moussa K, Ostergren PO, Platt S, Prattala R, Rasmussen NK, Mackenbach JP. Trends in smoking behaviour between 1985 and 2000 in nine european countries by education. J Epidemiol Community Health. 2005;59:395-401

Gorelick PB. Cerebrovascular disease in african americans. Stroke. 1998;29:2656-2664

Grau AJ, Weimar C, Buggle F, Heinrich A, Goertler M, Neumaier S, Glahn J, Brandt T, Hacke W, Diener HC. Risk factors, outcome, and treatment in subtypes of ischemic stroke: The german stroke data bank. Stroke. 2001;32:2559-2566

Gretarsdottir S, Thorleifsson G, Reynisdottir ST, Manolescu A, Jonsdottir S, Jonsdottir T, Gudmundsdottir T, Bjarnadottir SM, Einarsson OB, Gudjonsdottir HM, Hawkins M, Gudmundsson G, Gudmundsdottir H, Andrason H, Gudmundsdottir AS, Sigurdardottir M, Chou TT, Nahmias J, Goss S, Sveinbjornsdottir S, Valdimarsson EM, Jakobsson F, Agnarsson U, Gudnason V, Thorgeirsson G, Fingerle J, Gurney M, Gudbjartsson D, Frigge ML, Kong A, Stefansson K, Gulcher JR. The gene encoding phosphodiesterase 4D confers risk of ischemic stroke. Nature Genetics. 2003;35:131-138

Gschwendtner A, Bevan S, Cole JW, Plourde A, Matarin M, Ross-Adams H, Meitinger T, Wichmann E, Mitchell BD, Furie K, Slowik A, Rich SS, Syme PD, MacLeod MJ, Meschia JF, Rosand J, Kittner SJ, Markus HS, Muller-Myhsok B, Dichgans M. Sequence variants on chromosome 9p21.3 confer risk for atherosclerotic stroke. Ann Neurol. 2009;65:531-539

Hallström B, Jönsson AC, Nerbrand C, Petersen B, Norrving B, Lindgren A. Lund stroke register: Hospitalization pattern and yield of different screening methods for first-ever stroke. Acta Neurol Scand. 2007;115:49-54

Hallström B, Jönsson AC, Nerbrand C, Norrving B, Lindgren A. Stroke incidence and survival in the beginning of the 21st century in southern Sweden: Comparisons with the late 20th century and projections into the future. Stroke. 2008;39:10-15

Hamrefors V, Hedblad B, Engström G, Almgren P, Sjögren M, Melander O. A myocardial infarction genetic risk score is associated with markers of carotid atherosclerosis. J Intern Med. 2011

Hannon GJ, Beach D. P15ink4b is a potential effector of tgf-beta-induced cell cycle arrest. Nature. 1994;371:257-261

Harismendy O, Notani D, Song X, Rahim NG, Tanasa B, Heintzman N, Ren B, Fu XD, Topol EJ, Rosenfeld MG, Frazer KA. 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. Nature. 2011;470:264-268

Helgadottir A, Manolescu A, Thorleifsson G, Gretarsdottir S, Jonsdottir H, Thorsteinsdottir U, Samani NJ, Gudmundsson G, Grant SF, Thorgeirsson G, Sveinbjornsdottir S, Valdimarsson EM, Matthiasson SE, Johannsson H, Gudmundsdottir O, Gurney ME, Sainz J, Thorhallsdottir M, Andresdottir M, Frigge ML, Topol EJ, Kong A, Gudnason V, Hakonarson H, Gulcher JR, Stefansson K. The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke. Nat Genet. 2004;36:233-239

Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Jonasdottir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiasdottir S, Jonsdottir T, Palsson S, Einarsdottir H, Gunnarsdottir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdottir U, Kong A, Stefansson K. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science. 2007;316:1491-1493

Hirsh J, O'Donnell M, Weitz JI. New anticoagulants. Blood. 2005;105:453-463

Holdt LM, Beutner F, Scholz M, Gielen S, Gabel G, Bergert H, Schuler G, Thiery J, Teupser D. *ANRIL* expression is associated with atherosclerosis risk at chromosome 9p21. Arterioscler Thromb Vasc Biol. 2010;30:620-627

Hosmer DW, Lemeshow S. Confidence interval estimation of interaction. Epidemiology. 1992;3:452-456

Ishii N, Maier D, Merlo A, Tada M, Sawamura Y, Diserens AC, Van Meir EG. Frequent co-alterations of TP53, P16/*CDKN2A*, $p14^{arf}$, PTEN tumor suppressor genes in human glioma cell lines. Brain Pathol. 1999;9:469-479

Jerrard-Dunne P, Cloud G, Hassan A, Markus HS. Evaluating the genetic component of ischemic stroke subtypes: A family history study. Stroke. 2003;34:1364-1369

Jin SL, Conti M. Induction of the cyclic nucleotide phosphodiesterase *PDE4B* is essential for LPS-activated TNF-alpha responses. Proc Natl Acad Sci U S A. 2002;99:7628-7633

Jood K, Ladenvall C, Rosengren A, Blomstrand C, Jern C. Family history in ischemic stroke before 70 years of age: The Sahlgrenska Academy Study on Ischemic Stroke. Stroke. 2005;36:1383-1387

Kalinina N, Agrotis A, Antropova Y, Ilyinskaya O, Smirnov V, Tararak E, Bobik A. Smad expression in human atherosclerotic lesions: Evidence for impaired TGF-beta/Smad signaling in smooth muscle cells of fibrofatty lesions. Arterioscler Thromb Vasc Biol. 2004;24:1391-1396

Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand S, Engert JC, Samani NJ, Schunkert H, Erdmann J, Reilly MP, Rader DJ, Morgan T, Spertus JA, Stoll M, Girelli D, McKeown PP, Patterson CC, Siscovick DS, O'Donnell CJ, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Melander O, Altshuler D, Ardissino D, Merlini PA, Berzuini C, Bernardinelli L, Peyvandi F, Tubaro M, Celli P, Ferrario M, Fetiveau R, Marziliano N, Casari G, Galli M, Ribichini F, Rossi M, Bernardi F, Zonzin P, Piazza A, Mannucci PM, Schwartz SM, Siscovick DS, Yee J, Friedlander Y, Elosua R, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Kathiresan S, Meigs JB, Williams G, Nathan DM, MacRae CA, O'Donnell CJ, Salomaa V, Havulinna AS, Peltonen L, Melander O, Berglund G, Voight BF, Kathiresan S, Hirschhorn JN, Asselta R, Duga S, Spreafico M, Musunuru K, Daly MJ, Purcell S, Voight BF, Purcell S, Nemesh J, Korn JM, McCarroll SA, Schwartz SM, Yee J, Kathiresan S, Lucas G, Subirana I, Elosua R, Surti A, Guiducci C, Gianniny L, Mirel D, Parkin M, Burtt N, Gabriel SB, Samani NJ, Thompson JR, Braund PS, Wright BJ, Balmforth AJ, Ball SG, Hall AS, Schunkert H, Erdmann J, Linsel-Nitschke P, Lieb W, Ziegler A, Konig I, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann HE, Schreiber S, Schunkert H, Samani NJ, Erdmann J, Ouwehand W, Hengstenberg C, Deloukas P, Scholz M, Cambien F, Reilly MP, Li M, Chen Z, Wilensky R, Matthai W, Qasim A, Hakonarson HH, Devaney J, Burnett MS, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Knouff CW, Waterworth DM, Walker MC, Mooser V, Epstein SE, Rader DJ, Scheffold T, Berger K, Stoll M, Huge A, Girelli D, Martinelli N, Olivieri O, Corrocher R, Morgan T, Spertus JA, McKeown P, Patterson CC, Schunkert H, Erdmann E, Linsel-Nitschke P, Lieb W, Ziegler A, Konig IR, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann HE, Schreiber S, Holm H, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Engert JC, Do R, Xie C, Anand S, Kathiresan S, Ardissino D, Mannucci PM, Siscovick D, O'Donnell CJ, Samani NJ, Melander O, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Altshuler D. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet. 2009;41:334-341

Khan FA, Zia E, Janzon L, Engstrom G. Incidence of stroke and stroke subtypes in malmo, sweden, 1990-2000: Marked differences between groups defined by birth country. Stroke. 2004;35:2054-2058

Khan U, Porteous L, Hassan A, Markus HS. Risk factor profile of cerebral small vessel disease and its subtypes. J Neurol Neurosurg Psychiatry. 2007;78:702-706

Kim DH, Ahn WY, Kim DK, Choe BK, Kim SK, Jo DJ, Kim JY, Chung JH, Jeong YS, Yun DH, Yoo SD, Kim HS, Baik HH. A promoter polymorphism (rs17222919, -1316t/g) of *ALOX5AP* is associated with intracerebral hemorrhage in korean population. Prostaglandins Leukot Essent Fatty Acids. 2011;85:115-120

Kristoffersson U. Medicinsk genetik : En introduktion. Lund: Studentlitteratur; 2003

Knudson AG, Jr. Hereditary cancer, oncogenes, and antioncogenes. Cancer Res. 1985;45:1437-1443

Kostulas K, Gretarsdottir S, Kostulas V, Manolescu A, Helgadottir A, Thorleifsson G, Gudmundsson LJ, Thorsteinsdottir U, Gulcher JR, Stefansson K, Hillert J. *PDE4D* and *ALOX5AP* genetic variants and risk for ischemic cerebrovascular disease in sweden. Journal of the Neurological Sciences. 2007;263:113-117

Kuhlenbaumer G, Berger K, Huge A, Lange E, Kessler C, John U, Funke H, Nabavi DG, Stogbauer F, Ringelstein EB, Stoll M. Evaluation of single nucleotide polymorphisms in the phosphodiesterase 4D gene (*PDE4D*) and their association with ischaemic stroke in a large german cohort. Journal of Neurology Neurosurgery and Psychiatry. 2006;77:521-524

Kuo L, Davis MJ, Cannon MS, Chilian WM. Pathophysiological consequences of atherosclerosis extend into the coronary microcirculation. Restoration of endothelium-dependent responses by L-arginine. Circ Res. 1992;70:465-476

Lachance J. Detecting selection-induced departures from hardy-weinberg proportions. Genet Sel Evol. 2009;41:15

Lamy C, Giannesini C, Zuber M, Arquizan C, Meder JF, Trystram D, Coste J, Mas JL. Clinical and imaging findings in cryptogenic stroke patients with and without patent foramen ovale: The PFO-ASA study. Atrial Septal Aneurysm. Stroke. 2002;33:706-711

Landells LJ, Szilagy CM, Jones NA, Banner KH, Allen JM, Doherty A, O'Connor BJ, Spina D, Page CP. Identification and quantification of phosphodiesterase 4 subtypes in CD4 and CD8 lymphocytes from healthy and asthmatic subjects. Br J Pharmacol. 2001;133:722-729

Lehnart SE, Wehrens XH, Reiken S, Warrier S, Belevych AE, Harvey RD, Richter W, Jin SL, Conti M, Marks AR. Phosphodiesterase 4D deficiency in the ryanodine-receptor complex promotes heart failure and arrhythmias. Cell. 2005;123:25-35

Lemmens R, Abboud S, Robberecht W, Vanhees L, Pandolfo M, Thijs V, Goris A. Variant on 9p21 strongly associates with coronary heart disease, but lacks association with common stroke. Eur J Hum Genet. 2009;17:1287-1293

Li C, Hedblad B, Rosvall M, Buchwald F, Khan FA, Engstrom G. Stroke incidence, recurrence, and case-fatality in relation to socioeconomic position: A population-based study of middle-aged swedish men and women. Stroke. 2008;39:2191-2196

Lin HF, Liao YC, Liou CW, Liu CK, Juo SH. The phosphodiesterase 4d gene for early onset ischemic stroke among normotensive patients. J Thromb Haemost. 2007;5:436-438

Lindholm E, Melander O, Almgren P, Berglund G, Agardh CD, Groop L, Orho-Melander M. Polymorphism in the *MHC2TA* gene is associated with features of the metabolic syndrome and cardiovascular mortality. PLoS One. 2006;1:e64

Linsel-Nitschke P, Gotz A, Medack A, Konig IR, Bruse P, Lieb W, Mayer B, Stark K, Hengstenberg C, Fischer M, Baessler A, Ziegler A, Schunkert H, Erdmann J. Genetic variation in the arachidonate 5-lipoxygenase-activating protein (*ALOX5AP*) is associated with myocardial infarction in the german population. Clin Sci (Lond). 2008;115:309-315

Liu H, Maurice DH. Phosphorylation-mediated activation and translocation of the cyclic amp-specific phosphodiesterase *PDE4D3* by cyclic amp-dependent protein kinase and mitogen-activated protein kinases. A potential mechanism allowing for the coordinated regulation of *PDE4D* activity and targeting. J Biol Chem. 1999;274:10557-10565

Liu H, Palmer D, Jimmo SL, Tilley DG, Dunkerley HA, Pang SC, Maurice DH. Expression of phosphodiesterase 4D (*PDE4D*) is regulated by both the cyclic amp-dependent

protein kinase and mitogen-activated protein kinase signaling pathways. A potential mechanism allowing for the coordinated regulation of *PDE4D* activity and expression in cells. J Biol Chem. 2000;275:26615-26624

Lloyd-Jones DM, Leip EP, Larson MG, D'Agostino RB, Beiser A, Wilson PW, Wolf PA, Levy D. Prediction of lifetime risk for cardiovascular disease by risk factor burden at 50 years of age. Circulation. 2006;113:791-798

Lowe SW, Sherr CJ. Tumor suppression by ink4a-arf: Progress and puzzles. Curr Opin Genet Dev. 2003;13:77-83

Lunetta KL. Genetic association studies. Circulation. 2008;118:96-101

Lõhmussaar E, Gschwendtner A, Mueller JC, Org T, Wichmann E, Hamann G, Meitinger T, Dichgans M. *ALOX5AP* gene and the *PDE4D* gene in a Central European population of stroke patients. Stroke. 2005;36:731-736

Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. N Engl J Med. 1994;330:1041-1046

Mariotti C, Savarese N, Suomalainen A, Rimoldi M, Comi G, Prelle A, Antozzi C, Servidei S, Jarre L, DiDonato S, Zeviani M. Genotype to phenotype correlations in mitochondrial encephalomyopathies associated with the A3243G mutation of mitochondrial DNA. J Neurol. 1995;242:304-312

Martin MA. Bootstrap hypothesis testing for some common statistical problems: A critical evaluation of size and power statistics. Computational Statistics & Data Analysis. 2007;51:6321-6342

Martinez-Fernandez E, Gil-Peralta A, Garcia-Lozano R, Chinchon I, Aguilera I, Fernandez-Lopez O, Arenas J, Campos Y, Bautista J. Mitochondrial disease and stroke. Stroke. 2001;32:2507-2510

Matsushita T, Kubo M, Yonemoto K, Ninomiya T, Ashikawa K, Liang BL, Hata J, Doi Y, Kitazono T, Ibayashi S, Iida M, Kiyohara Y, Nakamura Y. Lack of association between variations of *PDE4D* and ischemic stroke in the Japanese population. Stroke. 2009;40:1245-1251

Medin J, Nordlund A, Ekberg K. Increasing stroke incidence in Sweden between 1989 and 2000 among persons aged 30 to 65 years: Evidence from the Swedish hospital discharge register. Stroke. 2004;35:1047-1051

Meschia JF, Brott TG, Brown RD, Crook R, Worrall BB, Kissela B, Brown WM, Rich SS, Case LD, Evans EW, Hague S, Singleton A, Hardy J. Phosphodiesterase 4D and 5-lipoxygenase activating protein in ischemic stroke. Annals of Neurology. 2005;58:351-361

Meschia JF. Clinically translated ischemic stroke genomics. Stroke. 2004;35:2735-2739

Morris JA, Gardner MJ. Calculating confidence intervals for relative risks (odds ratios) and standardised ratios and rates. Br Med J (Clin Res Ed). 1988;296:1313-1316

Murphy R, Turnbull DM, Walker M, Hattersley AT. Clinical features, diagnosis and management of maternally inherited diabetes and deafness (MIDD) associated with the 3243A>G mitochondrial point mutation. Diabet Med. 2008;25:383-399

NCBI 2012.1: SNP cluster report. Available from: http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=4977574, accessed July 6, 2012

NCBI 2012.2: SNP gene summary. Available from: http://www.ncbi.nlm.nih.gov/gene/100048912, accessed July 6, 2012

NIH 2011: NIH National Library of Medicine 2011 MeSH search. Available from: http://www.nlm.nih.gov/cgi/mesh/2011/MB_cgi?mode=&term=Second+Messenger+Systems, accessed July 10, 2012

Nakayama T, Asai S, Sato N, Soma M. Genotype and haplotype association study of the STRK1 region on 5q12 among Japanese - A case-control study. Stroke. 2006;37:69-76

National Stroke Association 2012: Available from http://www.stroke.org/site/PageServer?pagename=aamer, accessed June 28, 2012

Nature-Genetics-Editorial. Census, race and science. Nat Genet. 2000;24:97-98

Nelson DL, Cox MM, Lehninger AL. Lehninger principles of biochemistry, 5th edition. New York: W. H. Freeman; 2008.

Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, Najjar SS, Zhao JH, Heath SC, Eyheramendy S, Papadakis K, Voight BF, Scott LJ, Zhang F, Farrall M, Tanaka T, Wallace C, Chambers JC, Khaw KT, Nilsson P, van der Harst P, Polidoro S, Grobbee DE, Onland-Moret NC, Bots ML, Wain LV, Elliott KS, Teumer A, Luan J, Lucas G, Kuusisto J, Burton PR, Hadley D, McArdle WL, Brown M, Dominiczak A, Newhouse SJ, Samani NJ, Webster J, Zeggini E, Beckmann JS, Bergmann S, Lim N, Song K, Vollenweider P, Waeber G, Waterworth DM, Yuan X, Groop L, Orho-Melander M, Allione A, Di Gregorio A, Guarrera S, Panico S, Ricceri F, Romanazzi V, Sacerdote C, Vineis P, Barroso I, Sandhu MS, Luben RN, Crawford GJ, Jousilahti P, Perola M, Boehnke M, Bonnycastle LL, Collins FS, Jackson AU, Mohlke KL, Stringham HM, Valle TT, Willer CJ, Bergman RN, Morken MA, Doring A, Gieger C, Illig T, Meitinger T, Org E, Pfeufer A, Wichmann HE, Kathiresan S, Marrugat J, O'Donnell CJ, Schwartz SM, Siscovick DS, Subirana I, Freimer NB, Hartikainen AL, McCarthy MI, O'Reilly PF, Peltonen L, Pouta A, de Jong PE, Snieder H, van Gilst WH, Clarke R, Goel A, Hamsten A, Peden JF, Seedorf U, Syvanen AC, Tognoni G, Lakatta EG, Sanna S, Scheet P, Schlessinger D, Scuteri A, Dorr M, Ernst F, Felix SB, Homuth G, Lorbeer R, Reffelmann T, Rettig R, Volker U, Galan P, Gut IG, Hercberg S, Lathrop GM, Zelenika D, Deloukas P, Soranzo N, Williams FM, Zhai G, Salomaa V, Laakso M, Elosua R, Forouhi NG, Volzke H, Uiterwaal CS, van der Schouw YT, Numans ME, Matullo G, Navis G, Berglund G, Bingham SA, Kooner JS, Connell JM, Bandinelli S, Ferrucci L, Watkins H, Spector TD, Tuomilehto J, Altshuler D, Strachan DP, Laan M, Meneton P, Wareham NJ, Uda M, Jarvelin MR, Mooser V, Melander O, Loos RJ, Elliott P, Abecasis GR, Caulfield M, Munroe PB. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet. 2009;41:666-676

Nilsson-Ardnor S, Wiklund PG, Lindgren P, Nilsson AK, Janunger T, Escher SA, Hallbeck B, Stegmayr B, Asplund K, Holmberg D. Linkage of ischemic stroke to the *PDE4D* region on 5q in a Swedish population. Stroke. 2005;36:1666-1671

Olsson S, Jood K, Blomstrand C, Jern C. Genetic variation on chromosome 9p21 shows association with the ischaemic stroke subtype large-vessel disease in a swedish sample aged </= 70. Eur J Neurol. 2011.1;18:365-367

Olsson S, Jood K, Melander O, Sjögren M, Norrving B, Nilsson M, Lindgren A, Jern C. Lack of association between genetic variations in the *KALRN* region and ischemic stroke. Clin Biochem. 2011.2;44:1018-1020

Olsson S, Melander O, Jood K, Smith JG, Lövkvist H, Sjögren M, Engström G, Norrving B, Lindgren A, Jern C. Genetic variant on chromosome 12p13 does not show association to ischemic stroke in 3 Swedish case-control studies. Stroke. 2011.3;42:214-216

Oudin A, Strömberg U, Jakobsson K, Stroh E, Bjork J. Estimation of short-term effects of air pollution on stroke hospital admissions in Southern Sweden. Neuroepidemiology. 2010;34:131-142

Pan X, Arauz E, Krzanowski JJ, Fitzpatrick DF, Polson JB. Synergistic interactions between selective pharmacological inhibitors of phosphodiesterase isozyme families *PDE III* and *PDE IV* to attenuate proliferation of rat vascular smooth muscle cells. Biochem Pharmacol. 1994;48:827-835

Pantoni L. Cerebral small vessel disease: From pathogenesis and clinical characteristics to therapeutic challenges. Lancet Neurol. 2010;9:689-701

Preuss M, Konig IR, Thompson JR, Erdmann J, Absher D, Assimes TL, Blankenberg S, Boerwinkle E, Chen L, Cupples LA, Hall AS, Halperin E, Hengstenberg C, Holm H, Laaksonen R, Li M, Marz W, McPherson R, Musunuru K, Nelson CP, Burnett MS, Epstein SE, O'Donnell CJ, Quertermous T, Rader DJ, Roberts R, Schillert A, Stefansson K, Stewart AF, Thorleifsson G, Voight BF, Wells GA, Ziegler A, Kathiresan S, Reilly MP, Samani NJ, Schunkert H. Design of the coronary artery disease genome-wide replication and meta-analysis (cardiogram) study: A genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. Circ Cardiovasc Genet. 2010;3:475-483

Quarta G, Stanzione R, Evangelista A, Zanda B, Di Angelantonio E, Marchitti S, Di Castro S, Di Vavo M, Volpe M, Rubattu S. Phosphodiesterase 4D and 5-lipoxygenase activating protein genes and risk of ischemic stroke in Sardinians. European Journal of Human Genetics. 2009;17:1448-1453

Reilly MP, Li M, He J, Ferguson JF, Stylianou IM, Mehta NN, Burnett MS, Devaney JM, Knouff CW, Thompson JR, Horne BD, Stewart AF, Assimes TL, Wild PS, Allayee H, Nitschke PL, Patel RS, Martinelli N, Girelli D, Quyyumi AA, Anderson JL, Erdmann J, Hall AS, Schunkert H, Quertermous T, Blankenberg S, Hazen SL, Roberts R, Kathiresan S, Samani NJ, Epstein SE, Rader DJ. Identification of *ADAMTS7* as a novel locus for coronary atherosclerosis and association of *ABO* with myocardial infarction in the presence of coronary atherosclerosis: Two genome-wide association studies. Lancet. 2010;377:383-392

Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, Guiducci C, Perola M, Jula A, Sinisalo J, Lokki ML, Nieminen MS, Melander O, Salomaa V, Peltonen L, Kathiresan S. A multilocus genetic risk score for coronary heart disease: Case-control and prospective cohort analyses. Lancet. 2010;376:1393-1400

Rosand J, Bayley N, Rost N, de Bakker PI. Many hypotheses but no replication for the association between *PDE4D* and stroke. Nat Genet. 2006;38:1091-1092; author reply 1092-1093

Ross R. The pathogenesis of atherosclerosis: A perspective for the 1990s. Nature. 1993;362:801-809

Saleheen D, Bukhari S, Haider SR, Nazir A, Khanum S, Shafqat S, Anis K, Frossard P. Association of phosphodiesterase 4D gene with ischemic stroke in a Pakistani population. Stroke. 2005;36:2275-2277

Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barrett JH, Konig IR, Stevens SE, Szymczak S, Tregouet DA, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H. Genomewide association analysis of coronary artery disease. N Engl J Med. 2007;357:443-453

Sasieni PD. From genotypes to genes: Doubling the sample size. Biometrics. 1997;53:1253-1261

Schaefer AM, Taylor RW, Turnbull DM, Chinnery PF. The epidemiology of mitochondrial disorders--past, present and future. Biochim Biophys Acta. 2004;1659:115-120

Schenker N. Qualms about bootstrap confidence intervals. J Am Statistical Association. 1985;80:360-361

Schunkert H, Erdmann J, Samani NJ. Genetics of myocardial infarction: A progress report. Eur Heart J. 2010;31:918-925

Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Muhleisen TW, Muhlestein JB, Munzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nothen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schafer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C, Witteman JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, Marz W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J, Samani NJ. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet. 2011;43:333-338

Smith JG, Melander O, Lövkvist H, Hedblad B, Engström G, Nilsson P, Carlson J, Berglund G, Norrving B, Lindgren A. Common genetic variants on chromosome 9p21 confers risk of ischemic stroke: A large-scale genetic association study. Circ Cardiovasc Genet. 2009;2:159-164

Song Q, Cole JW, O'Connell JR, Stine OC, Gallagher M, Giles WH, Mitchell BD, Wozniak MA, Stern BJ, Sorkin JD, McArdle PF, Naj AC, Xu Q, Gibbons GH, Kittner SJ. Phosphodiesterase 4d polymorphisms and the risk of cerebral infarction in a biracial population: The stroke prevention in young women study. Human Molecular Genetics. 2006;15:2468-2478

Statistics Sweden 2012: Available from: http://www.ssd.scb.se/databaser/makro/start.asp, accessed June 25, 2012

Staton JM, Sayer MS, Hankey GJ, Attia J, Thakkinstian A, Yi Q, Cole VJ, Baker R, Eikelboom JW. Association between phosphodiesterase 4d gene and ischaemic stroke. Journal of Neurology Neurosurgery and Psychiatry. 2006;77:1067-1069

Strachan T, Read AP. Human molecular genetics 3. London: Garland Science; 2004

Sun Y, Huang YY, Chen X, Liu Y, Lu XZ, Shi YY, Tang W, Yang JD, Chen WY, Zhao XZ, Gao LH, Li S, Feng GY, He L. Association between the *PDE4D* gene and ischaemic stroke in the Chinese Han population. Clinical Science. 2009;117:265-272

Swanberg M, Lidman O, Padyukov L, Eriksson P, Akesson E, Jagodic M, Lobell A, Khademi M, Borjesson O, Lindgren CM, Lundman P, Brookes AJ, Kere J, Luthman H, Alfredsson L, Hillert J, Klareskog L, Hamsten A, Piehl F, Olsson T. *MHC2TA* is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. Nat Genet. 2005;37:486-494

Tentschert S, Greisenegger S, Wimmer R, Lang W, Lalouschek W. Association of parental history of stroke with clinical parameters in patients with ischemic stroke or transient ischemic attack. Stroke. 2003;34:2114-2119

The Swedish National Board of Health and Welfare 2012: Available from http://www.socialstyrelsen.se/Lists/ Artikelkatalog/Attachments/18623/2012-3-6.pdf, accessed June 28, 2012, in Swedish

The world health organization MONICA project (monitoring trends and determinants in cardiovascular disease): A major international collaboration. Who MONICA project principal investigators. J Clin Epidemiol. 1988;41:105-114

Thorvaldsen P, Asplund K, Kuulasmaa K, Rajakangas AM, Schroll M. Stroke incidence, case fatality, and mortality in the WHO MONICA project. World Health Organization monitoring trends and determinants in cardiovascular disease. Stroke. 1995;26:361-367

Tobis MJ, Azarbal B. Does patent foramen ovale promote cryptogenic stroke and migraine headache? Tex Heart Inst J. 2005;32:362-365

Toivanen S. Income differences in stroke mortality: A 12-year follow-up study of the Swedish working population. Scand J Public Health. 2011;39:797-804

Tournier-Lasserve E, Iba-Zizen MT, Romero N, Bousser MG. Autosomal dominant syndrome with strokelike episodes and leukoencephalopathy. Stroke. 1991;22:1297-1302

van Gijn J, Kerr RS, Rinkel GJ. Subarachnoid haemorrhage. Lancet. 2007;369:306-318

van Rijn MJE, Slooter AJC, Schut AFC, Isaacs A, Aulchenko YS, Snijders P, Kappelle LJ, van Swieten JC, Oostra BA, van Duijn CM. Familial aggregation, the *PDE4D* gene, and ischemic stroke in a genetically isolated population. Neurology. 2005;65:1203-1209

Warlow C, Sudlow C, Dennis M, Wardlaw J, Sandercock P. Stroke. Lancet. 2003;362:1211-1224

Welin L, Svardsudd K, Wilhelmsen L, Larsson B, Tibblin G. Analysis of risk factors for stroke in a cohort of men born in 1913. N Engl J Med. 1987;317:521-526

Wellek S. Tests for establishing compatibility of an observed genotype distribution with hardy-weinberg equilibrium in the case of a biallelic locus. Biometrics. 2004;60:694-703

Verny C, Amati-Bonneau P, Letournel F, Person B, Dib N, Malinge MC, Slama A, Le Marechal C, Ferec C, Procaccio V, Reynier P, Bonneau D. Mitochondrial DNA A3243G mutation involved in familial diabetes, chronic intestinal pseudo-obstruction and recurrent pancreatitis. Diabetes Metab. 2008;34:620-626

Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, Blow MJ, Cohen JC, Rubin EM, Pennacchio LA. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. Nature. 2010;464:409-412

Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: The Framingham Study. Stroke. 1991;22:983-988

Wolf PA. Prevention of stroke. Lancet. 1998;352 Suppl 3:SIII15-18

Woo D, Kaushal R, Kissela B, Sekar P, Wolujewicz M, Pal P, Alwell K, Haverbusch M, Ewing I, Miller R, Kleindorfer D, Flaherty M, Chakraborty R, Deka R, Broderick J. Association of phosphodiesterase 4D with ischemic stroke - A population-based case-control study. Stroke. 2006;37:371-376

Woolf B. On estimating the relation between blood group and disease. Ann Hum Genet. 1955;19:251-253

WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447:661-678

Xu XW, Li X, Li JJ, Ou R, Sheng WL. Meta-analysis of association between variation in the *PDE4D* gene and ischemic cerebral infarction risk in Asian populations. Neurogenetics. 2010;11:327-333

Yoon D, Park SK, Kang D, Park T, Park JW. Meta-analysis of homogeneous subgroups reveals association between *PDE4D* gene variants and ischemic stroke. Neuroepidemiology. 2011;36:213-222

You RX, Thrift AG, McNeil JJ, Davis SM, Donnan GA. Ischemic stroke risk and passive exposure to spouses' cigarette smoking. Melbourne Stroke Risk Factor Study (MERFS) group. Am J Public Health. 1999;89:572-575

Zee RYL, Brophy VH, Cheng S, Hegener HH, Erlich HA, Ridker PM. Polymorphisms of the phosphodiesterase 4D, cAMP-specific (*PDE4D*) gene and risk of ischemic stroke - a prospective, nested case-control evaluation. Stroke. 2006;37:2012-2017

– – –