



# LUND UNIVERSITY

## Population-based studies of risk factors for subarachnoid hemorrhage

Söderholm, Martin

2016

[Link to publication](#)

*Citation for published version (APA):*

Söderholm, M. (2016). *Population-based studies of risk factors for subarachnoid hemorrhage*. [Doctoral Thesis (compilation), Department of Clinical Sciences, Malmö]. Lund University: Faculty of Medicine.

*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



The background of the slide is a high-magnification, black and white micrograph of brain tissue, likely showing a subarachnoid hemorrhage. The image displays a dense network of dark, irregular structures against a lighter, textured background. A white rectangular box is positioned in the upper left quadrant, containing the title and author information.

# Population-based studies of risk factors for subarachnoid hemorrhage

MARTIN SÖDERHOLM | FACULTY OF MEDICINE | LUND UNIVERSITY 2016







**LUND UNIVERSITY**  
Faculty of Medicine

Lund University, Faculty of Medicine  
Doctoral Dissertation Series 2016:118  
ISBN 978-91-7619-344-0  
ISSN 1652-8220



# Population-based studies of risk factors for subarachnoid hemorrhage

Martin Söderholm



**LUND**  
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Medicine, Lund University, Sweden.

To be defended at Jubileumsaulan, Jan Waldenströms gata 5, Skånes

Universitetssjukhus, Malmö, Friday 11 November 2016, 1.00pm.


*Faculty opponent*

Associate Professor Peter Appelros,

Örebro University and Örebro University hospital

Organization LUND UNIVERSITY	Document name DOCTORAL DISSERTATION			
	Date of issue: November 11, 2016			
Author(s): Martin Söderholm	Sponsoring organization			
Title and subtitle: Population-based studies of risk factors for subarachnoid hemorrhage				
<p><b>Abstract</b></p> <p>Subarachnoid hemorrhage (SAH), the least common type of stroke, is a highly fatal vascular disease, in most cases caused by a rupture of an arterial intracranial aneurysm. SAH is more common in women than in men and the mean age at SAH is lower compared to other strokes, often affecting people of working ages. However, the causes of SAH are not fully elucidated. The aim was to study potential risk factors for SAH in population-based settings.</p> <p>Lung volumes assessed in spirometry was studied in association to incidence of SAH in 28,000 subjects from the general population (150 SAH events). Stroke incidence was also evaluated in 100,000 patients with chronic obstructive pulmonary disease (COPD) based on nation-wide register data, comparing the rates to the general population. Leukocyte count was studied for SAH incidence in a cohort study. In nested case control studies, we studied serum levels of fibroblast growth factor 23 (FGF23), and exome-wide coding genetic variation, in relation to SAH risk.</p> <p>Subjects with moderately reduced lung function had increased incidence of SAH, even after adjusting for smoking and in non-smokers. COPD patients had an increased risk of all stroke subtypes; ischemic stroke, intracerebral hemorrhage, and SAH. High leukocyte count and high levels of FGF23, which both may indicate systemic inflammation, was associated with risk of incident SAH. No coding genetic variants were strongly associated with SAH, but a few previously reported common loci for intracranial aneurysms were detected also for SAH.</p> <p>The results suggest that mechanisms related to reduced lung function, and low-grade inflammation may be involved in development of SAH. It is still not known if the association between COPD and stroke is due to shared risk factors or if there is a causal relationship. However, risk factor interventions in COPD patients seem important to reduce stroke risk.</p>				
Key words: Subarachnoid haemorrhage, stroke, ischemic stroke, intracerebral haemorrhage, epidemiology, population study, forced expiratory volume in 1 second, COPD, leukocytes, inflammation, fibroblast growth factors, exome, genome-wide association study, nested case controls study, cohort study				
Classification system and/or index terms (if any)				
Supplementary bibliographical information		Language: English		
ISSN and key title: 1652-8220		ISBN: 978-91-7619-344-0		
Recipient's notes	Number of pages 92	Price		
	Security classification			

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 5/10/2016

# Population-based studies of risk factors for subarachnoid hemorrhage

Martin Söderholm



**LUND**  
UNIVERSITY

Coverphoto by Set Söderholm, “Frottage av motorsågad träyta”.

Copyright (Martin Söderholm)

Faculty of Medicine | Department of Clinical Sciences Malmö

Lund University, Faculty of Medicine Doctoral Dissertation Series 2016:118

ISBN 978-91-7619-344-0

ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University, Lund 2016



# Table of contents

List of papers	7
Abbreviations	9
Introduction	11
General aspects of stroke epidemiology	12
Subarachnoid hemorrhage	14
Diagnosis	15
Causes of SAH	15
Intracranial aneurysms	15
Pathophysiology of aneurysms	17
Incidence and prognosis	18
Risk factors	20
Complications and Treatment	23
Family history and genetic studies of SAH	25
Further studies of genetic determinants	27
Risk factors for ischemic stroke and ICH	27
Lung function, COPD and stroke	28
COPD	29
Aims	31
Methods	33
Study populations	33
The Malmö preventive project	33
The Malmö Diet and Cancer Study	33
The COPD cohort	35
Assessment of exposure	36
Paper I	36
Paper II	37
Paper III	37
Paper V	38
Ascertainment of outcome	38
Paper I-III and V	38
Paper IV	40



Study design and statistical analysis	40
Cohort studies	40
Interaction and sensitivity analyses	42
Nested case-control studies	43
Paper I	43
Paper II	44
Paper III	45
Paper IV	46
Paper V	46
Results	49
Paper I	49
Sensitivity analyses	50
Paper II	52
Paper III	53
Paper IV	54
Sensitivity analyses	54
Paper V	58
Discussion	59
Methodological considerations	59
Study designs	59
Internal and external validity	61
Selection bias	61
Classification of exposure	62
Classification of outcome	63
Confounding	66
Certain methodological considerations in genetic studies	67
Generalizability	68
General Discussion	69
Lung function and risk of SAH	69
COPD and risk of stroke	70
Markers related to inflammation and SAH	72
Coding genetic variation and SAH	74
Conclusions	75
Populärvetenskaplig sammanfattning på svenska	77
Acknowledgements	79
References	81

# List of papers

This doctoral thesis is based on the following original papers, which are referred to in the text by their Roman numerals:

- I. Söderholm M, Zia E, Hedblad B, Engström G. Lung function as a risk factor for subarachnoid hemorrhage: a prospective population-based study. *Stroke* 2012; 43: 2598-2603. © American Heart Association, Inc. Wolters Kluwer Health Lippincott Williams & Wilkins
- II. Söderholm M, Zia E, Hedblad B, Engström G. Leucocyte count and incidence of subarachnoid haemorrhage. A prospective cohort study. *BMC Neurology* 2014; 14: 71. © Söderholm M, et al.
- III. Söderholm M, Engström G. Fibroblast growth factor 23 and incidence of subarachnoid haemorrhage. Nested case control study. *Stroke* 2015; 46: 3260-3262. © American Heart Association, Inc. Wolters Kluwer Health Lippincott Williams & Wilkins
- IV. Söderholm M, Inghammar M, Hedblad B, Egesten A, Engström G. Incidence of stroke and stroke subtypes in chronic obstructive pulmonary disease. *European Journal of Epidemiology* 2016; 31: 159-168. © Springer Science+Business Media Dordrecht
- V. Söderholm M, Zia E, Lorentzen E, Olsson M, Jern C, Melander O, Engström G. Coding genetic variation and subarachnoid hemorrhage: A population-based nested case control study. *Manuscript*.

The papers I-IV were reprinted with permission from the publishers.



# Abbreviations

CI	Confidence interval
COPD	Chronic obstructive pulmonary disease
CSF	Cerebrospinal fluid
CT	Computed tomography
FEV1	Forced expiratory volume in 1 second
FGF23	Fibroblast growth factor 23
FVC	Forced vital capacity
GWAS	Genome wide association study
HR	Hazard ratio
ICD	International classification of diseases
ICH	Intracerebral hemorrhage
MDC	Malmö Diet and Cancer study
MDC-CC	Malmö Diet and Cancer study Cardiovascular cohort
MONICA	Monitoring the trends and determinants in cardiovascular disease
MPP	Malmö Preventive Project
MRI	Magnet resonance imaging
OR	Odds ratio
QC	Quality control
PAR	Population attributable fraction
PPV	Positive predictive value
SAH	Subarachnoid hemorrhage
SMC	Smooth muscle cell
WHO	World Health Organization





# Introduction

John Snow (1813-1858) investigated the cholera epidemic in the quarters of Soho, London in 1854, thereby undertaking one of the first population-base studies that have been recorded<sup>1</sup>. At that time, cholera was considered an air-borne toxin or a result of bad air. Based on observations of which drinking water companies delivered water to the different water pumps and mapping of cholera cases, Snow could trace the source to a pump on Broad Street that was contaminated by sewage. John Snow is considered one of the “fathers of modern epidemiology”.

According to ‘A dictionary of epidemiology’ (Edited by Porta, M, 2014), epidemiology is

“The study of the occurrence and distribution of health-related events, states, and processes in specified populations, including the study of the determinants influencing such processes, and the application of this knowledge to control relevant health problems.”

Numerous epidemiological studies have been performed since John Snow reported from his pioneer studies of cholera. Epidemiological studies of stroke have provided much knowledge about risk factors for cerebrovascular diseases. However, the risk factors for subarachnoid hemorrhage (SAH), the least common type of stroke, have received less scientific attention. SAH is a devastating, sudden-onset vascular disease, with a 30-50% mortality rate and a high risk of severe disability in survivors. The age at SAH is generally lower compared to other stroke types, and SAH is more common in women than in men<sup>2</sup>. Because of the disease characteristics, SAH has a large socio-economic impact that is comparable to that of all ischemic strokes<sup>3</sup>.

Although the impact of SAH on both the individual and society is large, the causes of the disease remain poorly understood. Epidemiological studies of SAH may provide better knowledge about its risk factors, and methods to identify high risk groups. These studies can also help elucidating the disease-causing mechanisms.

In the present thesis, potential risk factors for SAH are studied in population-based settings, with the aim to improve the knowledge about risk groups and disease mechanisms.

## General aspects of stroke epidemiology

Subarachnoid hemorrhage is the least common type of stroke, and accounts for around 5% of all strokes. SAH is a bleeding in the subarachnoid space around the brain, in most cases caused by the rupture of an arterial intracranial aneurysm<sup>2</sup>. Ischemic stroke and intracerebral hemorrhage (ICH) are the other main types, accounting for about 85 and 10 % of all strokes, respectively<sup>4</sup>. Ischemic stroke is caused by obstruction of the blood flow, for example by a distal blood clot or focal atherosclerosis with thrombus formation, leading to infarction in the affected part of the brain. Intracerebral hemorrhage occurs from rupture of small intraparenchymal arteries and results in bleeding within the brain. Stroke is defined by the World Health Organization (WHO) as a sudden onset of focal or global neurological deficits, lasting for more than 24 hours or until death, without any evident non-vascular cause<sup>5</sup>. Global instead of focal loss of neurological function is most often seen in the case of SAH<sup>5</sup>.

Overall, stroke is one of the most common causes of death worldwide, and a major cause of severe acquired disability in the adult population<sup>6, 7</sup>. Stroke incidence increases exponentially with age, although stroke occurs in all age groups and the health burden of stroke is substantial also in the age below 65 years<sup>8</sup>. Men have a higher risk of stroke than women, whereas women are affected at higher age and more often have severe strokes<sup>9, 10</sup>. In the latest global burden of disease study, stroke ranked second of individual causes of death after ischemic heart disease and third with regard to disability adjusted life years, i.e. the sum of life years lost to disease and years living with disability<sup>6, 7</sup>.

During the last 50 years, incidence of stroke has tended to decrease or be stable in high income countries and increase in low and middle income countries<sup>4, 8</sup>. The causes of this pattern likely reflect life style changes, an increased prevalence of modifiable risk factors, such as smoking, hypertension and diabetes, and a higher life expectancy, especially in low and middle income countries<sup>4</sup>. In Sweden, the estimated number of individuals affected by strokes each year is between 27,000 and 35,000, of whom between 1000 and 2000 have a subarachnoid hemorrhage<sup>9</sup>. Between 1998 and 2014, 21,000 to 28,000 individuals per year were admitted to hospital because of a first-ever stroke<sup>11</sup>. The mortality from stroke has declined over the last years, in Sweden, and in other Western countries, perhaps as a result of better stroke care, whereas the incidence rate is rather stable<sup>9, 12</sup>. With a growing, and older population, the absolute number of strokes will increase worldwide, as will survivors of stroke, and the burden of disability caused by stroke<sup>8</sup>. Thus, strategies of prevention and treatment of stroke are of great importance and will continue to be a challenge in the nearest future.

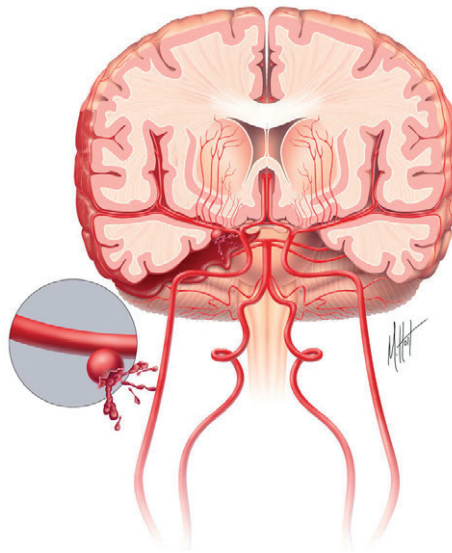
The epidemiologic characteristics of the three stroke types have both similarities and differences. For instance, SAH and ICH are associated with very high mortality rates (20–50 %) <sup>13, 14</sup>, whereas ischemic stroke has lower mortality, around 10-15%, within 30 days <sup>4</sup>. SAH occurs at a relatively lower age, compared to ischemic stroke and ICH, and is unlike most other vascular diseases more common in women than in men <sup>15</sup>. Some risk factors are shared for all stroke subtypes (e.g. hypertension and age), whereas other major vascular risk factors (e.g. smoking and diabetes) may be differentially related to ischemic stroke, ICH and SAH <sup>16-18</sup>.

In epidemiological studies of common conditions, such as ischemic heart disease, and ischemic stroke, a high number of risk determinants have been identified. This research has provided important knowledge about the trends of disease and disease-related disability, risk factors, as well as risk assessment tools to identify high risk groups, and strategies of intervention to reduce the burden of disease. A lower number of epidemiological studies have evaluated risk factors for less common vascular diseases, including ICH and SAH. This may be due to the lower incidences of these diseases, which makes them more difficult to study in prospective population-based studies, and the high mortality causing difficulties with inclusion of patients and assessment of risk factors. Because SAH has some clinical and etiological characteristics that differ from other stroke subtypes, it is not always included in studies of the other stroke types and sometimes considered an own category separate from ischemic stroke and ICH.

The focus of this doctoral thesis is the study of potential risk factors for subarachnoid hemorrhage in the general population. The other stroke types are also included in some parts. Therefore, the background description will put most emphasis on SAH. Ischemic stroke and ICH will be only briefly described.

# Subarachnoid hemorrhage

Bleeding into the subarachnoid space, subarachnoid hemorrhage (SAH), occurs when one of the arteries along the base of the brain, in the subarachnoid space, ruptures (**Figure 1**). The subarachnoid space, below the arachnoid mater and outside of the pia mater (the soft meninges), encloses the large vessels of the brain, it contains cerebrospinal fluid and is in connection with the ventricle system. Spontaneous SAH is most often caused by a rupture of an arterial aneurysm, which is a balloon-like focal dilation of the blood vessel<sup>2</sup>. When SAH occurs, blood spreads along the surface of the brain, exciting the nerves in the pia and arachnoid mater (lepto-meningeal nerves), and causing the intracranial pressure to rise. For the patient, SAH will commonly cause severe, sudden-onset headache reaching the maximum intensity instantaneously or within a few minutes (thunderclap headache), nausea and vomiting, deteriorating consciousness and sometimes, focal neurological deficits<sup>19</sup>. SAH is one of the deadliest vascular emergencies and results in death in 30-50% of patients<sup>13</sup>. However, the clinical picture at SAH varies greatly, and between one third and half of SAH patients present with thunderclap headache as the only symptom and are alert and have normal neurological examination<sup>19</sup>. These cases are difficult to clinically distinguish from benign thunderclap headache<sup>19</sup>. Patients with a small initial bleed from an aneurysm, so called “warning leak”, are at high risk of a new bleed and urgent handling of these cases is necessary<sup>2</sup>.



**Figure 1.** Subarachnoid haemorrhage from ruptured aneurysm. (Copyright Matthew Holt. Reprinted with permission).

## Diagnosis

Computed tomography (CT) of the brain has a high sensitivity for diagnosing SAH. Because SAH may still be missed on CT, especially if it is performed a long time after onset of symptoms, lumbar puncture is normally performed to rule out SAH if CT is negative. The diagnosis of SAH is strongly supported in the presence of red blood cells in CSF ( $>2000$  erythrocytes/mm<sup>3</sup>) together with confirmation of xanthochromia in spectrophotometry after centrifugation (i.e. yellow-stained fluid with increased absorbance at pre-specified wavelengths). Spectrophotometry is based on detecting bilirubin, which is formed only in vivo from breakdown of hemoglobin, and bilirubin in the CSF is thus a certain sign of SAH and not of a traumatic tap. Lumbar puncture should be done earliest 12 hours after onset of headache, because it will take 6-12 hours for bilirubin to form<sup>20, 21</sup>.

Interestingly, recent studies have shown that CT has a near 100% sensitivity for ruling out SAH if it is performed in the first 6 hours after onset of headache<sup>22, 23</sup>. According to new guidelines, lumbar puncture could, and should therefore be withheld in the investigation of some of the patients with acute headache and negative CT within 6 hours, if they also fulfill a number of other criteria<sup>24</sup>.

If SAH is diagnosed, further investigation with CT, magnetic resonance imaging (MRI), or digital subtraction angiography should be performed to detect intracranial aneurysms causing the SAH<sup>25</sup>.

## Causes of SAH

The most common cause of non-traumatic SAH is a ruptured arterial intracranial aneurysm, and aneurysmal SAH accounts for some 75-85% of cases<sup>26-28</sup>. Non-aneurysmal perimesencephalic SAH is a benign form of SAH with a good prognosis. Causes of non-aneurysmal perimesencephalic SAH are unknown but bleeding from veins has been suggested<sup>29</sup>. Rare causes of SAH include inflammatory conditions (for example mycosis, Borreliosis and auto-immune vasculitis), which can cause vessel weakness, sometimes aneurysms, and rupture, other structural vascular abnormalities (e.g. arterial dissection, cavernous angiomas, arterio-venous malformations, tumors), coagulopathies and certain drugs (e.g. cocaine and amphetamine).

## Intracranial aneurysms

Aneurysms are localized balloon-like dilations of blood vessels. The most common form of intracranial arterial aneurysm is the saccular or berry aneurysm attached by

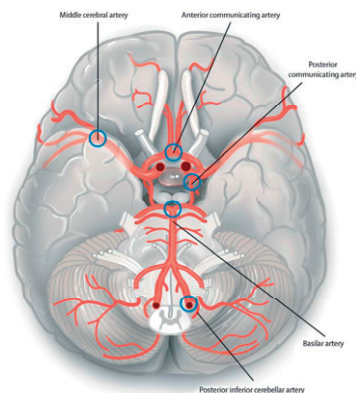


a neck. They are commonly located at arterial branching sites on the large arteries of the circle of Willis or on one of its large branches (35% anterior cerebral artery/anterior communicating artery, 30% internal carotid artery, 20% middle cerebral artery, 10% posterior communicating artery, and 5% vertebro-basilar arteries, **Figure 2**)<sup>30</sup>. Aneurysms are formed during life and the prevalence increases with age<sup>31</sup>. The overall prevalence of intracranial aneurysm in the adult general population has been estimated to 3.2% (95% CI 2-5%) in a recent meta-analysis, with higher figures in women<sup>31</sup>. About 30% have multiple aneurysms<sup>31</sup>.

Unruptured aneurysms are often detected incidentally at investigation for different diseases, e.g. cerebral infarction, headache or transient neurological symptoms. The natural history of unruptured aneurysms is, however, poorly elucidated. Most unruptured aneurysms are small (<5 mm) and the risk of rupture is strongly related to the size of the aneurysm: A meta-analysis found rupture rates of 0.4% per year in aneurysms <7 mm<sup>32</sup>, another prior meta-analysis 0.7% per year for aneurysm <10mm, and 4% for size of >10 mm<sup>33</sup>. Besides size of the aneurysm, location in the posterior circulation, higher age, female sex, smoking, previous history of SAH, and having symptoms from the aneurysm are also factors associated with increased risk of rupture<sup>32, 34</sup>. Although small aneurysms have a low rupture risk, they are the most common, and therefore causes the largest amount of SAHs.

Thus, the risk assessment of which aneurysms will cause SAH is difficult. Present decision guidelines about when to offer intervention for unruptured aneurysm, and when instead following up with repeated imaging, are mostly based on aneurysm size, location and risks of the procedure, but patient's risk factors should also be considered<sup>25</sup>. More studies are needed to find predictive factors for risk of SAH in persons with known aneurysms, so that preventive treatment with endovascular or surgical intervention may be offered to those that will have the largest benefit.

**Figure 2.** The base of the brain and the large arteries with the most common locations of intracranial aneurysms marked by rings. Reprinted from The Lancet, Vol 369, Van Gijn, J., Kerr, R., Rinkel, G., Subarachnoid haemorrhage, p 306-318, Copyright (2007), with permission from Elsevier.



## Pathophysiology of aneurysms

The pathogenesis of aneurysm formation and rupture is not fully understood. Histological studies including immunohistochemistry, gene expression and experimental studies of animal models have outlined some characteristics and suggested mechanisms of intracranial aneurysm formation and rupture.

The intracranial arteries are composed of the innermost intimal layer containing the endothelium. The intima is surrounded by the internal elastic lamina, the media, with elastic fibers and smooth muscle cells (SMCs), and the adventitia with fibroblasts and collagen. The intracranial arteries lack the external elastic lamina of extracranial arteries. The stability and strength of the vessel wall is given of the extracellular matrix (ECM), which is built up of a fine network of collagen, elastin and other elastic fibers, surrounded by proteoglycans and glycoproteins<sup>35</sup>.

In principle, aneurysms can form and rupture when the shear stress from blood pressure on the vessel walls exceeds the tensile strength of the wall. In animal models induced hypertension, disturbed hemodynamics and impaired collagen integrity of the vessel wall are needed for aneurysm formation<sup>36</sup>. Aneurysms form at bifurcations where shear stress is highest. They are probably the result of increased shear stress in combination with a weakened vessel wall due to for example inflammatory activation at the surface<sup>36</sup>.

In histological studies, aneurysm tissue is characterized by lack of internal elastic lamina, loss of endothelium, myointimal hyperplasia from mural cell proliferation and migration into the intima, and also, decreased number, or loss of mural cells<sup>37-39</sup>. Infiltration and activation of inflammatory cells, predominantly macrophages, T-cells and neutrophils, and activation of the humoral immune response is also observed<sup>36</sup>, as is increased proteolytic activity of collagenases, for example matrix metalloproteinase 2 and 9<sup>40, 41</sup>. The inflammatory infiltration is most pronounced in ruptured compared to unruptured aneurysms<sup>38</sup>. Atherosclerosis is also frequently observed in intracranial aneurysms, with 50% showing advanced atherosclerosis, sometimes also calcification<sup>42</sup>. An immune-reaction to cholesterol particles is seen in aneurysm walls during development and degeneration, although it is not known if the processes of atherosclerosis are a primary cause of aneurysm formation and/or rupture, or is only a secondary phenomenon<sup>35, 36, 42</sup>.

The rupture of aneurysms is associated with morphological changes that are believed to precede the rupture<sup>38</sup>. These include increased inflammatory infiltration, loss of SMCs and fibroblasts, increased degradation of ECM, loss of endothelium and intimal thrombus formation. These are features usually associated with vascular remodeling and repair after injury but for unknown reasons these repair mechanisms will in some cases instead increase the degradation and lead to rupture<sup>38</sup>. Collagenases as well as molecules involved in inflammatory cell infiltration, and

oxidative stress, are overexpressed in gene expression analyses of ruptured compared to unruptured aneurysms<sup>43</sup>. However, it is still not known if the inflammatory reactions seen in aneurysms are primary causes of formation and rupture, or merely an epi-phenomenon<sup>36</sup>.

## **Incidence and prognosis**

SAH accounts for between 1 and 10 % of all strokes and the overall worldwide age-standardized incidence rate is between 7 and 13 per 100,000 person years in most countries<sup>15</sup>. For comparison, age-standardized incidence rates of ischemic stroke and ICH are mostly between 50-300 and 10-30 per 100,000 person-years, respectively<sup>4, 44-46</sup>. Incidence rates for SAH have been higher in Finland and Japan in many studies, and lower in Central and South America and China<sup>15, 47</sup>, although it is not known if this is due to methodological reasons of case finding, or true incidence differences<sup>48, 49</sup>. The WHO MONICA study showed clear regional differences with highest rates in northern Sweden and Finland, using the same definition of SAH in all populations<sup>47</sup>.

In Sweden, between 1000 and 1400 patients were admitted to hospital with a diagnosis of SAH per year from 1998 to 2014<sup>11</sup>. There is a regional difference in incidence of SAH in Sweden with highest rates in the north (15 per 100,000)<sup>50, 51</sup> and lower in the south (10 per 100,000 person years)<sup>27, 47</sup>. Incidence rates declined in men but not in women between 1985 and 2000 in northern Sweden<sup>50</sup>. However, incidence rates of SAH have been stable over the last decades in other studies<sup>15</sup>.

Another distinct feature of SAH is the lower age at onset, compared to the other types of stroke. The mean age is around 60 years<sup>15, 27</sup>, compared to about 75 years for ischemic stroke and ICH<sup>9, 44</sup>. Higher age is, however, also a risk factor for SAH, and SAH is uncommon under the age of 30 years. The incidence rate pattern seems more plateau-like between the ages of 55 and 85 years (**Table 1**). Data in the elderly (>85 years) is limited<sup>15, 50</sup>.

The case-fatality rate after SAH is high: The proportion who had died 28 days after SAH was between 30 and 60 % in European studies, as meta-analyzed by Nieuwkamp et al<sup>13</sup>. The WHO MONICA study showed mortality of 42% worldwide, restricted to the younger age group (25-64 years)<sup>47</sup>. In Sweden, studies have shown 36%, 45% and 64% 30 day mortality in 1983-1985<sup>46</sup>, 1985-2000<sup>50</sup>, and 1999-2000<sup>44</sup>, respectively. SAH is a known cause of sudden deaths and 10-20% of SAH patients die suddenly before reaching the hospital or in the emergency department<sup>50, 52, 53</sup>.

In a study from Sweden, including all SAHs between 1987 and 2002 (n=17,000) registered in the national Inpatient and Causes of death Registers, 32% had died one

month after SAH, and 35% three months after<sup>54</sup>. A tendency towards decreased mortality was observed 1987-2002<sup>54</sup>, which is in line with the results from the MONICA study in northern Sweden (1985-2000)<sup>50</sup>.

**Table 1.** Age specific incidence rates per 100 000 person years for SAH in southern Sweden 1996 (Nilsson et al.<sup>27</sup>) and from meta-analysis of 20 studies (De Rooij et al.)<sup>15</sup>.

Age, years	southern Sweden			Meta-analysis
	Men	Women	Total	Total
0-14	0	0	0	2.0 <25 years
15-24	0	1.4	0.7	2.0 <25 years
25-34	0	2.7	1.3	7.7
35-44	4.1	8.4	6.2	10.5
45-54	13.2	22.2	17.6	19.5
55-64	13.8	22.2	18.0	24.8
65-74	16.5	25.3	21.2	25.4
75-84	12.8	27.1	21.2	26.2
>85	12.3	27.8	23.0	31.3

### *Symptoms after SAH*

In patients surviving SAH, cognitive impairment is common, including deficits of memory, executive function and language. 4-12 % have deficits of activities of daily living, whereas 44-95% have deficits in so called instrumental activities of daily living, for example, dealing with finances, shopping, housekeeping. Around 40% of SAH survivors that worked before the SAH event are not able to return to their previous work after SAH<sup>55</sup>.

Because of the lower age at onset, and the higher mortality compared to other strokes, SAH is estimated to cause one third (27 %) of potential life years lost before the age of 65 due to all cerebrovascular diseases. This is comparable with that of ischemic stroke (38%) or ICH (34%)<sup>56</sup>. Hence, although the relatively low incidence compared to other strokes, the morbidity, mortality and socioeconomic burden of SAH is substantial.

## **Risk factors**

Higher age, female sex, hypertension, smoking, high alcohol intake and family history are established risk factors for SAH<sup>17, 25</sup>. For other potential risk factors, results are less consistent.

### *Female sex and hormonal factors*

SAH is around 1.25-1.5 times more common in women than in men, thus contrasting other types of stroke, and most other vascular diseases, for which male sex usually is a risk factor<sup>15, 17, 27</sup>. The sex difference in SAH risk is most pronounced after the age of 60 years<sup>15</sup>. The reasons for the difference between men and women are not known. It has been hypothesized that postmenopausal hormonal changes may be part of the explanation. Interestingly, hormonal replacement therapy has been associated with decreased risk of SAH<sup>17, 57</sup>. Oral contraceptives, on the other hand, have been suggested to increase the risk of SAH in a meta-analysis<sup>3</sup>. Recently, it has been proposed that smoking and female sex may have a synergistic effect on SAH risk and that this may be part of the explanation of the observed higher incidence in women<sup>58</sup>.

### *Hypertension*

Hypertension has been associated with a 2.5 fold risk increase in a meta-analysis, and estimates were similar and consistent in both longitudinal and case control studies<sup>17</sup>. There was, however, no uniform definition of hypertension in the meta-analysis. In many studies, the actual levels of blood pressure are not recorded, and blood pressure are sometimes measured after SAH, which may result in an overestimation of the risk associated with hypertension because blood pressure may be raised as an effect of SAH. Because of these shortcomings, it is not well studied if there is a dose-response relationship between blood pressure levels and SAH. A Norwegian cohort study found that systolic blood pressure of 130-160 mmHg doubled the risk of SAH, and over 160 mmHg increased the risk by 3-fold, compared to having blood pressure <130 mmHg<sup>59</sup>. A Finnish cohort study found a similar trend with most clearly increased risk in subjects with systolic blood pressure >160 mmHg, whereas 'diagnosed hypertension' had a hazard ratio of 1.5<sup>60</sup>. One study used a strict definition of hypertension; confirmation of hypertension by a physician before SAH, or evidence of target organ damage of hypertension. In that study, no association between hypertension and SAH was found, and the authors questioned hypertension as a risk factor<sup>61</sup>.

### *Smoking*

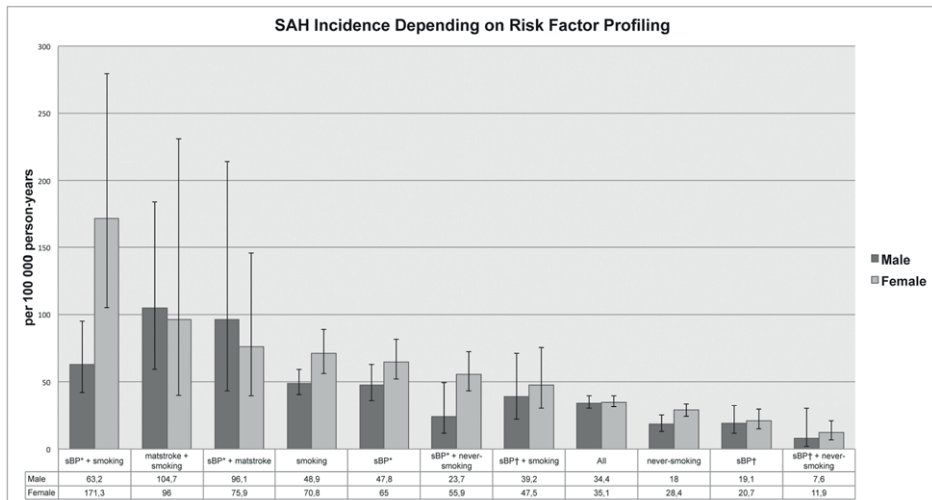
Smoking is a strong and consistent risk factor for SAH. In a population-based Australian case control study, current smoking conferred an odds ratio (OR) of 5



compared to never smokers, and there was a dose-response relationship between amount of smoked cigarettes and SAH risk<sup>62</sup>. Likewise, in prospective cohort studies, current smoking has been strongly associated with incident SAH, with risk ratios between 3 and 6 compared to never smokers<sup>59, 60, 63, 64</sup> and of around 2 compared to non-smokers (i.e. never and former smokers)<sup>63</sup>. Estimates of the population attributable fraction (PAR) of smoking have been about 20-40%<sup>62, 65, 66</sup>. The high risk associated with current smoking decreases substantially after quitting smoking, meaning that smoking cessation is likely of great value to reduce the incidence of SAH. However, in most studies, former smokers are still at an increased risk compared to never smokers<sup>17, 59, 60, 62, 64, 67</sup> (risk ratio of around 2 in meta-analysis<sup>17</sup>), although the time since stopping smoking, and years as a smoker differs. In the American Nurses' Health Study of relatively young women, the increased risk in former smokers seemed to persist for at least five years after cessation<sup>64</sup>. For comparison, the excess risk of ischemic stroke in former smokers disappeared after two years<sup>64</sup>. In the Australian study, the excess risk of SAH from former smoking diminished after one year. Exposure to smoke at home was not associated with SAH risk<sup>62</sup>.

Two recent prospective population-based studies have shown substantially higher effect estimates for smoking on SAH risk in women compared to men<sup>58, 60</sup>. This observation could be part of the explanation for the sex difference in SAH risk. **Figure 3** shows the joint effect of risk factors for SAH from the 'FINRISK' study by Korja et al<sup>60</sup>. However, other results regarding the sex difference for smoking have been inconsistent<sup>17</sup>. Studies of interactions require large samples. A Swedish collaboration is currently meta-analyzing data from a large number of cohort studies, including >2500 SAH events. Preliminary results show a higher effect of smoking in women (OR 2.6) than in men (OR 1.7,  $p$  for interaction <0.01)<sup>68</sup>.

The biological effects of smoking on SAH risk are not fully understood. Mechanisms might be inflammatory activation in arterial walls, endothelial dysfunction, or inhibition of anti-proteases causing increased degradation of ECM<sup>36</sup><sup>69</sup>. Increased collagenase activity and ECM degradation could be a primary cause of smoking, or a result of the macrophage activation induced by smoking<sup>36</sup>. Alpha-1-antitrypsin, an important anti-protease is inhibited by smoking<sup>69</sup>. Smoking may also cause blood pressure peaks with sudden increases of shear stress, which could be of importance for SAH. Smokers have generally, however, lower blood pressure than non-smokers<sup>70</sup>.



**Figure 3.** Rates of SAH by combination of risk factors in men and women. From Korja et al., Plos One, 2013<sup>60</sup>. sBP, high systolic blood pressure; matstroke, stroke in mother.

### *Alcohol drinking*

Heavy alcohol drinking has been a risk factor for SAH in several studies<sup>17, 60</sup>. Intake of >150 g/week was associated with a two-fold increased risk compared to those drinking no alcohol in a meta-analysis<sup>17</sup>. In a large Finnish cohort study, drinking  $\geq 87$  grams of alcohol per week was associated with a two-fold increased risk compared to abstainers<sup>60</sup>. However, results for high alcohol intake in the Australian and Asian populations have been inconsistent<sup>71-73</sup>. Because the majority of high alcohol consumers are also smokers, often heavy smokers, there might be residual confounding from smoking explaining part of the association between alcohol intake and SAH, and it is difficult to do subgroup analysis of alcohol and SAH risk in never smokers<sup>60, 66</sup>.

### *Metabolic factors, diet and physical activity*

Hypercholesterolemia, diabetes mellitus, and overweight have been suggested as protective factors for SAH in a meta-analysis<sup>17</sup>. However, results must be considered uncertain<sup>17</sup>. Results for cholesterol levels and SAH risk are conflicting and reasons for a negative association are unclear. Many case control studies measured cholesterol in the acute phase of SAH, or did not adjust for other risk factors<sup>74</sup>. Prospective studies, have instead reported positive associations between cholesterol and risk of SAH, especially in men<sup>60, 74-76</sup>. Diabetes may cause vessel stiffness and an inverse association is seen for aortic aneurysms. If this is the case for SAH is not known. High dietary intakes of fat and salt have been associated with risk of SAH<sup>77</sup>. Reduced-fat milk, fruits, soy and green tea have instead been

associated with a lower risk<sup>77, 78</sup>. High intensity physical activity may trigger SAH<sup>79</sup>. A high level of long-term physical activity is instead associated with a decreased risk<sup>17, 80</sup>.

### *Circulating biomarkers*

Few studies have examined circulating biomarkers in association to future risk of SAH. Two studies with small numbers of SAH cases (n=29 and 33, respectively) examined markers of systemic inflammation, hsCRP<sup>81</sup> and fibrinogen<sup>82</sup>, but none of these were significantly associated with SAH incidence.

Fibroblast growth factor 23 is a hormone secreted from osteoblasts involved in calcium homeostasis. It is increased in chronic kidney failure and high levels have been associated with inflammation and mortality, respectively, in patients with chronic kidney disease<sup>83, 84</sup>. One prospective population-based study found that high levels of FGF23 was associated with incidence of ICH, but not with ischemic stroke. SAH was not evaluated because of a low number of cases<sup>85</sup>. Another study found an association between FGF23 and cardioembolic ischemic stroke<sup>86</sup>. We hypothesized that FGF23 levels are associated with increased risk of incident SAH because of a possible relation to both ICH and inflammatory markers. Also, one genetic risk locus for intracranial aneurysms is located close to the gene encoding Klotho. Klotho is a regulator of fibroblast growth factors and vascular calcification and vascular aging<sup>87 88</sup>.

### *Shortcomings of prior risk factor studies for SAH*

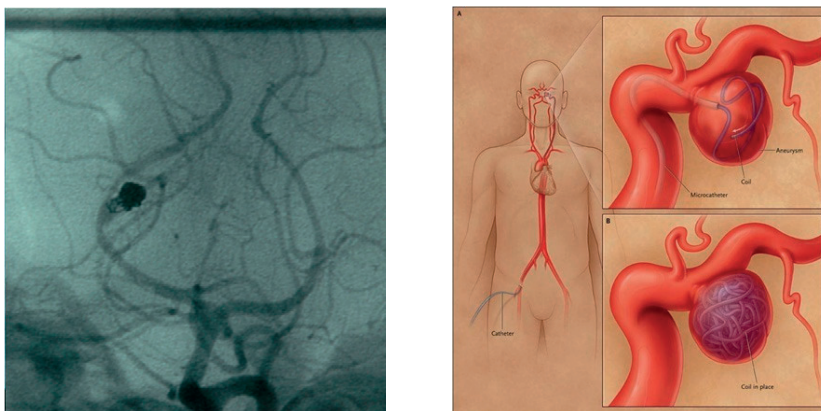
Generally, several of the studies of risk factors for SAH are small and few have prospective cohort designs, i.e. exposure is measured before disease. Many results come from case control studies. Many of these assessed risk factors in the acute phase of SAH or after SAH, which could induce recall bias and lead to findings of reverse causation. Controls come from different settings; for example, in some studies hospital controls have been used. Moreover, fatal cases have been excluded in some studies and only few studies include also cases of SAH dying before reaching the hospital. In other studies, proxy information has been used for fatal and severely disabled cases. The heterogeneous methodologies of risk factor studies make results from meta-analysis difficult to interpret<sup>74, 89</sup>.

## **Complications and Treatment**

Rebleeding and delayed cerebral ischemia are common and serious complications to SAH associated with worse outcome<sup>90</sup>. Delayed cerebral ischemia most often develops 4 to 14 days after SAH and affects at least 10-30% of patients, but probably a clearly higher proportion<sup>22</sup>. It causes focal or global loss of neurological function.

One cause of ischemia is vasospasm, which is in turn primarily due to blood surrounding the large vessels, although the mechanisms for vasospasm are complicated<sup>22</sup>. Other factors, for example global ischemia during the course of SAH also seem to be of importance for delayed cerebral ischemia<sup>91</sup>. Initial loss of consciousness and large amounts of blood, respectively, associate with high risk of delayed cerebral ischemia. Prevention of cerebral vasospasm by the use of calcium-antagonists reduces cerebral ischemia and have a favorable effect on outcome after SAH<sup>92</sup>. Hydrocephalus due to blood obstructing the flow of CSF in the ventricular system is also a complication after SAH<sup>2</sup>.

To prevent rebleeding after aneurysmal SAH, occlusion of the aneurysm should be done as soon as possible<sup>25</sup>. Serious rebleeding often occurs already shortly after the initial bleed (around 15% in the first 24 hours<sup>93</sup>). Endovascular coiling, i.e. occluding the aneurysm with platinum spirals (coils), or microsurgical neck clipping in open surgery, is performed to occlude the aneurysm. Randomized trials have shown reduced rate of death and dependency in patients undergoing endovascular coiling compared to microsurgical clipping<sup>94</sup>. Endovascular intervention is therefore preferable although clipping is still the method of choice for some patients, because the location and morphology (e.g. large, broad-based) of the aneurysm make them unsuitable for endovascular techniques<sup>2, 94</sup>. Antifibrinolytic drugs (e.g. tranexamic acid) reduce the risk of rebleeding<sup>95</sup>, but are not clearly favorable on overall outcome after SAH, maybe because they could also increase the risk of cerebral ischemia and venous thrombosis<sup>96</sup>. However, short-term treatment with antifibrinolytic drugs before aneurysm occlusion is done is probably beneficial for both rebleeding and long-term outcome<sup>95, 97</sup>.



**Figure 4.** Coiled aneurysm of the posterior cerebral artery with a small residual aneurysm sac. From Pillai et al. Journal of Medical Case Reports 2007 1:168. (Creative commons attribution license.)

## Family history and genetic studies of SAH

First-degree relatives of SAH patients are at two to seven-fold increased risk of having a SAH<sup>98-101</sup>, and in those with two or more first-degree relatives the risk is even stronger<sup>100</sup>. However, only some 5% of patients with SAH have a first-degree relative with SAH<sup>100</sup>. To some part, the excess risk in first-degree relatives is explained by shared acquired risk factors, such as smoking, hypertension and high alcohol intake, which are common and account for a large part of SAH<sup>65, 66, 102</sup>. However, studies also support a genetic influence on risk of SAH and intracranial aneurysms<sup>66</sup>. A large twin study reported a heritability, i.e. the proportion of disease variance explained by genetic factors, of 41%, thus suggesting that other risk factors account for the largest part of SAH but genetic factors have moderate influence<sup>103</sup>. This estimate is similar to other so called complex vascular diseases, such as ischemic stroke (17-38%)<sup>104-106</sup> and coronary artery disease (40-50%)<sup>107</sup>, and cancers like prostate (42%) and breast (27%) cancer<sup>108</sup>.

Some genome-wide association studies (GWAS) have identified and replicated several loci related to intracranial aneurysm. Whereas the GWAS approach is hypothesis-free and investigates a high number of variants, often single nucleotide polymorphisms, throughout the genome, candidate gene studies are based on hypotheses about which genes that are important for the studied disease and studies only variants in that genes<sup>109, 110</sup>. Candidate gene studies have also found some risk variants for aneurysms and SAH, but mostly they have not been replicated, as has been the case also for other diseases<sup>111</sup>. Generally, GWA studies have been most successful in identifying risk loci for complex diseases, compared to candidate gene studies<sup>110</sup>. A reason for this is that it is difficult to formulate hypotheses because the links between the genetic variation and the biological processes in disease are very complex.

The most robust associations from GWAS for intracranial aneurysms have been found for polymorphisms in or near the cyclin dependent kinase inhibitor 2A and B (*CDKN2A/CDKN2B*) genes on chromosome 9p21, the SRY Box-17 (*SOX17*) gene on chromosome 8q12, and the endothelin receptor type A (*EDNRA*) gene on chromosome 4q31<sup>87, 111, 112</sup>. The chromosome 9p21 locus has been associated with aneurysmal SAH in a south Swedish sample<sup>113</sup>, and is also related to coronary heart disease, large-artery atherosclerotic ischemic stroke and abdominal aortic aneurysm<sup>114</sup>. Both the chromosome 8 and 9 loci probably confer especially high risks of intracranial aneurysms/SAH in smokers<sup>115, 116</sup>. Discovered and replicated loci from GWA studies for intracranial aneurysms are listed in **Table 2**.



**Table 2.** Loci associated with intracranial aneurysms in genome-wide association studies.

Locus	Nearest genes
9p21.3 <sup>87, 114, 115, 117</sup>	<i>CDKNA2BAS</i>
8q11.23 <sup>87, 115, 117, 118</sup>	<i>SOX17</i>
8q12.1 <sup>87, 118</sup>	<i>SOX17</i>
4q31.23 <sup>112</sup>	<i>EDNRA</i>
10q24.32 <sup>87</sup>	<i>CNNM2</i>
12q22 <sup>87, 112</sup>	<i>FGD6</i>
13q13.1 <sup>87</sup>	<i>KL/STARD13</i>
18q11.2 <sup>87</sup>	<i>RBBP8</i>
2q33.1 <sup>117</sup>	<i>ANKRD4/HSPD1</i>
7p21.1 <sup>118</sup>	<i>TWIST1/FERD3L/HDAC9</i>
20p12.1 <sup>112</sup>	<i>RRBP1</i>

The pathophysiological mechanisms by which the discovered risk variants influence aneurysms and SAH are not fully understood, but there are some suggested functions. Two tumor suppressor genes, the *CDKNA2A* and *CDKNA2B*, that regulate cell cycle and cell proliferation, are located in this region. Some of the 9p21 polymorphisms are located in the *CDKN2B antisense RNA1 (CDKN2BAS)*, which has probably a regulatory function<sup>119</sup>. Reduced expression of *CDKN2B* is seen in blood leukocytes, vascular SMCs, and atherosclerotic plaques of subjects carrying the 9p21 risk variants<sup>120</sup>. The effect of the 9p21 variants on atherosclerosis and aneurysm may be caused by increased vascular SMC proliferation, and in some cases, increased apoptosis of SMCs<sup>121, 122</sup>. Increased apoptosis of SMCs causes weakening of the arterial wall and is an important characteristic in intracranial aneurysm histopathology<sup>123</sup>.

The SOX7, SOX17 and SOX18 proteins (SoxF family) are transcription factors functionally closely related to each other, and with partially redundant biological functions. They have important roles in vascular endothelial development and maintenance. The involvement of SOX17 in aneurysm development and SAH is supported by a recent experimental study, in which *SOX17* knock-out mice had more cerebral arterial wall weakening, dilation of arteries, increased aneurysm development and hemorrhage occurrence, due to endothelial pathology, compared to their littermates<sup>124</sup>.

The EDNRA variant associated with intracranial aneurysms is located directly upstreams of the gene in a regulatory region. The endothelin receptor has well-characterized functions on the regulation of vascular tonus, and proliferation of

endothelial and SMCs, and may thus be directly involved in aneurysm formation, and also atherosclerosis<sup>112</sup>.

Candidate gene studies have suggested the influence of genes related to ECM integrity, inflammation and endothelial function<sup>111</sup>. A meta-analysis found evidence of association for six loci from candidate gene studies with intracranial aneurysms or SAH, including polymorphisms in the serine protease encoding gene *SERPINA3*, the collagen type 1 and 3 genes *COL1A2* and *COL3A1*, the proteoglycan genes versican and heparin sulfate proteoglycan 2<sup>111</sup>. Some of these associations have been restricted to one ethnicity, and need further confirmation<sup>111</sup>.

### **Further studies of genetic determinants**

Only a small part of the heritability for intracranial aneurysms can be explained by known loci from recent GWA studies. No genome-wide study of SAH has been performed, to our best knowledge. Further studies with different approaches are needed to find the specific variants associated with intracranial aneurysms and SAH risk. Studying variation within the protein-coding part of the genome (exome variation) may be one way of finding new risk loci for SAH. Commercial so called exome arrays are available and genotype a large number of coding variants throughout the exomes of all genes (exome-wide)<sup>125</sup>. Coding non-synonymous variants can cause missense (a change of an amino acid), nonsense (introduction of a stop codon) or splice site mutations (changes the splice site), and although these variants are mostly uncommon their effect may be high on disease risk because of the potential direct effects on protein functions<sup>109</sup>.

## **Risk factors for ischemic stroke and ICH**

As for SAH, smoking is an important risk factor also for ischemic stroke, whereas results are uncertain for ICH<sup>16, 18, 126</sup>. In the 'INTERSROKE' study, which is a large international case control study of risk factors for ischemic stroke and ICH, current smoking conferred an odds ratio of 1.9 for ischemic stroke in the whole study (population attributable fraction of 46%) and 3.3 restricted to cases and controls from Western Europe, North America and Australia. Smoking was not a risk factor for ICH in the whole sample (OR 1.14, 99% CI 0.95-1.14) or in any region<sup>18</sup>. A study from the MPP<sup>16</sup> and a systematic review<sup>126</sup> are in line with these results. However smoking was a risk factor for lobar ICH in the MPP study<sup>16</sup>.

Hypertension is the risk factor accounting for most cases of ischemic stroke and ICH, with ORs of 2.8 for ischemic stroke (PAR=45%) and 4.1 for ICH (PAR=56%)

in the 'INTERSTROKE' study, defining hypertension as blood pressure  $\geq 140/90$  or self-reported hypertension<sup>18, 127</sup>.

In addition to smoking and hypertension, influential risk factors for ischemic stroke include diabetes, hypercholesterolemia, overweight, physical inactivity, heavy alcohol consumption, psychosocial factors, and atrial fibrillation. For ICH, risk factors in addition to hypertension, include overweight, high alcohol intake and psychosocial stress<sup>16, 18, 126</sup>.

Several of the risk factors for SAH are also risk factors for other vascular diseases. This is reflected by results showing that patients surviving SAH have an increased long-term risk of mortality from vascular diseases<sup>54</sup>, in particular other cerebrovascular events<sup>52</sup>, compared to the general population.

## Lung function, COPD and stroke

Reduced lung function assessed in spirometry, expressed as either a low forced expiratory volume in 1 second (FEV1) or a low forced vital capacity (FVC), has been associated with increased incidence of several cardiovascular diseases<sup>128, 129</sup>, and with cardiovascular and all-cause mortality<sup>130, 131</sup>. The associations have persisted also in studies of life-long never smokers<sup>131, 132</sup>.

For stroke as outcome, studies have shown that reduced lung function reflected by a low FEV1 and FVC, respectively, is associated with an increased incidence of ischemic stroke, also in never smokers<sup>133, 134</sup>. Low FEV1 is associated with stroke from all causes, especially fatal stroke, in the general population independently of smoking<sup>130, 135</sup>. FEV1 has also been inversely associated with all-cause stroke in hypertensive subjects, and in subjects with both ischemic heart disease and hypertension, adjusting for smoking<sup>136, 137</sup>. A relationship between asthma all-cause stroke has been suggested<sup>138</sup>.

The causes for the association between reduced lung function and cardiovascular diseases are not known. Systemic inflammation, increased ECM degradation, hypoxia, and oxidative stress have been suggested mechanisms<sup>132</sup>. It is not studied whether reduced lung function is associated with incidence of SAH, although reduced lung function is associated with several other vascular diseases and may share pathophysiological mechanisms, such as degradation of ECM, with SAH.

## COPD

Chronic obstructive pulmonary disease (COPD) is a common disease with a prevalence around 10% in the adult population worldwide and in Sweden<sup>139, 140</sup>. COPD was recently ranked as the third most common cause of death worldwide in the Global burden of disease study<sup>7</sup>, and prevalence of COPD is expected to increase in the next years<sup>140</sup>. COPD is characterized by chronic airflow limitation in combination with respiratory symptoms. It is diagnosed based on a FEV1 to FVC ratio below 0.7<sup>140</sup>. COPD is associated with chronic airway inflammation, causing degradation of elastic tissues and emphysema.

Smoking is the main risk factor for COPD. Other important risk factors include age, chronic asthma, repeated lower respiratory-tract infections, occupational exposure to dust and fumes<sup>141</sup>, and indoor air pollution from solid fuels like wood and charcoal<sup>142</sup>. Genetic factors probably also influence the susceptibility of COPD. Deficiency of alpha-1-antitrypsin, a major anti-protease, is the most well established genetic factor but on a population level it explains only a small fraction of cases<sup>143</sup>. Although smoking is the major risk factor, some studies have shown that relatively large proportions COPD are never (around 25%) or non-smokers (10-45%)<sup>142, 144</sup>. In a Swedish study, 90% of men and 65% of women with COPD were ever-smokers<sup>143</sup>. In severe COPD, 100% and 85% were ever-smokers<sup>143</sup>.

Cardiovascular disease is a major cause of hospitalizations and deaths in COPD patients<sup>145</sup>. COPD is associated with increased incidence of coronary heart disease, heart failure and cardiovascular mortality<sup>146-149</sup>. To some part, these associations are explained by shared risk factors, however, also adjusting for all risk factors, associations have remained in several studies<sup>146-149</sup>.

COPD has been associated with increased stroke risk in some studies, although results have been somewhat inconsistent, and various designs have been used. One study found a high effect estimate for COPD (HR 2.8) adjusting for sex and smoking, but this was largely restricted to younger age groups<sup>150</sup>. Other studies have found an increased risk of around 50%<sup>147, 148, 151, 152</sup>, but after adjustment for comorbidities, the association was not significant in all studies<sup>147, 151</sup>. Thus, it remains unclear if COPD is associated with an excess risk of stroke. Also, the main stroke types (ischemic stroke, ICH, SAH) have not been studied in COPD.



# Aims

The overall aim was to study risk factors for SAH in the general population, and in some cases, to relate these associations to the other stroke types. The specific aims of the included papers are described below.

- I. To investigate the association between reduced lung function assessed in spirometry and incident SAH in subjects from the general population, and to assess if smoking influences this association.
- II. To evaluate if the level of blood leukocytes is associated with risk of incident SAH.
- III. To evaluate if serum levels of FGF23 are associated with risk of incident SAH.
- IV. To investigate whether patients with COPD are at increased risk of stroke, including all main stroke types, in comparison with reference individuals from the general population.
- V. To find genetic variants of the protein-coding regions of the genome associated with risk of SAH.



# Methods

## Study populations

### **The Malmö preventive project**

The Malmö preventive project (MPP) was a health screening project with the aim to identify individuals with high risk for cardiovascular disease, and people with alcohol abuse, and to offer preventive interventions<sup>153, 154</sup>. Men and women were invited by mail to the screening program between 1974 and 1992. In the first years of the program (1974-1981), predominantly men were examined and in the later years (1982-1991), predominantly women. Invited were all men living in Malmö, born 1921, 1926-1942, 1944, 1946, 1948-1949, and all women, born 1926, 1928, 1930, 1932-1936, 1938, 1941-1942, 1949. A total of 22,444 men and 10,092 women participated, corresponding to an overall attendance rate of 71%. The mean age was 44 years (range 27-61) in men and 50 years (range 28-58) in women. The screening visit included a physical examination, a panel of laboratory tests, and a computer-based questionnaire<sup>153, 154</sup>.

There was no risk-reducing effect on overall, or cause-specific mortality of the MPP screening and intervention program when invited subjects were compared to non-invited<sup>154</sup>, which indicate that the interventions probably were not effective. However, the MPP has been extensively used in research purposes.

Individuals that were invited to the MPP but did not participate were more often born abroad, and living alone, and they had lower educational and socioeconomic levels, compared to those who participated. Total mortality as well as incidence of non-fatal ischemic heart disease was higher in non-participants<sup>154</sup>.

### **The Malmö Diet and Cancer Study**

The Malmö Diet and Cancer Study (MDC) is a population-based cohort study performed in the city of Malmö, Sweden, with the main objective to study the effect of diet and life-style factors on cancer incidence and mortality<sup>155, 156</sup>. All men, born 1923-1945, and all women, born 1923-1950, living in Malmö were invited to



participate in the MDC study by advertisements in public places and local newspapers, and personal invitation letters, as described in detail elsewhere<sup>156, 157</sup>. Individuals with insufficient skills in the Swedish language, or mental incapacity were excluded<sup>156</sup>. In the invited birth cohorts, there were a total of 74,138 individuals, of which 68,905 were eligible for the study, and 30,447 responded and were initially included. A total of 28,449 (11,246 men and 17,203 women) had a complete baseline examination and 28,098 had also a complete diet assessment. This corresponds to a participation rate of about 41 %<sup>156</sup>. The part of the MDC with complete diet assessments are also included in the WHO project European Prospective Investigations into Cancer and nutrition (EPIC).

### *Representativity of the MDC*

About 60 % of participants were women in the MDC study. Women were also somewhat younger than men (mean age of 57 and 59 years). The reasons for these differences were that additional invitations were sent out to younger women (born 1946-1950) after redefining the cohort in 1995, in order to improve the ability for studying determinants of breast cancer<sup>156</sup>. About 11% of the MDC cohort was born outside of Sweden, most of these in the neighboring countries. In a study assessing the representativity of the MDC, it was found that incidence of cancer and mortality, including mortality from cardiovascular disease, was higher in non-participants compared to participants. The prevalence of smoking and obesity, and socio-demographic characteristics were similar in participants and non-participants, based on a comparison in some of the included birth cohorts (1923, 1933, 1943) with results from a health survey with 76 % participation rate<sup>156</sup>. Because incidence of life-style related cancers, e.g. lung cancer, was increased, it should, however, be suspected that life-style related risk factors were more common in non-participants. The results thus indicate that the MDC represents a somewhat healthier part of the population<sup>156</sup>.

### *Baseline examinations*

In the MDC, all participants underwent physical examinations and filled out a self-administered questionnaire. At the first visit, blood pressure, height, weight and body composition were measured, and the questionnaire was handed out. Venous blood samples were collected and blood separated within one hour, and serum and plasma samples were biobanked at -80 or -140 degrees<sup>158</sup>. At the second visit, most often about two weeks later, the questionnaire was returned and checked for completeness by trained study nurses. A detailed diet assessment was completed as described elsewhere<sup>159</sup>. Some of the participants were randomly selected for additional examinations, including ultrasound of the carotid arteries and additional blood samples, and were included in the MDC Cardiovascular cohort (MDC-CC), for which the main goal was to study the epidemiology of carotid disease<sup>160</sup>.

## **The COPD cohort**

The COPD cohort was collected by use of several Swedish nation-wide registers. All individuals between 40 and 84 years of age, having a discharge diagnosis of COPD in the National Inpatient Register, either as primary or as secondary diagnosis, between 1987 and 2003 were identified (n=132,017). For each individual with COPD one reference individual was randomly selected from the Total Population Register, matched for sex, year of birth, and county of residence during the year of the first COPD diagnosis.

The Swedish National Inpatient Register has been recording all discharges from hospitals in Sweden since 1987<sup>161</sup>. Each discharge includes one primary diagnosis, and in many cases also several secondary diagnoses. Diagnoses are made by physicians in routine care and contra-signed by board-certified specialist. The general quality of the register is high<sup>161</sup>. The specific validity has been evaluated for different diagnoses. For COPD diagnoses, a validation study showed that about 10% are misclassified or uncertain<sup>162</sup>. The validity of stroke diagnoses is discussed in the methodological discussion.

We excluded 11,131 (8.4 %) of COPD patients and 2678 (2.0 %) of reference subjects, who died less than 30 days after the date of inclusion. Also, 1736 (1.3 %) reference subjects who had been hospitalized with a COPD diagnosis prior to inclusion, were excluded, as were 682 (0.5%) COPD patients because of data inconsistencies.

All individuals with a stroke diagnosis in the national Inpatient Register (ICD 9 codes: 430, 431, 434 and 436, ICD 10: I60, I61, I63, I64) or diagnosis of “late effects of cerebrovascular disease” (ICD 9: 438 and ICD 10: I69.0-I69.4) before or on the start of follow-up were additionally excluded because they were considered to have a history of stroke (4715 [4.1%] of COPD patients and 8624 [7.4%] of reference subjects).

To keep matched pairs of COPD and reference individuals after exclusions, 8161 (6.2 %) COPD patients and 18,785 (14.2 %) referents had to be additionally excluded because they missed their counterpart.

The National Population Censuses (Folk- och bostadsräkningen) was used to achieve information about socioeconomic status. The Censuses were conducted every fifth year between 1960 and 1990 and participation rate was over 97% between 1980 and 1990. Socioeconomic data was categorized into manual, non-manual, other types of occupation, and being outside the workforce (i.e. old age pensioners, housewives, students, unemployed or those with missing information). Information from The Total Population Register, which holds data about vital status, date of death, country of birth, and dates of immigration or emigration, was linked

to the study population. Information about comorbidities was achieved from the Inpatient Register using ICD 9 and 10 codes listed at discharge.

The final study cohort consisted of 103,419 COPD patients and the same number of reference subjects, born between 1903 and 1963. Among both COPD patients and referents, 54% were men and the mean age was 70 years. 62% of COPD patients had COPD as primary diagnosis. 88% of COPD patients and 91% of reference subjects were born in Sweden. Prevalence of cardiovascular comorbidities were generally higher in COPD patients than reference subjects, and 32% of the COPD group compared to 57% of the reference group had not been hospitalized before inclusion in the study.

## Assessment of exposure

### Paper I

In MPP, spirometry was performed during most of, but not the entire, screening period, including 94 % (21,186) of men and 71 % (7748) of women. In men, the screening included whole birth cohorts. Characteristics in subjects that were screened and not screened with spirometry, respectively, have been described<sup>128</sup>.

Spirometry was performed using a Spirotron apparatus (Drägerwerk AG, Lübeck, Germany) with the person in a standing position having no nose clip. For analysis, the lung volumes were standardized for age and height using equations derived from linear regressions of never smokers in the present cohort<sup>129, 163</sup>. FEV1, FVC and FEV1/FVC are expressed as their percentages of the predicted values.

Blood pressure was measured twice in the right arm after 10 minutes of rest. Cholesterol, blood glucose and erythrocyte sedimentation rate (ESR) were analyzed in fasting blood samples. Diabetes was defined as fasting venous blood glucose >6.1 mmol/L and/or a positive answer to the question “Do you have diabetes mellitus?”. Smoking habits were based on self-report and smoking was categorized into current or non-smoking. If a person answered the question “Are you mostly engaged in sedentary activities in spare time, for example, watching TV, reading, going to the movies?” positively, he/she was classified as being physical inactivate in spare time. Individuals were considered high alcohol consumers if they had more than two affirmative answers to the modified shortened version of the Michigan Alcoholism Screening Test<sup>164</sup>, which can identify a high proportion of heavy alcohol consumers<sup>164</sup>. A history of myocardial infarction or stroke was recorded based on the questionnaire and data from the inpatient register. Angina pectoris was considered prevalent if the subject reported a doctor’s diagnosis or use of nitrates.

Individuals with a history of stroke (n=30), myocardial infarction (n=258), or angina pectoris (n=466) at baseline were excluded. Subjects with missing information about blood pressure, alcohol consumption, body mass index, cholesterol, diabetes, physical inactivity, or ESR, were excluded (n=350). Fifty-five with ESR >50 mm (n=55) were excluded because this may indicate severe ongoing disease. 27,771 subjects, 20 534 men and 7237 women, were included.

## **Paper II**

Leukocyte count was measured in MPP for individuals screened between 1974 and 1981. During this time, men born 1921, 1926–1940, 1942, 1944, 1946 and 1949 were examined, and women born in 1926, 1931, 1938, and 1949. In total, 19,930 subjects, 17,198 men (78 % of all men in MPP) and 2,732 women (31 % of all women in MPP) had leukocyte counts measured. Leucocyte count was determined with automatic counters at the Malmö hospital laboratory. Other baseline data was measured and defined as described for **Paper I**. Smoking status included also former smoking in **Paper II**, i.e. previous daily smoking for at least 6 months.

Ten individuals with a history of stroke were excluded, and 119 because of missing information about blood pressure, cholesterol or diabetes. Excluded were also seven individuals with leukocyte counts  $>20.0 \times 10^9$  cells/L in order not to include possible false or unreasonable values. A total of 19,794 subjects (2,711 women and 17,083 men) were included.

## **Paper III**

Levels of FGF23 were analyzed in stored serum samples of SAH cases and control subjects selected for the nested-case control study that is described under Study design. Levels of FGF23 were analyzed with the Proseek multiplex CVD I assay, a proximity extension immuno-assay<sup>165</sup>. The intra- and inter-assay coefficients of variation were 9 % and 26 %, respectively<sup>165, 166</sup>. Assessment of the other baseline data is described under Study population.

Exposure data in **Paper IV** is described under Study population.

## Paper V

Genotyping was performed with the Illumina OmniExpressExome Beadchip version 1.0 at Broad institute, Boston, using the autocall algorithm for genotype calling. This genotyping array includes a standard genome-wide array, and an exome part, which is designed with focus on coding variants, i.e., missense, nonsense and splice site mutations<sup>125</sup>. The exome part of the array includes 244,194 single nucleotide variants (SNVs).

Quality control (QC) of genotype data was performed in a combined dataset of SAH cases, other stroke cases from the MDC and stroke-free subjects from the MDC Cardiovascular cohort. QC on the individual-level excluded subjects with 1) a call rate <95%, 2) discrepancy between reported and genotyped sex, 3) excess of heterozygosity, 4) relatedness of up to second degree (identity by descent sharing,  $\pi\text{-hat} > 0.1875$ ), and 5) population outliers based on inspection of the first two principal components. SNVs were excluded if they had a call rate <95%, a p value <  $10^{-6}$  for departure from Hardy-Weinberg equilibrium, or if they were on non-autosomal chromosomes. All variants on the exome part of the genotype array that passed QC were considered for inclusion in the study. To evaluate some previously reported loci for intracranial aneurysm, we used data for ten common variants in the 9p21 region<sup>115</sup>, from the ‘GWAS part’ of the genotyping array.

## Ascertainment of outcome

### Paper I-III and V

Cases of SAH treated in hospital and SAH patients followed-up in the outpatient clinics in Malmö were identified by use of the Malmö Stroke Register. The stroke register was started in 1989 with the aim to record all stroke cases in the city of Malmö<sup>45</sup>. Cases of stroke have been identified continuously through systematic search among inpatients, outpatients and in the emergency department of the hospital in Malmö. In addition, stroke cases in primary care have also been searched for<sup>167</sup>. A research nurse, supported by a stroke neurologist, has validated all diagnoses of stroke through review of medical records and sometimes with patient interviews. Stroke was defined in accordance with the WHO as sudden onset of focal or global (in the situation of SAH) loss of neurological function, lasting for more than 24 hours or leading to death before then, without any apparent non-vascular causes<sup>5</sup>. Diagnosis of cerebral infarction was made if there was a typical clinical presentation and CT confirmed an infarction or excluded other causes. Diagnosis of intracerebral hemorrhage and SAH required confirmation in CT and/or

autopsy, and for SAH, diagnosis could also be made based on lumbar puncture. If neither CT nor autopsy were done the stroke was considered unspecified (stroke not specified as hemorrhage or infarction).

Cases of SAH treated in hospitals outside of Malmö were identified by the national Inpatient Register based on the International Classification of Diseases (ICD) codes for non-traumatic SAH, 430 (ICD 8 and 9) and I60 (ICD 10). Patients with SAH from Malmö are normally treated in the neurosurgical clinic in Lund in the acute phase. Information about which hospital and which clinic the patient was discharged from, and what operations were done during hospitalization, were also obtained for cases identified in the Inpatient Register.

Cases of SAH dying unexpectedly outside of hospital, or in emergency rooms shortly after arriving to hospital, and therefore not registered in the hospital-based register, were identified with the Causes of Death register.

In Sweden, ICD8 was used 1969-1986, ICD 9 1987-1996 and ICD10 from 1997. We did not distinguish aneurysmal and non-aneurysmal SAHs in the present outcome classification. Numbers for each of the sources used to identify cases are shown in **Table 3**.

Over approximately twenty years follow-up time in MDC (until Dec 31, 2010, median 16 years), the incidence rate of SAH per 100,000 person-years was 15 in men and 22 in women. In men aged 45-54, 55-64 and 65-73 years at baseline, incidence rates were 14, 15 and 17, and in women corresponding rates were 19, 24 and 24 per 100000 person-years. Mean age at SAH was 66 years (range 48-88).

In MPP, over some 30 years follow-up (until Dec 31, 2008, median 28 years), rates of SAH were 19 and 26 per 100,000 person-years for men and women, respectively. Mean age at SAH was 59 years (range 29-80).

**Table 3.** Numbers of SAH events identified with different methods in the MDC and MPP cohorts.

	MDC	MPP
Total	89	149
Malmö Stroke Register	57	49
Inpatient Register	16 (12 Ns)	88 (72 Ns)
Causes of death register	16 (13 autopsy)	12 (11 autopsy)
Fatal cases (<28 days)	36 (40%)	42 (28%)

MDC; Malmö Diet and Cancer study, MPP; Malmö Preventive Project, Ns; Neurosurgical clinic.

## Paper IV

Diagnoses of stroke during follow-up were identified in the national Inpatient Register using the ICD codes for SAH, ICH, ischemic stroke and stroke not specified as hemorrhage or infarction; ICD9: 430, 431, 434 and 436, ICD10: I60, I61, I63, I64.

## Study design and statistical analysis

An overview of study designs and statistical methods is shown in **Table 4**.

**Table 4.** Overview of study design and methods used in each paper.

	Study population	Study design	N subjects	N events	Years included in follow-up	Statistical methods used for main analysis
Paper I	MPP	Cohort study	28,253	149 SAH	1974-2008	Cox regression, Kaplan Meier curve
Paper II	MPP	Cohort study	19,794	95 SAH	1974-2008	Cox regression, Kaplan Meier curve
Paper III	MDC	Nested case control	311	79 SAH	1991-2010	Conditional logistic regression
Paper IV	All COPD 1987-2003, referents from general population	Cohort study	206, 838	17,402 strokes	1987-2010	Stratified Cox regression
Paper V	MDC	Nested case control	416	104 SAH	1991-2008	Logistic regression (Firth's bias corrected), SKAT-O

## Cohort studies

A cohort is a group defined by a certain characteristic that is followed over time. In cohort studies, a defined group of individuals is followed over time and death and/or certain events are recorded during follow-up. The incidence rates can be calculated and compared according to an exposure of interest.

In the cohort study, the time to event is the essential measurement. The umbrella term for the kind of analysis handling time-to-event data is survival analysis. Each individual contributes with time-at-risk from inclusion until experiencing the event of interest. If a subject leaves the study for other reasons it is said to be censored. This occurs if a subjects dies from other causes than the event of interest, withdraws from the study, or is lost to follow-up for other reasons, for example moves out of the country making registration of events impossible. Censoring also occurs of those still at risk at the end of the study because most studies do not follow all individuals

until they die (right-censored). Censoring is a kind of missing data since the exact time to event is not known for censored individuals. The incidence rate is calculated by dividing the number of events with the total time-at-risk for all individuals, and is expressed as a rate, e.g per 100,000 person-years<sup>168</sup>.

The survival function is defined as  $S(t) = P(T > t)$ , and gives the probability that a subject survives longer than the time  $t$ . In the start of a study ( $t = 0$ ) it gives the probability of survival of 1 and after an infinite period of time ( $t=\infty$ ) the probability of survival is 0. As events occur during follow-up the probability of staying alive or event free changes, but this will be dependent on the time interval for which the survival is calculated. The hazard function can overcome this by estimating the instantaneous rate of occurrence of the event in subjects that have survived and are event-free at that time point<sup>169</sup>.

### *Kaplan-Meier survival curve*

The probability of survival (or in many cases of not having an event) to a certain time point can be calculated using the Kaplan-Meier estimator, which is commonly plotted for graphical presentation. The Kaplan-Meier estimate describes the probability of survival past a certain time point and is calculated at each failure time (time of an event) by dividing the number of individuals surviving past the time point by the number at risk before that time point<sup>169, 170</sup>. The log-rank test can compare survival between different exposure groups<sup>169</sup>.

### *Cox regression*

Cox proportional hazards regression (here referred to as only Cox regression) is frequently used for survival analysis and was introduced by James Cox in 1972<sup>169, 171</sup>. This model uses the hazard function but does not make any assumptions about the shape of the baseline hazard, meaning that it does not have to be fitted to a pre-defined distribution. The hazard rate for the  $j$ th subject is given by the formula:

$$h(t|x_j) = h_0(t)\exp(x_j\beta_x)$$

In Cox regression the time scale used for calculating time-to event might be time on study, the attained age of the subjects, or calendar time<sup>172</sup>. Different groups of exposure are defined at baseline but may also change over time so that the same subject can contribute follow-up time to different exposure groups. The hazard ratio (HR) between exposed and non-exposed is obtained in Cox regression and models can be adjusted for covariates.

The Cox model assumes that the effect of an exposure is constant with time (proportional hazards assumption). When applying Cox regression, this assumption should be tested<sup>168, 172</sup>. Different methods can be used. Graphical methods include plotting the log-transformed values of the Cox survival function over time (-log-log plots) and visually inspect if the curves for exposed and unexposed are parallel. If



not parallel, a violation of the assumption is suspected. A formal, statistical method to test the assumption is based on plotting the scaled Schoenfeld residuals of the Cox model with time and test the correlation between these and survival time, which will give a p value (goodness of fit test). If Schoenfeld residuals are not correlated with time the proportionality assumption is valid. The proportionality assumption can also be tested by comparing the effect estimates between different intervals of follow-up time by using a likelihood ratio test, comparing the model with an interaction with time intervals to a model without this interaction term. If the proportional hazards assumption is not met other ways of analyzing the data should be considered, such as presenting the results separately in different time intervals, include time-dependent covariates in the model, use of other regression models, or use an adapted Cox model weighing the time scale<sup>173</sup>. The Cox model and other common survival methods assume non-informative or independent censoring, meaning that the risk of event is assumed to be the same in censored and non-censored subjects in each risk set<sup>169</sup>.

**Paper I, II and IV** has the design of cohort studies. Study subjects were categorized according to exposure at the baseline examination and then followed-up until the event of SAH, death, emigration from Sweden, or last follow-up date.

## **Interaction and sensitivity analyses**

Interaction or effect modification is the phenomenon of two factors acting synergistically on risk of disease. The joint effect of such risk factors will be greater than the sum of the individual effect of each factor. Statistical interaction can be evaluated by a multiplicative interaction term. However, this is limited by power because number of cases will be lower in each risk stratum and by the use of a specific statistical model. Biological interaction may be present by two factors working together on disease risk, although no statistically significant interaction term is found in a regression model<sup>168, 174</sup>.

Sensitivity analyses are performed to evaluate if results will hold when changing some characteristics of the analysis. Examples are restricting the analysis to a certain subgroup, for example women over 65 years, or men with missing information about a covariate, in order to evaluate certain hypotheses, or address any concerns raised by the main results.

## **Nested case-control studies**

A case control study compares an exposure between cases and controls. An adequate selection of cases and controls is of high importance to avoid selection bias. The general rules for selection of controls are that controls should come from the same population that cases came from (i.e. the source population) and that should be selected independently of exposure status<sup>175</sup>. If cases and controls are both part of the same cohort study the study is referred to as nested case control study. This may be a convenient way of studying exposures that are not available in the whole cohort but can be determined in all incident cases and randomly sampled disease-free individuals (controls), matched by follow-up time<sup>176</sup>.

Matching in case control studies have both benefits and drawbacks. Matching can minimize the variance in cases and controls, and thus improve the efficiency (precision). However, matching will introduce selection bias that should be controlled for by stratifying (conditioning) the analysis on matching variables<sup>175</sup>.

### *Conditional logistic regression*

Logistic regression is used to study categorical outcome variables in relation to one or several independent variables (covariates) and is often the choice of statistical method for case control studies. It works with odds of disease. Conditional (matched) logistic regression stratifies the analysis by each pre-defined group and thus cases and controls are compared only within each group/strata. This method is recommended for matched case control studies to take into account the matching variables<sup>168, 177</sup>.

**Paper III and V** used a nested case control design, within the MDC cohort study. In the whole MDC as a cohort study, a total of 89 cases occurred during a mean follow-up of 15.5 years (439,818 person-years at risk, until Dec 31, 2010).

## **Paper I**

In paper I, a cohort study design was used for studying measures of lung function and the incidence of SAH. Lung volumes, FEV1, FVC and the FEV1/FVC ratio were divided into quartiles with equal numbers of men and women in each group. Lung volume variables were also studied as continuous variables per 10% decrease of FEV1, FVC and the FEV1/FVC ratios, respectively. Risk factor distributions were compared between the quartiles of FEV1 and FVC, respectively, using one-way analysis of variance (ANOVA) and logistic regression.

The Kaplan-Meier estimator and Cox regression models was used to study the association between lung volumes and incident SAH. The proportional hazards

assumption was evaluated by inspection of Kaplan-Meier curves and by adding a time-dependent covariate to the analysis. Time-on-study was used as time scale. Subjects were censored if they died from other causes than SAH, emigrated from Sweden, or were free of an event at end of follow-up (Dec 31, 2008). Both unadjusted models, and models adjusted for important potential confounders (age, smoking, blood pressure, blood pressure medication, diabetes, total cholesterol, body mass index, high alcohol intake, physical inactivity) were implemented. Models were stratified for year of screening (3-year groups) to account for differences during the long screening period.

## **Paper II**

A cohort study design was applied to study the association between leukocyte count and incidence of SAH. Subjects were follow-up from baseline until the event of SAH, death, emigration or Dec 31, 2008.

Leucocytes were divided into quartiles with equal numbers of men and women in each quartile, to adjust for the association between sex and SAH. Baseline characteristics were described and compared between quartiles of leukocytes using ANOVA and logistic regression. Adjusted Cox proportional hazards regression models estimated hazard ratios for risk of SAH according to quartile of leukocytes and also per standard deviation increase of leukocytes. Time-on-study was used as time scale. Models were adjusted for important risk factors selected a priori based on knowledge about exposure and outcome; age, smoking, blood pressure, blood pressure medication, diabetes, body mass index, alcohol intake, total cholesterol, and physical inactivity. History of myocardial infarction or angina pectoris, and FEV1, was also added to the full model, but these factors had no clear influence on the estimate for the association between leucocytes and SAH. The proportionality assumption was checked by inspection of incidence rates plotted over time and a goodness of fit test, using Schoenfeld residuals. Multiplicative interaction terms between leukocytes and each of the other variables, respectively, were evaluated in the Cox model.

Harrell's concordance statistics was used to estimate the ability of the risk factor adjusted Cox model to predict SAH<sup>178</sup>. The discriminatory ability of a variable or a model can be described by c statistics, which is equivalent to the area under the receiver operating characteristic curve (a plot of 1-specificity by sensitivity). A value of 0.5 has the same predictive ability as chance and 1 is perfect prediction.

## Paper III

A nested case control design was used, including all incident cases of SAH from the MDC (until Dec 31, 2010). The date of SAH was considered the index date. For each case, one to three control subjects still at risk for SAH on the index date, were randomly selected, matched to cases by age (5 year groups) and sex. This was done using the Stata command sttocc (StataCorp, College Station, TX, USA). Characteristics of cases and controls are shown in **Table 5**.

Serum levels of FGF23 was analysed in biobanked blood samples for the included cases and controls. FGF23 is right-skewed and was log-transformed and divided into quartiles before analysis. The association between FGF23 and SAH was studied by conditional logistic regression adjusting for smoking, blood pressure, blood pressure medication, diabetes, body mass index, alcohol intake and leucocyte count.

**Table 5.** Characteristics in incident cases and controls in the nested case control study.

	Cases	Controls
Number	79	232
FGF23, median (IQR)	3.23 (1.41)	2.97 (1.00)
Age, years, mean (SD)	58 (8)	58 (8)
Women, %	72	74
Systolic BP, mmHg, mean (SD)	143 (17.6)	137 (19.6)
Diastolic BP, mmHg, mean (SD)	87 (9.0)	84 (9.5)
BP drug, %	20.3	12.1
Body Mass Index, kg/m <sup>2</sup> , mean (SD)	24.7 (4.3)	25.0 (3.7)
Smoking, %	50.6	28
High alcohol intake, %	1.3	4.3
Leukocyte count*, 10 <sup>9</sup> /L, mean (SD)	6.8 (1.8)	6.1 (1.7)
GFR <sup>a</sup> , ml/min per 1.73 m <sup>2</sup> , mean (SD)	74.1 (25)	76.5 (14)

FGF23; Fibroblast growth factor 23, IQR; interquartile range, SD; standard deviation, BP; blood pressure, GFR; estimated glomerular filtration rate

<sup>a</sup>Estimated with the Cockcroft-Gault formula in 14 of the cases and 72 of the controls, which are included in a randomly selected subsample, the MDC-CC.

## Paper IV

Paper IV is a matched cohort study, in which COPD patients (exposed) and reference individuals (non-exposed) were follow-up for incidence of stroke. Cox models, adjusting for comorbidities and total length of prior hospitalizations, were used to study the association of COPD with incidence of stroke subtypes. Analysis was stratified by matching variables; year of birth, year of inclusion, sex, and county of residence, and time-on-study was used as time scale.

Follow-up was started 30 days after inclusion to the study and each individual was followed until the event of a stroke, death, emigration, or until ten years after inclusion, however, at longest until Dec 31, 2010. Reference subjects being diagnosed with COPD during follow-up were censored.

The proportional hazards assumption was evaluated by inspection of –log–log plots and by comparing a model with the interaction between COPD status and follow-up time intervals (0-2, 2-5 and 5-10 years), to a model without this interaction, using a likelihood ratio test. There was strong evidence of a time-dependent association between COPD and all-cause stroke ( $p<0.001$ ), ischemic stroke ( $p<0.001$ ) and ICH ( $p\leq 0.035$ ), but not for SAH ( $p=0.47$ ). Results were therefore calculated for each time interval separately.

Interactions were evaluated for COPD with all covariates, respectively, by comparing a model with the interaction to a model of main effects, using a likelihood ratio test. Sensitivity analyses were performed 1) by counting both primary and secondary diagnoses of stroke as outcome, 2) based on COPD as primary or secondary diagnosis, and 3) including only the reference subjects that had been hospitalized (an in-patient control group).

## Paper V

A nested case control material was constructed by including all prevalent and incident cases of SAH until 31 Dec, 2008 in the MDC Study. Three controls per case were randomly selected from SAH-free subjects that took part in the MDC-CC, matched by sex and age (5-year groups).

Genotype data from the exome part of the genotyping array was analyzed by two different approaches; gene-based and single-variant analysis. The gene-based analysis was performed using the optimal combination of the Sequence Kernel Association test and the burden test (SKAT-O)<sup>179</sup>. Genes with two or more variants, and with a cumulative minor allele count of  $\geq 5$ , were included. An upper minor allele frequency limit of 10% was applied. A total of 36,195 SNVs in 7636 genes were included.

Single variant analysis was performed of SNVs with a minor allele count  $>5$  ( $n=39,509$ ) with logistic regression models, using the Firth's bias corrected test statistics and assuming an additive effect of each allele.

All analyses were adjusted for age and sex. Bonferroni correction was applied to correct for multiple testing. The significance level in gene-based analysis was  $p < 5.4 \times 10^{-6}$  (corrected for the number of genes) and  $p < 1.3 \times 10^{-6}$  in the single-variant analysis (corrected for the number of variants).





# Results

## Paper I

Baseline characteristics and risk factors by quartiles of FEV1 are displayed in **Table 6**. Low FEV1 was clearly associated with current smoking, diabetes, high alcohol consumption and physical inactivity.

**Table 6.** Baseline characteristics and risk factors by quartiles of forced expiratory volume in 1 second (FEV1).

	Q4	Q3	Q2	Q1 (low)	All
FEV1, % of predicted, range	>106	96-107	85-96	< 86	
n	6942	6944	6943	6942	27771
Women, %	26	26	26	26	26
Age, years, mean	45±7	44±7	44±7	45±7	44±7
Height, cm	173±9	174±9	174±9	173±9	174±9
Current smoking, %	33.3	41.8	51.4	63.6	47.5
Systolic blood pressure, mmHg	126±15	126±15	125±15	127±16	126±15
Antihypertensive medication, %	3.5	4.1	4.1	5.6	4.3
High alcohol consumption, %	12.1	13.1	14.5	16.1	13.9
Diabetes, %	1.7	1.8	2	3.5	2.3
Cholesterol, mmol/L	5.6±1	5.6±1	5.6±1	5.7±1	5.6±1
Body mass index, kg/m <sup>2</sup>	24±3	24±3	24±4	25±4	24±3
Physical inactivity, %	37	42	47	50	44
ESR, mm	6.0±5	6.1±6	6.4±6	7.4±7	6.5±6

Values are mean±standard deviation or percentages.

During 725,245 person-years of follow-up, 98 men and 47 women had a SAH. The overall crude incidence was 18.3 and 26.5 per 100 000 person-years in men and women, respectively.

Low lung function, by means of FEV1, and low FEV1/FVC ratio (airway obstruction), respectively, was associated with increased incidence of SAH. The association remained after adjusting for all covariates, including smoking and hypertension: Subjects with FEV1 in the lowest compared with the highest quartile had a hazard ratio (HR) of 2.24 (95% CI 1.32-3.81) for SAH. The corresponding

result for the FEV1/FVC ratio was 1.92 (1.14-3.23). FVC was not associated with SAH (**Table 7**). The Kaplan-Meier survival curve by quartiles of FEV1 is shown in **Figure 5**.

## Sensitivity analyses

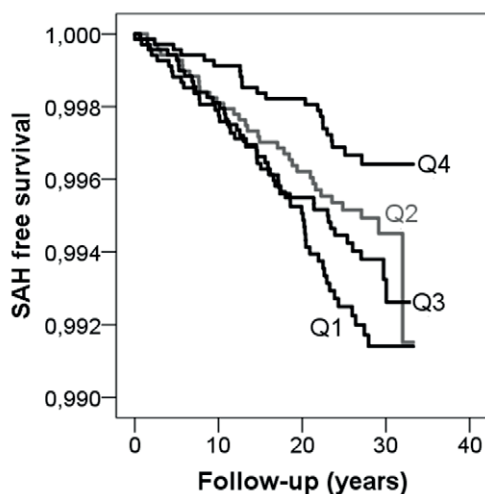
### *Smoking*

Estimates were similar or somewhat higher in non-smokers as in the whole cohort (FEV1: HR [lowest versus highest quartile]: 2.59, 95% CI: 1.06-6.32; FEV1/FVC: 2.34, 1.01-5.39; per 10% decrease of FEV1, HR: 1.21, 1.05-1.40, and FEV1/FVC, 1.34, 1.09-1.35) whereas estimates were somewhat lower in smokers (FEV1: HR [in lowest versus highest quartile]: 1.52, 95% CI 0.82-2.84, FEV1/FVC: 1.53, 0.77-3.04, per 10% decrease FEV1, 1.07, 0.95-1.21, and FEV1/FVC, 1.14, 0.96-1.35). There was no evidence of effect modification by smoking for any of the lung volumes ( $p=0.2$  for both). In a subgroup with information about former smoking, the group of never smokers ( $n=9625$ ) had also similar or slightly higher effect estimates compared to the whole cohort (for example, per 10% decrease of FEV1, HR: 1.29, 95% CI: 1.09-1.54, FEV1/FVC: 1.33, 1.02-1.74) adjusted in the full model.

In this data, current smoking was strongly associated with SAH with a HR of around two compared to non-smokers (HR: 1.84, 95% CI: 1.32-2.57). Of interest, the effect of smoking was evidently higher in women than in men (HR: 4.1, 2.2-7.9 versus 1.3, 0.9-2.0,  $p$  for interaction=0.002).

### *Hypertension and sex*

Individuals with hypertension had somewhat higher estimates than those that were normotensive (FEV1: HR [lowest vs. highest quartile]: 2.46, 95% CI: 1.05-5.76 in hypertensive and 1.96, 0.99-3.89 in normotensive subjects, FEV1/FVC: HR: 2.61, 1.15-5.93, in hypertensive and, 1.66, 0.81-3.40 in normotensive subjects). The association between FEV1 and SAH also tended to be stronger in men compared with women (HR [lowest vs. highest quartile]: 2.68, 95% CI: 1.41–5.10 in men and 1.47, 0.47–3.77 in women; per 10% decrease: 1.17, 1.04-1.30 in men and 1.03, 0.87-1.23 in women) whereas results for FEV1/FVC ratio seemed relatively similar in men and women (HR [lowest vs. highest quartile]: 1.68, 95% CI: 0.90–3.14, in men and 2.42, 0.94–6.22, per 10% decrease 1.21, 1.03-1.42, in men, and 1.20, 0.93-1.55 in women). However, there was no statistical evidence of effect modification by sex or hypertension ( $p>0.7$  for all).



**Figure 5.** Kaplan-Meier survival estimates by quartiles of FEV1; Q1, first (lowest) quartile, to Q4 (highest).

**Table 7.** Results for lung volumes and SAH incidence.

	Q4	Q3	Q2	Q1 (low)	P trend	Per 10 % decrease
<b>FEV1</b>						
N all/n SAH	6942/21	6944/42	6943/34	6942/48		
HR (95% CI)	1	2.10 (1.24-3.55)	1.75 (1.01-3.02)	2.64 (1.57-4.43)	0.001	1.17 (1.07-1.28)
HR*	1	2.01 (1.19-3.40)	1.58 (0.91-2.74)	2.24 (1.32-3.81)	0.014	1.13 (1.03-1.24)
<b>FEV<sub>1</sub>/FVC</b>						
N all/n SAH	6942/22	6944/30	6943/50	6942/43		
HR	1	1.38 (0.80-2.39)	2.33 (1.41-3.86)	2.08 (1.24-3.49)	0.001	1.25 (1.09-1.43)
HR*	1	1.39 (0.80-2.40)	2.30 (1.39-3.81)	1.92 (1.14-3.23)	0.003	1.22 (1.06-1.40)
<b>FVC</b>						
N all/n SAH	6942/33	6944/34	6943/42	6942/36		
HR	1	1.05 (0.65-1.69)	1.33 (0.84-2.11)	1.21 (0.75-1.94)	0.286	1.08 (0.98-1.20)
HR*	1	1.01 (0.62-1.62)	1.22 (0.77-1.94)	1.04 (0.64-1.68)	0.689	1.05 (0.94-1.16)

\*Adjusted for sex, age, smoking, systolic blood pressure, antihypertensive medication, high alcohol consumption, cholesterol, diabetes, body mass index, physical inactivity and erythrocyte sedimentation rate.

## Paper II

High leucocyte count was strongly associated with current smoking, high alcohol consumption, and low FEV1. Mean age in men was 44 years (range 27–61) and in women 42 years (range 28–54). In the studied sample, 95 SAH events (78 in men and 17 in women) occurred during a total of 530,562 person years of follow-up (crude incidence 17.2 per 100,000 person years in men and 22.1 in women).

Higher leucocyte count was associated with increased incidence of SAH, even after adjusting for all covariates: the HR per one standard deviation (SD) increase ( $2.01 \times 10^9$  cells/L) was 1.26 (95% CI 1.05-1.53),  $p=0.014$ . Results were similar in men (HR: 1.28, 95% CI: 1.05-1.55,  $n=17,083$ ) and women (1.37, 0.95-1.97,  $n=2711$ ). In **Table 8**, results are shown by quartiles of leukocytes and per SD increase, in never, former and current smokers. The results were somewhat different for subgroups defined by smoking status; a clear association between leukocytes and SAH was observed among smokers. There was no evidence of effect modification by smoking status ( $p$  for interaction  $>0.2$ ). Numbers of events are low in some of the subgroups. Adjusted models were performed if the number of events was more than 50.

By means of Harrell's c statistics, the ability to predict SAH of the multivariable model increased somewhat after adding leucocytes (0.62 versus 0.64).

**Table 8.** Association between leucocyte count and incidence of SAH in selected subgroup.

	Q1	Q2	Q3	Q4	Per SD increase
<b>All</b>					
N all/n SAH	5110/15	4864/23	4794/27	5026/30	19794/95
HR (95 % CI)	1	1.65 (0.86-3.16)	2.00 (1.06-3.76)	2.22 (1.19-4.12)	1.30 (1.09-1.54)
HR*	1	1.63 (0.85-3.14)	1.92 (1.00-3.66)	2.05 (1.06-3.96)	1.26 (1.05-1.53)
<b>Never smokers</b>					
N all/n SAH	2563/5	1858/10	1317/8	665/2	6403/25
HR	1	2.80 (0.96-8.18)	3.18 (1.04-9.72)	1.60 (0.31-8.24)	1.31 (0.82-2.11)
<b>Former smokers</b>					
N all/n SAH	1249/8	1085/5	783/3	420/2	3555/18
HR	1	0.74 (0.24-2.26)	0.62 (0.16-2.34)	0.81 (0.17-3.80)	0.93 (0.49-1.78)
<b>Current smokers</b>					
N all/n SAH	1288/2	1908/8	2635/16	3913/26	9836/52
HR	1	2.72 (0.58-12.8)	3.88 (0.89-16.9)	4.47 (1.06-18.8)	1.33 (1.07-1.65)
HR*	1	2.73 (0.58-12.8)	3.91 (0.90-17.0)	4.51 (1.07-19.0)	1.32 (1.06-1.63)

\* Adjusted for age, sex, systolic blood pressure, antihypertensive medication, smoking status, high alcohol consumption, body mass index, diabetes, total cholesterol and physical inactivity.

## Paper III

High levels of FGF23 were associated with higher blood pressure and higher leucocyte count, but not with smoking (**Table 5**). There was a positive association between FGF23 and incident SAH. The OR was 3.28 (95% CI, 1.34-8.00) for those with FGF23 in the highest quartile compared to the lowest ( $p$  for trend over quartiles = 0.006), adjusting for smoking, systolic blood pressure, blood pressure medication, body mass index, alcohol intake, and leucocytes. The adjusted OR for SAH per one unit of FGF23 on the log scale was 2.96 (1.25–7.01,  $p=0.014$ ). Because of the nested case control design and the matched analysis, the OR will equal the rate ratio<sup>176, 177</sup>.

## Paper IV

Among COPD patients and reference individuals, there were a total of 17,402 incident strokes during 1,348,973 person years of follow-up (7945 in COPD patients and 9457 in reference individuals). Median follow-up time was 5.1 years (548,938 person years) in COPD patients and 9.7 years (800,034 person years) in reference individuals.

The incidence rate of all-cause stroke was 14.5 (95 % CI 14.2–14.8) per 1000 person years in COPD and 11.8 (11.6–12.1) in reference subjects (12.9, 95% CI 12.7–13.1 in all). COPD was associated with increased incidence of all-cause stroke after adjusting for all covariates (**Table 9**). The HR for all-cause stroke in COPD versus referents was higher during the first 2 years of follow-up (1.46, 95% CI 1.37–1.55) compared with between two and five years (1.15, 1.09–1.22) and after five years of follow-up (1.16, 1.10–1.22), respectively. Incidences of ischemic stroke and ICH were increased in COPD, and these associations showed a similar change over time as for all-cause stroke (**Table 9**). After 5 years of follow-up there was still evidence of a higher incidence of ischemic stroke, but not ICH, in COPD compared to reference subjects. There was evidence of an increased incidence of SAH for COPD patients, and this association did not change over time (adjusted HR 1.46, 95% CI 1.16–1.85).

### Sensitivity analyses

The association between COPD and stroke incidence was similar or somewhat stronger in an analysis including both primary and secondary diagnoses of stroke (n=20,231). HR for all-cause stroke and ischemic stroke was 1.33; 95% CI: 1.29–1.37, and 1.27; 1.22–1.32, after full adjustment.

We also assessed whether both primary and secondary COPD diagnoses were associated with stroke. This analysis showed somewhat higher estimates for stroke in those with COPD as secondary diagnosis (HR all-cause stroke: 1.43; 95% CI: 1.36–1.51 for the whole follow-up) compared to those with COPD as primary diagnosis (HR: 1.13; 1.08–1.18; p for interaction<0.001).

Because all subjects in the exposed, but not in the non-exposed group had been hospitalized (at COPD diagnosis), a sensitivity analysis evaluated the association between COPD and stroke using an in-patient control group, i.e. including only hospitalized reference subjects. We found in this analysis that the HRs for all-cause stroke (1.21; 95% CI: 1.16–1.27), ischemic stroke (1.16; 1.10–1.23), ICH (1.31; 1.13–1.51) and SAH (1.40; 1.00–1.97), were similar compared to the main analysis after full adjustment. There were no significant differences of HRs for all-cause

stroke, ICH or SAH, according to if the reference individuals were inpatients or non-hospitalized subjects from the general population (p for interaction=0.14, 0.95 and 0.37, respectively). For IS, there was moderate evidence of effect modification (p for interaction=0.043).



**Table 9.** The association between COPD and stroke incidence for the whole follow-up and in different time intervals.

		Follow-up 0-2 years		Follow-up 2-5 years		Follow-up 5-10 years		
	COPD	Reference	COPD	Reference	COPD	Reference	COPD	Reference
N	103,419	103,419	103,419	103,419	76,273	95,020	52,206	81,767
All-cause stroke								
No. of events	7,945	9,457	2,734	1,972	2,618	3,106	2,593	4,379
IR, /1000 py	14.5 (14.2-14.8)	11.8 (11.6-12.1)	15.6 (15.0-16.2)	9.9 (9.5-10.4)	13.7 (13.2-14.3)	11.7 (11.3-12.1)	14.2 (13.6-14.7)	13.0 (12.6-13.4)
HR	1.35 (1.31-1.40)	1 (1.51-1.69)	1.60 (1.51-1.69)	1 (1.19-1.33)	1.26 (1.19-1.33)	1 (1.19-1.33)	1.26 (1.20-1.33)	1 (1.20-1.33)
HR <sup>a</sup>	1.24 (1.19-1.28)	1 (1.37-1.55)	1.46 (1.37-1.55)	1 (1.09-1.22)	1.15 (1.09-1.22)	1 (1.09-1.22)	1.16 (1.10-1.22)	1 (1.10-1.22)
Ischemic stroke								
No. of events	5,580	6,855	1,801	1,374	1,801	2,158	1,978	3,323
IR, /1000 py	10.2 (9.9-10.4)	8.6 (8.4-8.8)	10.3 (9.8-10.8)	6.9 (6.6-7.3)	9.5 (9.0-9.9)	8.1 (7.8-8.5)	10.8 (10.3-11.3)	9.9 (9.5-10.2)
HR	1.32 (1.27-1.37)	1 (1.41-1.62)	1.51 (1.41-1.62)	1 (1.16-1.32)	1.24 (1.16-1.32)	1 (1.19-1.35)	1.27 (1.19-1.35)	1 (1.19-1.35)
HR <sup>a</sup>	1.20 (1.15-1.25)	1 (1.27-1.47)	1.37 (1.27-1.47)	1 (1.05-1.19)	1.12 (1.05-1.19)	1 (1.05-1.19)	1.15 (1.08-1.23)	1 (1.08-1.23)
Intracerebral haemorrhage								
No. of events	799	948	269	186	278	318	252	444
IR, per 1000 py	1.46 (1.36-1.56)	1.18 (1.11-1.27)	1.5 (1.4-1.7)	0.9 (0.8-1.2)	1.5 (1.3-1.6)	1.2 (1.1-1.3)	1.4 (1.2-1.6)	1.3 (1.2-1.5)
HR	1.35 (1.22-1.49)	1 (1.36-1.98)	1.64 (1.36-1.98)	1 (1.13-1.58)	1.34 (1.13-1.58)	1 (1.13-1.58)	1.17 (0.99-1.38)	1 (0.99-1.38)
HR <sup>a</sup>	1.29 (1.16-1.43)	1 (1.29-1.89)	1.56 (1.29-1.89)	1 (1.08-1.53)	1.29 (1.08-1.53)	1 (1.08-1.53)	1.12 (0.95-1.33)	1 (0.95-1.33)

<b>Subarachnoid haemorrhage<sup>b</sup></b>										
No. of events	175	168	60	42	52	58	63	68		
IR, per 1000 py	0.32 (0.27-0.37)	0.21 (0.18-0.24)	0.34 (0.27-0.44)	0.21 (0.16-0.29)	0.27 (0.21-0.36)	0.22 (0.17-0.28)	0.34 (0.27-0.44)	0.20 (0.16-0.26)		
HR	1.48 (1.19-1.85)	1	1.58 (1.06-2.34)	1	1.23 (0.84-1.81)	1	1.66 (1.15-2.39)	1		
HR <sup>a</sup>	1.46 (1.16-1.85)	1	1.57 (1.05-2.36)	1	1.20 (0.81-1.79)	1	1.65 (1.13-2.40)	1		
<b>Stroke, not otherwise specified</b>										
No. of events	1,396	1,486	604	370	487	572	302	574		
IR, per 1000 py	2.53 (2.40-2.67)	1.86 (1.77-1.95)	3.4 (3.2-2.7)	1.9 (1.7-2.1)	2.6 (2.3-2.8)	2.2 (2.0-2.3)	1.6 (1.5-1.8)	1.6 (1.5-1.8)		
HR	1.49 (1.37-1.60)	1	1.89 (1.66-2.16)	1	1.32 (1.16-1.49)	1	1.26 (1.08-1.47)	1		
HR <sup>a</sup>	1.35 (1.24-1.47)	1	1.72 (1.50-1.97)	1	1.20 (1.05-1.36)	1	1.15 (0.98-1.35)	1		

IR incidence rate (95 % confidence interval), py person years, HR hazard ratio

<sup>a</sup>Adjusted for country of birth, socioeconomic position, history of asthma, diabetes, hypertension, ischemic heart disease, heart failure, atrial fibrillation, alcohol abuse, rheumatoid arthritis/systemic lupus erythematosus, kidney failure and total duration of hospital stay

<sup>b</sup>Effect of COPD-status not time-dependent

## Paper V

A total of 36,195 SNVs in 7636 genes were included in gene-based analysis and 39,509 SNVs in single variant analysis. No gene or single variant was associated with SAH at the pre-defined significance level. However, some suggestive findings were made. The strongest association in the gene-based analysis was for the *GLE1* gene (RNA export mediator), including seven SNVs. One missense SNV in this gene was suggestively associated with SAH in single-variant analysis (rs138310419,  $p=1.42 \times 10^{-4}$ , MAF=0.022). Of interest was also the finding that the *SOX7* gene was strongly associated with SAH risk in gene-based analysis ( $p=1.15 \times 10^{-4}$ ). This gene is functionally closely related to the *SOX17*, which is one of the previously reported associations for intracranial aneurysms.

A common intron variant in the *EFCA11* (EF-hand calcium binding domain 11) gene on chromosome 14q32.11 (rs8009621,  $p=8 \times 10^{-6}$ , MAF=0.42) and a rare missense SNV in the gene *TRAPPC2L* (trafficking protein particle complex 2-like) on chromosome 16q24.3 (rs114835271,  $p=3.3 \times 10^{-5}$ , MAF=0.0096) had strong associations with SAH in the single variant analysis.

Some of the previously identified risk loci for intracranial aneurysms were included also in the present study. There was a suggested association between the chromosome 18 locus near the *RBBP8* gene (rs11661542,  $p=0.0022$ ) and SAH<sup>87</sup>. The 9p21 risk variants<sup>115</sup> showed associations in the same direction as for intracranial aneurysms (for example, rs10757272, minor allele: T, MAF: 0.47, OR: 1.21, 95% CI: 0.88-1.67,  $p=0.24$ ). This association was stronger in analysis restricted to smokers ( $n=136$ , OR: 1.70, 95% CI: 1.00-2.89,  $p=0.052$ ). The two previously reported loci close to the *SOX17* gene<sup>87</sup> were not associated with SAH in the present study (rs10958409 and rs9298506,  $p>0.6$ ). No exome variants of the *SOX17* gene were available on the exome content of the chip.

# Discussion

SAH is a devastating vascular disease with high mortality and morbidity. Its causes are poorly understood and previous studies of risk factors for SAH are relatively few. In particular, the number of prior population-based prospective studies of risk factors for SAH is small and such studies are needed.

The optimal design of a study of risk factors for SAH has been discussed by Korja et al<sup>60</sup>, suggesting that large prospective unselected population-based cohort studies with a long follow-up time should be used, including broad risk factor data, ideally with measurements of risk factors at several time points during follow-up. Furthermore, all cases of SAH should be included, that is, also cases that die suddenly, and often before reaching the hospital, and results should be stratified by sex and age, given the association of these factors with SAH. Strengths of the studies of SAH in this thesis include some of these points, i.e. the population-based design in two large well-defined cohorts with a long follow-up time, and extensive risk factor data. Our case finding procedure also included cases of SAH dying outside of hospital.

Below, general methodological aspects will be discussed in detail and subsequently, the interpretation and implications of the findings will be discussed in the general discussion part.

## Methodological considerations

### Study designs

A cohort study has the advantage of the prospective assessment of risk factors at baseline and the subsequent follow-up for disease incidence. Thus, different exposures may be compared in relation to disease incidence, which may provide knowledge about risk factors preceding the event and potentially with a causal association to disease. These studies will not be prone to recall bias or bias from taking information from proxies. It is also possible to assess risk prediction models. To study relatively uncommon diseases, like SAH, a large cohort and a long follow-

up time are needed to include a reasonable number of events. The MPP and MDC studies are two such large cohorts providing the opportunity of study also SAH with a reasonable number of events. However, the numbers of events are still low, especially in certain subgroups, and this limits the possibilities of subgroup analyses.

The nested case control study is an efficient way of studying exposure that are not assessed in all participants in a cohort study, for example because of expensive or inconvenient tests. When controls are matched by follow-up time (concurrent sampling), an analysis stratified on matching variables will estimate the relative rate, assuming the rate is constant over follow-up<sup>177</sup>. The nested case control study will thereby equal the results that had been achieved if the whole cohort was analyzed<sup>176</sup>. Using a nested case control approach, the advantages of a cohort study will be gained and at the same time expenses can be cut, for example in studies of biomarkers (for example **Paper III**).

Matching can improve precision and thus efficiency of the study, and in some settings it can adjust for factors that are difficult to measure, for example in neighborhood or sibling controls. Disadvantages include a somewhat more difficult analysis, and that the effect of the matching factors cannot be estimated. There is also a risk of overmatching if matching factors are related to exposure but not to outcome, which will attenuate associations<sup>175</sup>. A matched design requires a “matched” analysis. Matched case control studies (**Paper III**) should be stratified for (conditioned on) the matching variables to avoid bias introduced by forcing cases and controls to be equal on certain characteristics. However, the same does not apply for matched cohort studies<sup>181-183</sup>. In the matched cohort study (**Paper IV**), the exposed and non-exposed groups will have the same characteristics with regard to the matching variables at study start, and no association will exist between exposure and matching variables. In the simple model without adjustment for other factors and little censoring, stratified analysis is therefore not necessary. However, when adjusting for other confounders, associations between these confounders and the matching factors may occur. Also, because of any censoring, the groups may change according to the distribution of the matching factors during follow-up. Therefore, stratifying for the matching variables (or adjusting in some cases) is recommended in the two latter situations<sup>175, 183</sup>. In **Paper IV**, analysis was stratified for matching variables.

## Internal and external validity

The internal validity of a study is determined by the amount of bias and confounding influencing the results and conclusions. The external validity refers to the generalizability of the study results to a broader population than that studied. Bias is the effect of systematic error that, in contrast to random error, will remain even if the sample size is increased. Bias may be divided into many subgroups, most of which will fit under the main groups selection and information<sup>184</sup>.

### Selection bias

Selection bias occurs if those selected for a study are different from those not selected in a systematic way. Selection bias commonly arise from different selection procedures of cases and controls in case control studies, i.e. if controls do not represent the source population that cases came from, or are not selected independently of exposure<sup>168, 175</sup>. Thus, cohort studies have normally not these problems because all subjects (diseased and non-diseased) represent the same source population. However, in cohort studies, selection bias may be caused by differential loss to follow-up, if those lost to follow-up or censored because of death from other causes, differ regarding the risk of having an event compared to those not lost to follow-up in the same risk set (e.g. informative censoring)<sup>185</sup>. In nested case control studies, controls are randomly sampled from the same cohort as cases come from, which do not introduce selection bias<sup>176</sup>.

In **Paper I**, subjects with spirometry data were included. In MPP, 94% of men but only 71% of women were screened with spirometry. In men, whole birth cohorts completed spirometry. All but one birth cohort of men (1926, 999 individuals) had spirometry at screening. Also, in screened, the proportion missing data was very low (1.2%). In women, however, spirometry was not fully confined to whole birth cohorts. On the other hand, whole birth cohorts of women born in 1926, -28, -31, -34, -36, -38, -41, and -49 completed spirometry (3% missing values), representing 83% (6414/7748) of all women with spirometry. A sensitivity analysis restricted to this group showed similar effect estimates for FEV1 and FEV1/FVC as compared to all women.

Leucocyte count was included in the screening protocol in the MPP for one part of the screening period (1974-1981), and **Paper II** studies these individuals. Participants with measurement of leukocyte count after 1981 were not systematically selected for any reason, and subjects with leukocyte count available should be considered a random sample of the cohort.

In **Paper I, II and IV**, censoring occurs during follow-up if subjects emigrate, or die from other causes than SAH. Loss to follow-up was restricted to a few

individuals emigrating from Sweden (n=604, 2%). A total of 7455 (26%) were censored because they died of other causes than SAH. Death from other causes than SAH could be regarded as a competing event, because the subject cannot suffer from a SAH after death from other causes. The influence of competing events is often important in studies of old people or if outcome is strongly related to age, or if there is informative or dependent censoring<sup>186, 187</sup>. However, for analysis of potential etiological associations (as opposed to predictive modeling) calculating the cause-specific hazard, by use of Cox regression with censoring of individuals having a competing event is most appropriate<sup>187</sup>. SAH does not vary strongly with age and occurs also in younger adults. Moreover, as we found a positive association between reduced lung function and SAH in **Paper I**, and hypothetically, a strong competing risk, hindering the event of SAH to occur in those dying of other causes, would most likely tend to obscure or dilute the association between lung function and SAH. A similar reasoning could be applied for results in **Paper II**.

In **Paper IV**, a high proportion of COPD patients die in the first years of follow-up (approximately 50% after five years). However, results are similar, or somewhat lower, in the interval of 5 to 10 years of follow-up compared to the interval 0-2 years. An influence of competing risks cannot be excluded as a reason for the lower estimates of the last part of follow-up when the individuals at high risk of stroke may have died. It is also possible that those susceptible to stroke have already had a stroke in the first years of follow-up. Those remaining in the analysis after five years could be those with mild COPD.

## **Classification of exposure**

### *Non-differential misclassification*

Non-differential misclassification will be present in most epidemiological studies. Misclassification of exposure will be non-differential if it does not depend on the outcome status, and misclassification of outcome will be non-differential if not dependent on exposure. For example, some individuals may be misclassified as having low lung function although they actually have high lung function, because of random variation in the way of performing lung volume tests or a temporary variation in the individual. As this would not be related to suffering from a stroke or not many years later, this misclassification will be non-differential. Non-differential misclassification will generally cause the effect estimates of two exposure categories to be more similar to each other, and almost exclusively cause bias towards the null, and the association studied will be underestimated<sup>168</sup>. In cohort studies this is referred to as regression dilution bias<sup>184, 188</sup>.

Spirometry procedures in the MPP did not follow the current recommendations<sup>189</sup>, because these guidelines were reported after the study protocol was developed

**(Paper I).** Lack of standardized procedures may cause some imprecision of lung volume measurements and some misclassification from higher to lower lung function (non-differential), which will potentially dilute associations. The validity of the lung function variables in MPP is supported by other studies from the MPP, showing strong associations between reduced lung function and pulmonary and cardiovascular diseases<sup>129, 163, 190</sup>.

Acute infections, inflammatory diseases or blood malignancies might increase levels of leucocytes **(Paper II)**. A temporary raise in leukocytes during screening will lead to non-differential misclassification, and because the values later will normalize this results in regression dilution and underestimation of the studied association. In inflammatory diseases leukocytes may be increased and, although uncommon in the general population, we cannot exclude that inflammatory conditions influenced the association between leukocytes and SAH.

#### *Paper IV*

In **Paper IV**, the main exposure was based on a diagnosis of COPD in the inpatient register. The validity of a diagnosis of COPD in this register was recently evaluated, reporting that around 10% of diagnoses are misclassified or uncertain. There was no difference with respect to level of diagnosis (primary or secondary)<sup>162</sup>. People with COPD diagnosed only in primary care may appear in the reference group. However, during follow-up, about 4% of the reference group (i.e the general population) had COPD and were censored for this reason (or changed exposure groups if diagnosed during the inclusion period). Around 3-5% of the general population has physician diagnosed COPD according to studies of Montnemery et al in 1998 and 2006<sup>191, 192</sup>. These figures thus imply that a large part of COPD subjects will be hospitalized over some period of time, for COPD or for other reasons<sup>162</sup>. Occurrence of COPD among reference subjects would bias the association between COPD and stroke towards the null.

## **Classification of outcome**

#### *Paper I-III and V*

Because all patients with diagnosed SAH should be treated in hospital, hospital-based identification of cases of SAH should be appropriate. In the Malmö Stroke Register, all cases of stroke in the Malmö hospital were included, and SAH diagnoses were validated by review of medical records. SAH is a known cause of sudden deaths outside of hospital, and sources where those cases are registered should also be included in studies of SAH. The Causes of death register were used for identifying SAH cases considered to have died from SAH before reaching the hospital or in the emergency department and thus not always registered in hospital-



based registers. These cases accounted for a similar proportion of all cases, as has been seen in other population-based studies,<sup>52, 53</sup> and almost all were based on autopsy. The autopsy frequency in subjects dying unexpectedly outside of hospital has been high in Malmö in the last decades, which means that close to all cases of SAH in this category should have been detected.

Because Malmö residents with SAH and need of neurosurgical care are treated in hospital in Lund, the Inpatient Register was used to identify those that were only registered in hospital in Lund. This might also include some subjects of MPP and MDC that have moved out of Malmö to neighboring municipalities. Diagnostic validity was supported by being treated and diagnosed in the neurosurgical clinic and by review of operations done at SAH. A few cases of SAH, i.e. those occurring before 1989 (MPP) and SAH in subjects that moved outside of Malmö and therefore treated in other hospitals than Malmö/Lund, were identified by use of the Inpatient Register. These cases have not been evaluated, but we decided to include them because, intuitively, we assumed that a hospital diagnosis of SAH is not made without careful clinical investigation, and strong grounds.

Concerning sensitivity, it cannot be excluded, that a few cases of SAH, for example in elderly persons with poor general health status residing in nursing homes, occurred without being subjected to hospital care, as has been the case for other stroke types<sup>44, 193</sup>. However, this will unlikely affect the results. We believe that a large proportion of SAH cases are identified and results thus will represent the whole spectrum of SAH. If any misclassification occurred, this will likely be non-differential, i.e. not different between exposed and non-exposed. Moreover, the age- and sex specific incidence rates in MPP and MDC are similar compared to a south Swedish study evaluating SAH incidence in 1996<sup>27</sup>. The case finding strategy to high extent followed the one recommended by Korja et al for studies of risk factors for SAH in population-based settings<sup>89</sup>.

#### *Paper IV*

Some validation studies have been performed of the diagnosis of stroke in the Inpatient Register, including ischemic stroke and intracerebral hemorrhage. Sensitivity (proportion identified) in the register is high (88-98%) for data from the 1980 to 1990, 2000 and 2002<sup>44, 161, 194-197</sup>. The proportion of acute stroke patients admitted to hospital has varied between studies and the lowest figure is probably that of 84%, in a study from Lund<sup>193</sup>. Of those not admitted to hospital, most were either men with mild stroke, or older patients in nursing homes.

The proportion of correct stroke diagnoses (positive predictive value, PPV) in the registry has differed noticeably between studies. Stegmayr et al<sup>194</sup> found a PPV around 70% between 1985 and 1989 in northern Sweden, whereas the PPV for first-ever stroke was 94% compared to the MONICA study in northern Sweden in

2004<sup>196</sup>. The PPV was lower for recurrent stroke diagnoses<sup>196</sup>. In a validation study from Örebro in 1999, 3% of strokes in the Inpatient Register were misclassified<sup>197</sup>.

In the present study, the population was restricted to participants without previous stroke, and stroke was only included as outcome if listed as primary diagnosis at discharge. A sensitivity analysis with respect to year of inclusion in the study showed similar or stronger results in the last years (1998-2003) when PPV was high for stroke in the Inpatient Register.

In **Paper IV**, we did not cover stroke cases that died outside of hospital and thus were not registered in hospital registers. Based on prior investigations, only around 1-4%<sup>167, 196</sup> of all strokes are expected to be in this category, thus having no substantial impact on results of overall or ischemic stroke. For ICH and SAH, the proportion is higher, so hypothetically this could influence results. However, there is no reason for suspecting that out of hospital deaths from stroke would be more common in non-COPD subjects, perhaps rather the opposite, and missing such cases would likely give stronger or unchanged results for ICH and SAH.

Ascertainment bias can occur if the probability of being diagnosed with an outcome is dependent on exposure. The typical situation in medical epidemiology is when either the exposed or unexposed group is subjected to additional investigations or screening, which will increase the probability of being diagnosed and thus give a falsely high estimate of their risk of the disease. In **Paper IV**, COPD patients have been hospitalized at least once whereas some, but not all, of the reference individuals, has been hospitalized. Theoretically, being hospitalized could increase the probability of being diagnosed with stroke, and if COPD subjects are more frequently hospitalized also during follow-up they may be more subjected to investigations such as brain CT, and thus have increased detection of stroke compared to reference individuals. However, symptoms of stroke are most often apparent and occur suddenly, and all subjects with suspected neurological symptoms will normally undergo investigations. Also, around 90% of stroke patients are admitted to hospital, regardless of having COPD or not<sup>196</sup>.

In our paper, several analytic approaches were applied to address the potential source of ascertainment bias: First, follow-up was started 30 days after the COPD subjects was discharged from hospital, meaning that stroke diagnoses during hospitalization for the index COPD admission, or in the first time after discharge, were recorded as a prevalent stroke and the subject was not included in the analysis. Second, analysis was adjusted for total duration of hospital stay, and we also performed a sensitivity analysis to study if the association between COPD and stroke differs when using a reference group from the general population compared to a hospitalized reference group. The associations between COPD and all stroke subtypes were still significant and estimates were similar to the main analysis, when using an inpatient reference group.

We also included only primary stroke diagnoses during follow-up, which would reduce the risk of including false positive diagnoses in patient admitted primarily for other reasons. Moreover, ascertainment bias will be more likely to occur for “less severe” diagnoses, such as hypertension and hypercholesterolemia. Taken together, we believe it is unlikely that ascertainment bias had an important impact on the association between COPD and stroke.

## Confounding

Confounding is sometimes classified as a type of bias, and sometimes as an own category of methodological issues needed to address for internal validity. Confounding occurs when an association between an exposure and an outcome is influenced of a third factor, so the effects of the studied exposure and the confounder will be mixed. A confounder is by definition a factor associated with outcome and exposure, but not being an intermediate factor on the causal pathway, i.e. it should not be an effect of the exposure. For example, smoking is associated with both reduced lung function and risk of SAH, and is therefore an important potential confounder when studying the association between lung function and SAH. If studying smoking and SAH, and for this example assuming that leucocytes is elevated due to smoking, and in turn is influencing the development of SAH, leucocytes is not a confounder because it is an effect of the exposure on the causal pathway (intermediate factor).

"The research process of learning about and controlling for confounding can be thought of as a walk through a maze toward a central goal. The path through the maze eventually permits the scientist to penetrate into levels that successively get closer to the goal: in [the example of maternal age and Down syndrome] the apparent relations between Down syndrome and birth order can be explained entirely by the effect of mother's age, but that effect in turn will ultimately be explained by other factors that have not yet been identified. As the layers of confounding are left behind, we gradually approach a deeper causal understanding of the underlying biology. Unlike a maze, however, this journey toward biologic understanding does not have a clear endpoint, in the sense that there is always room to understand the biology in a deeper way."

Rothman K. In Epidemiology. An introduction. Oxford University Press. 2002.

### *Choosing confounders for adjusted models*

The effect of a confounder could be handled by stratification, adjusting, matching, and propensity scores<sup>198, 199</sup>. In the present thesis confounding was mainly addressed by adjusting in regression models. Selecting covariates in a regression model in order to control for confounding is often difficult. In causal modeling, the interest is on the effect of the studied association, as compared to predictive modeling in

which the lowest number of covariates explaining most of the variation in outcome is the main goal. For causal modeling, selecting variables based on prior knowledge about the studied exposure and outcome is recommended rather than significance testing or automated selecting methods<sup>199</sup>. Variables very strongly correlated to each other should not be included at the same time (multicollinearity). One suggested approach is to have a set of variables that based on a priori knowledge are decided to be included as covariates, and then add potential additional confounders to the model to evaluate their effect on the estimate of the studied association (change-in-estimate approach). If a factor changes the estimate only marginally it could probably be omitted<sup>199</sup>.

Specific potential confounding effects and how it was addressed in each of the papers in this thesis is elaborated in the general discussion's section.

## **Certain methodological considerations in genetic studies**

### *Type I and II error in genetic analyses*

Type I error are false positive findings, i.e. the error occurring if a study finds an association but there is no true association. In contrast, Type II error is the error if no association is reported although there is actually a true association. If a high number of comparisons are made, the chance of finding a low p value and thus the risk of Type I error is high if standard significance levels are used. For example, if a significance level of  $p < 0.05$  is used and 20 markers are compared between cases and controls, the probability of getting at least one significant results is  $1 - (1 - 0.05)^{20} = 64\%$  due to chance alone. Thus, a lower significance level must be used in order to decrease type I error. However, a conservative adjustment for multiple testing will increase the risk of type II error. The so-called Bonferroni correction adjusts for the number of tests performed. (The significance level will thus be  $0.05/20 = 0.0025$  if 20 tests are performed). Bonferroni correction assumes independence of tested markers. Because of linkage disequilibrium, some degree of dependence is always observed in genome-wide data. In **Paper V**, Bonferroni adjustment was applied to correct for multiple comparisons. This is probably an overcorrection and true associations may be undetected.

### *Analysis of rare coding variants*

Because of the low frequencies of most of exome variants and their biological effects on protein function, analysis is often performed by aggregation of exome variants within the same gene. The combination of their effect can be done in different ways. In situations where the exact effects of tested variants are known, tests can be specifically adapted to fit the biological model. However, if testing data in many genes or exome-wide, this is not known. Burden tests combine the number

of risk alleles assuming the same effect of each allele, whereas the SKAT test can also take into account different directions and magnitudes of association for the individual risk alleles within the gene. The SKAT-O combines the SKAT test and a burden test and calculates their optimal combination, and has been shown to be a good option<sup>179</sup>.

## Generalizability

The MPP and MDC represent the population of Malmö in those years recruited. The MPP had a very high participation rate which is outstanding compared to many later cohort studies. The MDC had a participation rate about 40%. Differences between participants and non-participants of the MPP and MDC as described in the methods section, indicated that these cohorts represent the somewhat healthier part of the population, which is often the case in this kind of population-based studies<sup>154, 156</sup>. The distribution of life-style related factors such as diet, physical activity, alcohol intake and smoking may thus be more equal (less extreme values) than it would have if the attendance rate was 100%. It should be noted, however, that these cohorts are not in any way extremes with respect to healthfulness. For example, the smoking prevalence in MPP was about 45%, and some 40% answered affirmative to being mostly physical inactive in spare time<sup>154</sup>. In MDC, 60% had hypertension at baseline examinations, of which 45% were not treated with anti-hypertensive drugs and 25% were current smokers.

We believe in general results would be generalizable to other populations given that internal validity is high, which is probably more important than the exact representativity when studying potential risk factors. If firm associations are found for a risk factor in a certain population, it is often the case that the result will be generalizable to other populations given a high internal validity, because the biology is essentially the same<sup>168</sup>.

The allele frequencies differ based on ethnicity and genetic findings should be repeated in other ethnicities. COPD cases were taken from the Inpatient Register and whether results could be applied also to COPD diagnosed only in primary care is unclear. However, as argued in the methodological discussion, several points support that this is the case. Apart from these points results should generally be generalizable.

# General Discussion

## Lung function and risk of SAH

We found that moderately reduced lung function is associated with SAH incidence, an association that has not previously been studied, to our knowledge, although associations between reduced lung function and other vascular diseases such as coronary heart disease and ischemic stroke, have been reported. The associations between low FEV1 and low FEV1/FVC, respectively, and SAH remained after adjusting for smoking and in sensitivity analysis of non-smokers. Results for unadjusted lung volumes were similar to that of the standardized lung volumes.

### *Potential mechanisms*

The causes of the association between reduced lung function and cardiovascular disease are largely unknown. Systemic inflammation, or increased ECM degradation has been suggested as possible mechanisms. The analysis was adjusted for erythrocyte sedimentation rate, a measure of systemic inflammation, but more specific inflammatory markers may still be of interest for further evaluation. A common role of increased susceptibility to ECM degradation in both lung tissue destruction and aneurysm and SAH development could be a plausible mechanism. Genetic studies have suggested variants in genes associated with ECM maintenance to be associated with aneurysms and SAH<sup>111</sup>. Reduced lung function is associated with blood pressure variability<sup>200</sup> and this may be another reason for an association between reduced lung function and stroke in general, and perhaps SAH. It is not studied how blood pressure variability influences SAH risk specifically. Future studies should further address underlying causes of the association between lung function and SAH, for example by studying markers of ECM degradation, inflammation and blood pressure variability for SAH risk.

### *Influence of other risk factors*

Smoking is a substantial risk factor for SAH, and an important potential confounder to the association between reduced lung function and SAH. Information about current smoking was available for all individuals in **Paper I-III**. Results were significant in non-smokers. We did not include information about former smoking, which was, by random missing for some birth cohorts. However, when analyzing the subgroup with information about former smoking, results persisted after adjusting for former smoking, and also in subjects reporting never smoking. Further, we did not adjust for the amount of smoked cigarettes (e.g. pack-years), which has been associated with SAH risk<sup>62, 64</sup>. Influence of residual confounding from smoking cannot be fully ruled out. However, the validity of the smoking data in the MPP is

supported by a study showing strong relation to carboxylated hemoglobin levels<sup>201</sup>, and by the strong associations of smoking to other diseases, known to be strongly smoking-related, for example COPD, abdominal aortic aneurysm and peripheral arterial disease<sup>163, 190, 202, 203</sup>. The association between reduced FEV1 and SAH seems to be independent of smoking.

Hypertension is also an important risk factor for SAH. Blood pressure is inversely associated with lung function, both in cross-sectional and longitudinal studies<sup>204, 205</sup>. The association between lung function and SAH remained after adjusting for systolic blood pressure and antihypertensive treatment at baseline, and there was no effect modification by hypertension. Low FEV1 has been associated with future blood pressure increase<sup>205</sup>, and because we do not know whether subjects with low lung function had higher blood pressure increase during follow-up, this may be another potential reason for the association of lung function and SAH.

In consistence with prior results<sup>15, 27</sup>, the incidence of SAH was higher in women than in men, both in the MPP and MDC. Sex differences in risk factors for SAH have been proposed, but the estimates for FEV1 and FEV1/FVC ratio, were similar in men and women.

## **COPD and risk of stroke**

The study showed an association between diagnosis of COPD and incidence of all-cause stroke, ischemic stroke, ICH and SAH, respectively. The increased risk of stroke in COPD patients was highest in the first years of follow-up but persisted during the whole 10-year follow-up period for all-cause stroke, and ischemic stroke. The risk of ICH was clearly increased in the first 5 years of follow-up, whereas the risk of SAH was increased in COPD during 10 years without any change over time.

### *The results in relation to previous studies*

The magnitude of the estimate for total stroke risk in COPD found in this study is consistent with some previous findings<sup>147, 148, 152</sup>, although some studies found no clear association between COPD and stroke<sup>151</sup>. One study found a higher effect estimate, adjusting for smoking and sex<sup>150</sup>. The incidences of each stroke subtype in COPD have not been studied previously.

In line with our study, a recent 10-year follow-up of the Rotterdam study found an increased risk of stroke, both ischemic (HR 1.27; 1.02–1.59, n cases=701) and hemorrhagic (HR 1.70, 95% CI 1.01-2.84, n=107) in COPD<sup>206</sup>. Interestingly they adjusted for cardiovascular risk factors (except smoking), hsCRP, carotid intima-media thickness and oxygen saturation <95%, respectively, and results were substantially unchanged. However, adjusting for smoking, estimates were



attenuated by some 10% (HR 1.13, 95% CI 0.91–1.42 for ischemic and 1.54, 0.91–2.59 for hemorrhagic stroke). The increased incidence of hemorrhagic stroke, which is in line with our study, is interesting because smoking is not an evident risk factor for ICH<sup>18</sup>. Smoking is an important shared risk factor, but these results do not exclude an association between COPD and stroke independent of smoking.

#### *Reasons for the association between COPD and stroke*

It is still not known if the association between COPD and cardiovascular disease is causal, for example through mechanisms induced by COPD such as inflammation and oxidative stress, or if it is explained by residual confounding due to shared risk factors, for example smoking. Increased systemic inflammation has been found in COPD, and suggested as a link between COPD and cardiovascular disease. Low-grade systemic inflammation has also been associated with incidence of ischemic stroke and in this thesis, SAH, whereas the association of inflammation and ICH is uncertain.

The risk of SAH was increased in COPD in the present study over the whole follow-up period. This observation is in consistence with that of **Paper I** showing that low FEV/FVC ratio was associated with SAH. As discussed above ECM degradation or systemic inflammation, could be plausible links between lung function, COPD and SAH. COPD has been associated with atrial fibrillation<sup>146</sup>, which is an important risk factor for ischemic stroke. We adjusted for atrial fibrillation at baseline in the analysis, however, detection of atrial fibrillation and treatment with anticoagulants in COPD patients could be of importance to reduce the risk of stroke.

#### *Hypertension, smoking and exacerbations of COPD*

Large-scale registry-based studies have often the drawback of lacking individual information on all of the potential confounders. Smoking is an important risk factor for stroke, especially ischemic stroke and SAH. Information about smoking was not available, which is an important limitation. As mentioned above, it is of note that COPD was associated also with increased risk of ICH, although smoking is likely not a strong risk factor for ICH. Moreover, the vast majority of COPD patients are current or, most often, former smokers. The comparisons in the present study are not made between a current smoking and a nonsmoking group; population-based studies of COPD in Sweden have reported current smoking rates from 47% (in 1992) to 34% (in 2004) in COPD patients, compared to 33% and 13%, respectively, in non-COPD subjects in the same studies<sup>191, 207</sup>. In secondary care (2007) 23% were still smokers<sup>208</sup> compared to national estimates of around 16% (2005) in the whole population over 45 years of age<sup>209</sup>. In other studies of COPD/reduced lung function and CVD, estimates have been reduced with only around 10% after adjustment for smoking. In the present study, results were adjusted for variables strongly related to smoking status, e.g. socioeconomic status and cardiovascular comorbidities. It



cannot be excluded that smoking is one of the explanations for the association between COPD and stroke, but more studies are needed to elucidate this.

Hypertension is an important risk factor for all types of stroke. The results were adjusted for a diagnosis of hypertension in the Inpatient Register, but not for blood pressure levels, which would have been preferable. COPD has been associated with higher prevalence of hypertension in some studies<sup>210</sup> while others found no association<sup>211</sup>. The association between hypertension and COPD is thus somewhat uncertain and presumably not very strong. Also, smoking is often associated with lower blood pressure in epidemiological studies<sup>70</sup>. It is unlikely that hypertension alone can explain the association between COPD and stroke.

In a study by Donaldson et al<sup>212</sup>, exacerbations of COPD were associated with increased risk of stroke in up to 49 days, when exacerbations were defined as prescription of antibiotics in a large health registry of COPD patients. Another study showed no association between exacerbation frequency and stroke risk<sup>213</sup>. The most recent results from the Rotterdam study showed a clear risk increase of stroke (HR 6.6) after exacerbations compared to stable disease. It is uncertain if exacerbations could explain the increased risk of stroke in the first years observed in our study, because we started follow-up 30 days after discharge for COPD. We could not evaluate whether exacerbations may explain the increased risk of stroke in COPD over 10-years. It is of great interest for further studies to evaluate if exacerbations increase the risk of stroke and if this could be prevented.

From a clinical perspective, the present results show that patients with COPD have increased risk of stroke, regardless of the underlying reason. This is an important message, which should have an impact on the clinical evaluation of a patient with COPD as well as on directions for future research. Evaluation of risk factors for stroke seems important in COPD patients to reduce stroke incidence.

## **Markers related to inflammation and SAH**

Inflammatory activation has been observed in intracranial aneurysms and this phenomenon is more pronounced in ruptured compared to unruptured aneurysms. It is not known if this is an epi-phenomenon or one of the underlying mechanisms for aneurysms and SAH<sup>38</sup>. Some other findings that might point in the direction of an inflammatory component involved in SAH have also been done. For example, polymorphisms of the interleukin 6 gene, a stimulator of inflammatory plasma proteins, have shown association with intracranial aneurysm prevalence in some but not all studies<sup>214</sup>, and an increased risk of SAH has been found in patients with systemic lupus erythematosus<sup>215</sup>. Two prospective studies with small numbers of SAH cases (n=29 and 33, respectively) have examined markers of systemic

inflammation, hsCRP<sup>81</sup> and fibrinogen<sup>82</sup>, but none of these were significantly associated with SAH.

Our results were significant after adjusting for current and former smoking. However, in sensitivity analyses, the association between leukocytes and SAH was clear and significant only in current smokers, although there was no evidence of effect modification by smoking. The hazard ratios were still similar in smokers and never smokers and the numbers of events were small in subgroups. Leukocytes may be elevated as an effect of smoking and thereby perhaps an intermediate factor on the causal pathway between smoking and SAH. Leukocytes could be a marker of susceptibility to SAH in smokers. As discussed above, a high validity of smoking data in MPP has been reported. However, due to the relatively small numbers of events we could not adjust the association for amount of smoked cigarettes, which would have been preferable.

Larger samples need to be studied to find out if leukocytes are associated with SAH in non-smokers. Future studies should evaluate if other markers of inflammation are associated with incidence of SAH. If so, they may be evaluated for risk prediction, for example of the risk of SAH in individuals with known aneurysms, to improve the strategies for offering of preventive intervention.

In **Paper III** an association between FGF23 and SAH was found, adjusting for important confounders. It remains to be shown whether FGF23 is a causal risk factor for SAH or merely a risk marker for vascular diseases. FGF23 has been associated with incidence of ICH in a prior population-based study<sup>85</sup> and also with cardioembolic stroke in another study<sup>86</sup>. FGF23 is also associated with several specific inflammatory markers<sup>84</sup> and regulation of vascular function<sup>88</sup>, processes that may be of importance for SAH.

FGF23 is increased in chronic kidney disease. In a subsample with information about kidney function (n=86) there was no association between estimated glomerular filtration rate and SAH, and it seems unlikely that reduced kidney function could explain the association. FGF23 is involved in vascular calcification and vascular aging together with a factor called Klotho<sup>88</sup>. These are mechanisms potentially of interest for SAH. Interestingly, one of the GWA findings for intracranial aneurysms is a locus near the *KL* gene, encoding Klotho<sup>87</sup>. This variant should be tested for association with FGF23 levels and SAH in future studies.

## Coding genetic variation and SAH

No associations reached the pre-defined significance level in the study of exome variation and SAH. One important limitation is the small sample size, and this study has only the ability to identify strong variants for SAH risk. Larger studies are needed to evaluate if there are coding variants or genes associated with SAH at smaller effect sizes. Of note is that many of the rare variants with low p values for association with SAH had positive beta coefficients, i.e. more common in cases than controls. This is in line with the theory that those rare coding variants have deleterious effects on protein function and is thus associated with disease and not protective of disease. As discussed above, a Bonferroni correction might be too conservative in the present setting.

Gene-based tests (SKAT-O) were performed, which increase the power by combining the effect of many variants throughout a gene. Our sample is homogeneous with respect to the ethnic background and population outliers were excluded before analysis. The genetically homogenous sample will increase the ability of finding associations. We evaluated both aneurysmal and non-aneurysmal SAH together and were not able to separate those categories. Doing so could have increased the precision of our results if there are differences of risk variants between those groups. It is not known, however, to which degree their genetic background differ.

Some variants have been found for intracranial aneurysms. However, there are no previous GWA studies of SAH, to our knowledge. Because only a minority of intracranial aneurysms rupture it is important to study risk factors for SAH, which is the clinical outcome of most interest. Genetic variants related to SAH risk could also, hypothetically, be used in risk prediction of unruptured aneurysms.

Associations for variants at the 9p21 locus was in the same direction for SAH as in prior studies of aneurysms, and the effect was stronger in smokers (n=134), which is in line with previous findings<sup>115</sup>.

Some of the suggestive findings could have pathophysiological mechanisms perhaps related to SAH and are of interest for generating hypotheses. The *SOX7* gene was strongly associated with SAH risk, although not significantly after correction for multiple testing. This gene is a member of the Sox F family transcription factors together with *SOX17*. *SOX17* is a robust GWA finding for intracranial aneurysms<sup>115</sup>. The Sox-F proteins have in part redundant functions on endothelial maintenance and it could be speculated that *SOX7* alterations would be more susceptible to effects of variants in *SOX17*. One way of further evaluating this hypothesis could be a combined score of suboptimal variants within these genes, assessed by targeted sequencing.

# Conclusions

- Reduced lung function at baseline, by means of the FEV1, and FEV1/FVC ratio (airway obstruction), is associated with SAH incidence, independently of smoking.
- COPD patients have increased risks of all-cause stroke, and of all subtypes of stroke compared to a reference group from the general population. Residual confounding from smoking and hypertension cannot be excluded. However, treating stroke risk factors in COPD patients seems to be important to reduce the stroke incidence.
- Higher leukocyte count is associated with incidence of SAH, however, this may be restricted to smokers. The results support that systemic inflammation could be involved in SAH development. Leukocytes may mediate the effect of smoking and/or be a marker of susceptibility for SAH in smokers.
- Serum levels of Fibroblast growth factor 23 are associated with SAH incidence. It should be further studied whether FGF23 is a causal risk factor for SAH or if FGF23 could be used for risk prediction.
- No specific genes or variants were associated with SAH risk at the pre-defined Bonferroni corrected significance level. Further exome analyses should be performed in large samples to evaluate the suggestive findings of this study and to find coding variants with small effect sizes.



# Populärvetenskaplig sammanfattning på svenska

Slaganfall (eller stroke) är en vanlig sjukdom och drabbar ungefär 30,000 personer i Sverige varje år. Det finns tre huvudtyper av stroke varav ischemisk stroke eller blodpropp är den vanligaste och hjärnblödning inom själva hjärnan näst vanligast. Blödning under spindelvävshinnan (araknoidea), som omger hjärnan, kallas för subaraknoidalblödning. Detta är den minst vanliga typen av stroke och drabbar ungefär 1500 i Sverige varje år. Subaraknoidalblödning är en mycket allvarlig sjukdom och 30-50% av drabbade avlider första månaden efter insjuknandet, oftast de första dagarna. Hos överlevande är det vanligt med hjärnskador och symptom som språksvårigheter och försämrat minne. Medelåldern för de som drabbas av SAH är lägre jämfört med de andra stroke-typerna och SAH drabbar fler kvinnor än män. På grund av hög dödlighet, stor risk för restsymtom hos överlevande, och relativt låg ålder vid insjuknande orsakar SAH omfattande lidande och stor grad av bortfall från arbete och vardagsfunktioner, trots att sjukdomen har relativt låg förekomst.

SAH uppstår i de flesta fall när ett pulsåderbräck (aneurysm) på ett av hjärnans blodkärl brister. Kunskapen kring vad som är orsakerna bakom pulsåderbräck och varför dessa ibland orsakar SAH är bristfälligt kända. Rökning är en viktig riskfaktor för SAH. I studier har man föreslagit att inflammatoriska processer och nedbrytning av blodkärlen stödjevävnad kan vara av betydelse. Försämrade lungfunktion och sjukdomen KOL (kroniskt obstruktiv lungsjukdom) är tillstånd som utmärks av nedbrytning av stödjevävnad i lungorna. Låg lungfunktion och KOL har visats kopplade till risk för hjärtinfarkt och ischemisk stroke i tidigare studier även hos personer som inte röker. Det är inte studerat ifall inflammationsmarkörer som kan mätas i blodet är kopplade till risken för att få SAH i framtiden.

I avhandlingen studeras hur låg lungfunktion påverkar risken att drabbas av SAH hos personer som genomgick lungfunktions-mätningar (spirometri) i en befolkningsundersökning i Malmö mellan 1974 och 1992. Insjuknande i SAH har registrerats sedan studiens start fram till 2008. Risken för SAH och de andra stroke-typerna hos patienter som diagnostiserats med KOL, studeras i ett annat material som baseras på diagnoser i det nationella slutenvårdsregistret, där alla diagnoser vid

utskrivning från sjukhus finns. Vita blodkroppar i blodet som är relaterade till inflammation och en annan plasmamarkör som är kopplad inflammation, studeras också i förhållande till risk att få SAH och detta görs inom en annan stor befolkningsstudie från Malmö där individer undersökts och sedan följts upp avseende insjuknande i SAH. Vi undersöker i avhandlingen också genetiska varianter som kan påverka risken för SAH.

Resultaten visar att låg lungfunktion är kopplat till en högre risk att få SAH, och detta samband kan ses också hos icke-rökare. Personer med KOL har en förhöjd risk att drabbas av alla typerna av stroke jämfört med personer från befolkningen som inte har KOL, även när man tar hänsyn till ett stort antal andra sjukdomar som är vanligare hos KOL-patienter. Vi har också visat att inflammatoriska markörer är associerade med högre risk för SAH. Vi kunde inte hitta några säkra starka genetiska varianter kopplade till risk för SAH förutom att vi såg samband för några av de redan kända riskvarianterna, som hittats i tidigare studier. Dessa har ganska svaga samband med SAH men tros påverka risken tillsammans med andra riskfaktorer.

Fynden i denna avhandling indikerar att nedbrytning av stödjevävnad respektive inflammation i blodet eller i blodkärlen, kan vara mekanismer som är av stor vikt för utveckling av SAH. Om vidare studier kan identifiera specifika ämnen eller mekanismer i denna process kan det på sikt leda till nya sätt att förebygga eller behandla SAH. Resultaten visar också att patienter med KOL är en särskild riskgrupp för SAH och även de andra typerna av stroke, och detta bör leda till att särskild uppmärksamhet ges till förebyggande behandling av riskfaktorer för stroke inom denna grupp. Det behöver vidare studeras vad orsaken till sambanden mellan låg lungfunktion och SAH respektive KOL och stroke beror på. Fynden i denna avhandling kan på sikt bidra till att fler fall av SAH kan förhindras, till exempel genom förebyggande insatser till särskilda riskgrupper och att kunskapen kring sjukdomsmekanismerna klargörs.

# Acknowledgements

I would like to warmly thank all those who have helped me in the work with this thesis. In particular, I would like to thank:

Gunnar Engström, my supervisor, for solid support in all research related questions, a lot of encouragement, great knowledge about stroke epidemiology, and always being available for answering questions, commenting or discussing with an open mind subjects that have struck me when working with the thesis, may it be far-fetched hypotheses or a range of numbers. Your clear mentorship has provided a lot to this thesis and to my advance during the course.

Elisabet Zia, my supervisor, for bringing me into research in stroke already as a medical student, and for introducing me to neurology. Thank you also for a lot of helpful advice on varying things, and valuable input about stroke.

Malin Inghammar, co-author, for thorough knowledge about the principles of epidemiologic analysis, and for encouragement in my (sometimes tiresome) work with combining and understanding data from different registers.

Bo Hedblad for enthusiastic encouragement in cardiovascular research, careful review and valuable comments of my papers.

Ingela Jerntorp for careful work with the Malmö Stroke Register for many years.

Arne Egesten for valuable work with and comments on the COPD-stroke study.

Arne Lindgren, Christina Jern, Maja Olsson, Erik Lorentzen, Katarina Jood, Tara Stanne, Olle Melander and Peter Almgren for great collaboration in the South Swedish GWAS group, encouragement, and contribution with your broad knowledge in the fields of stroke genetics, on a epidemiological, molecular and analytic levels, helping me to get into stroke genetics.

Colleagues in the research group, Yan Borné, John Berntsson and Gustav Smith, for important comments and interesting discussions about all kinds of topics.

Maria Rosvall and Arkadiusz Siennicki-Lantz for input in my half time seminar.

All colleagues at the neurology clinic for valuable support, clinical inspiration and help.



Christer Nilsson and Jesper Petersson, at the neurology clinic, for letting research take place and time in the clinic.

Helene Pessah-Rasmussen for your encouragement of and interest in my thesis.

Ulf Ekelund, supervisor of my first research experience as a medical student summer project, for challenging ‘truths’ and doing research in an inspiring way.

All participants of the MPP and MDC. Anders Dahlin, Håkan Andersson and Kerstin Andersson for data handling.

My family for always supporting me.

Cecilia, Astrid and Julie, for all your love, patience and curiosity.

# References

1. Snow J. On the mode of communication of cholera, 2nd edition. John Churchill. London. 1860.
2. van Gijn J, Kerr RS, Rinkel GJ. Subarachnoid haemorrhage. *Lancet*. 2007;369:306-318
3. Johnston SC, Colford JM, Jr., Gress DR. Oral contraceptives and the risk of subarachnoid hemorrhage: A meta-analysis. *Neurology*. 1998;51:411-418
4. Feigin VL, Lawes CM, Bennett DA, et al. Worldwide stroke incidence and early case fatality reported in 56 population-based studies: A systematic review. *Lancet Neurol*. 2009;8:355-369
5. Asplund K, Tuomilehto J, Stegmayr B, et al. Diagnostic criteria and quality control of the registration of stroke events in the MONICA project. *Acta Med Scand Suppl*. 1988;728:26-39
6. Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease study 2010. *Lancet*. 2012;380:2197-2223
7. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease study 2010. *Lancet*. 2012;380:2095-2128
8. Feigin VL, Forouzanfar MH, Krishnamurthi R, et al. Global and regional burden of stroke during 1990-2010: Findings from the Global Burden of Disease study 2010. *Lancet*. 2014;383:245-254
9. Stegmayr BA, K. Ökad överlevnad vid stroke men oförändrad risk att insjukna. *Läkartidningen*. 2003
10. Appelros P, Stegmayr B, Terent A. Sex differences in stroke epidemiology: A systematic review. *Stroke*. 2009;40:1082-1090
11. Natinal Board on Health and Welfare. Statistikdatabas för diagnoser i slutenvården. <http://www.socialstyrelsen.se/statistik/statistikdatabas/diagnoserislutenvard>. Access date September 30, 2016.
12. Vaartjes I, O'Flaherty M, Capewell S, et al. Remarkable decline in ischemic stroke mortality is not matched by changes in incidence. *Stroke*. 2013;44:591-597
13. Nieuwkamp DJ, Setz LE, Algra A, et al. Changes in case fatality of aneurysmal subarachnoid haemorrhage over time, according to age, sex, and region: A meta-analysis. *Lancet Neurol*. 2009;8:635-642
14. Zia E, Engström G, Svensson PJ, et al. Three-year survival and stroke recurrence rates in patients with primary intracerebral hemorrhage. *Stroke*. 2009;40:3567-3573
15. de Rooij NK, Linn FH, van der Plas JA, et al. Incidence of subarachnoid haemorrhage: A systematic review with emphasis on region, age, gender and time trends. *J Neurol Neurosurg Psychiatry*. 2007;78:1365-1372
16. Zia E, Pessah-Rasmussen H, Khan FA, et al. Risk factors for primary intracerebral hemorrhage: A population-based nested case-control study. *Cerebrovasc Dis*. 2006;21:18-25

17. Feigin VL, Rinkel GJ, Lawes CM, et al. Risk factors for subarachnoid hemorrhage: An updated systematic review of epidemiological studies. *Stroke*. 2005;36:2773-2780
18. O'Donnell MJ, Chin SL, Rangarajan S, et al. Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): A case-control study. *Lancet*. 2016;388:761-775
19. Linn FH, Rinkel GJ, Algra A, et al. Headache characteristics in subarachnoid haemorrhage and benign thunderclap headache. *J Neurol Neurosurg Psychiatry*. 1998;65:791-793
20. Beetham R, UK Immunohistochemistry working group. Recommendations for CSF analysis in subarachnoid haemorrhage. *J Neurol Neurosurg Psychiatry*. 2004;75:528
21. Cruickshank A, Auld P, Beetham R, et al. Revised national guidelines for analysis of cerebrospinal fluid for bilirubin in suspected subarachnoid haemorrhage. *Ann Clin Biochem*. 2008;45:238-244
22. Rinkel GJ. Management of patients with aneurysmal subarachnoid haemorrhage. *Curr Opin Neurol*. 2016;29:37-41
23. Perry JJ, Stiell IG, Sivilotti ML, et al. Sensitivity of computed tomography performed within six hours of onset of headache for diagnosis of subarachnoid haemorrhage: Prospective cohort study. *BMJ*. 2011;343:d4277
24. National Board on Health and Welfare. Socialstyrelsen. Nationella riktlinjer för strokesjukvård 2014. Available at: <https://http://www.socialstyrelsen.se/SiteCollectionDocuments/nr-stroke-vetenskapligt-underlag-2014.pdf>. Accessed August 10, 2016.
25. Steiner T, Juvela S, Unterberg A, et al. European Stroke Organization guidelines for the management of intracranial aneurysms and subarachnoid haemorrhage. *Cerebrovasc Dis*. 2013;35:93-112
26. Ronkainen A, Hernesniemi J. Subarachnoid haemorrhage of unknown aetiology. *Acta Neurochir (Wien)*. 1992;119:29-34
27. Nilsson OG, Lindgren A, Ståhl N, et al. Incidence of intracerebral and subarachnoid haemorrhage in southern sweden. *J Neurol Neurosurg Psychiatry*. 2000;69:601-607
28. Epidemiology of aneurysmal subarachnoid hemorrhage in Australia and New Zealand: Incidence and case fatality from the Australasian Cooperative Research On Subarachnoid Hemorrhage study (ACROSS). *Stroke*. 2000;31:1843-1850
29. van Gijn J, van Dongen KJ, Vermeulen M, et al. Perimesencephalic hemorrhage: A nonaneurysmal and benign form of subarachnoid hemorrhage. *Neurology*. 1985;35:493-497
30. Keedy A. An overview of intracranial aneurysms. *McGill J Med*. 2006;9:141-146.
31. Vlak MH, Algra A, Brandenburg R, et al. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: A systematic review and meta-analysis. *Lancet Neurol*. 2011;10:626-636
32. Wermer MJ, van der Schaaf IC, Algra A, et al. Risk of rupture of unruptured intracranial aneurysms in relation to patient and aneurysm characteristics: An updated meta-analysis. *Stroke*. 2007;38:1404-1410
33. Rinkel GJ, Djibuti M, Algra A, et al. Prevalence and risk of rupture of intracranial aneurysms: A systematic review. *Stroke*. 1998;29:251-256
34. Wiebers DO, Whisnant JP, Huston J, 3rd, et al. Unruptured intracranial aneurysms: Natural history, clinical outcome, and risks of surgical and endovascular treatment. *Lancet*. 2003;362:103-110
35. Tulamo R, Frösen J, Hernesniemi J, et al. Inflammatory changes in the aneurysm wall: A review. *J Neurointervent Surgery*. 2010;2:120e130.

36. Frösen J, Tulamo R, Paetau A, et al. Saccular intracranial aneurysm: Pathology and mechanisms. *Acta Neuropathol.* 2012;123:773-786
37. Scanarini M, Mingrino S, Giordano R, et al. Histological and ultrastructural study of intracranial saccular aneurysmal wall. *Acta Neurochir (Wien).* 1978;43:171-182
38. Frösen J, Piippo A, Paetau A, et al. Remodeling of saccular cerebral artery aneurysm wall is associated with rupture: Histological analysis of 24 unruptured and 42 ruptured cases. *Stroke.* 2004;35:2287-2293
39. Kataoka K, Taneda M, Asai T, et al. Structural fragility and inflammatory response of ruptured cerebral aneurysms. A comparative study between ruptured and unruptured cerebral aneurysms. *Stroke.* 1999;30:1396-1401
40. Chyatte D, Lewis I. Gelatinase activity and the occurrence of cerebral aneurysms. *Stroke.* 1997;28:799-804
41. Bruno G, Todor R, Lewis I, et al. Vascular extracellular matrix remodeling in cerebral aneurysms. *J Neurosurg.* 1998;89:431-440
42. Kosierkiewicz TA, Factor SM, Dickson DW. Immunocytochemical studies of atherosclerotic lesions of cerebral berry aneurysms. *J Neuropathol Exp Neurol.* 1994;53:399-406
43. Marchese E, Vignati A, Albanese A, et al. Comparative evaluation of genome-wide gene expression profiles in ruptured and unruptured human intracranial aneurysms. *J Biol Regul Homeost Agents.* 2010;24:185-195
44. Appelros P, Nydevik I, Seiger A, et al. High incidence rates of stroke in Örebro, Sweden: Further support for regional incidence differences within Scandinavia. *Cerebrovasc Dis.* 2002;14:161-168
45. Jerntorp P, Berglund G. Stroke registry in Malmö, Sweden. *Stroke.* 1992;23:357-361
46. Norrving B, Löwenhielm P. Epidemiology of stroke in Lund-Örup, Sweden, 1983-85. Incidence of first stroke and age-related changes in subtypes. *Acta Neurol Scand.* 1988;78:408-413
47. Ingall T, Asplund K, Mahonen M, et al. A multinational comparison of subarachnoid hemorrhage epidemiology in the WHO MONICA stroke study. *Stroke.* 2000;31:1054-1061
48. Korja M, Lehto H, Juvela S, et al. Incidence of subarachnoid hemorrhage is decreasing together with decreasing smoking rates. *Neurology.* 2016
49. Linn FH, Rinkel GJ, Algra A, et al. Incidence of subarachnoid hemorrhage: Role of region, year, and rate of computed tomography: A meta-analysis. *Stroke.* 1996;27:625-629
50. Stegmayr B, Eriksson M, Asplund K. Declining mortality from subarachnoid hemorrhage: Changes in incidence and case fatality from 1985 through 2000. *Stroke.* 2004;35:2059-2063
51. Koffijberg H, Buskens E, Granath F, et al. Subarachnoid haemorrhage in Sweden 1987-2002: Regional incidence and case fatality rates. *J Neurol Neurosurg Psychiatry.* 2008;79:294-299
52. Korja M, Silventoinen K, Laatikainen T, et al. Cause-specific mortality of 1-year survivors of subarachnoid hemorrhage. *Neurology.* 2013;80:481-486
53. Huang J, van Gelder JM. The probability of sudden death from rupture of intracranial aneurysms: A meta-analysis. *Neurosurgery.* 2002;51:1101-1105
54. Nieuwkamp DJ, Algra A, Blomqvist P, et al. Excess mortality and cardiovascular events in patients surviving subarachnoid hemorrhage: A nationwide study in Sweden. *Stroke.* 2011
55. Al-Khindi T, Macdonald RL, Schweizer TA. Cognitive and functional outcome after aneurysmal subarachnoid hemorrhage. *Stroke.* 2010;41:e519-536
56. Johnston SC, Selvin S, Gress DR. The burden, trends, and demographics of mortality from subarachnoid hemorrhage. *Neurology.* 1998;50:1413-1418

57. Mhurchu CN, Anderson C, Jamrozik K, et al. Hormonal factors and risk of aneurysmal subarachnoid hemorrhage: An international population-based, case-control study. *Stroke*. 2001;32:606-612
58. Lindekleiv H, Sandvei MS, Njolstad I, et al. Sex differences in risk factors for aneurysmal subarachnoid hemorrhage: A cohort study. *Neurology*. 2011;76:637-643
59. Sandvei MS, Romundstad PR, Muller TB, et al. Risk factors for aneurysmal subarachnoid hemorrhage in a prospective population study: The HUNT study in Norway. *Stroke*. 2009;40:1958-1962
60. Korja M, Silventoinen K, Laatikainen T, et al. Risk factors and their combined effects on the incidence rate of subarachnoid hemorrhage--a population-based cohort study. *PLoS One*. 2013;8:e73760
61. Jimenez-Yepes CM, Londono-Fernandez JL. Risk of aneurysmal subarachnoid hemorrhage: The role of confirmed hypertension. *Stroke*. 2008;39:1344-1346
62. Anderson CS, Feigin V, Bennett D, et al. Active and passive smoking and the risk of subarachnoid hemorrhage: An international population-based case-control study. *Stroke*. 2004;35:633-637
63. Knekt P, Reunanen A, Aho K, et al. Risk factors for subarachnoid hemorrhage in a longitudinal population study. *J Clin Epidemiol*. 1991;44:933-939
64. Kawachi I, Colditz GA, Stampfer MJ, et al. Smoking cessation and decreased risk of stroke in women. *JAMA*. 1993;269:232-236
65. Ruigrok YM, Buskens E, Rinkel GJ. Attributable risk of common and rare determinants of subarachnoid hemorrhage. *Stroke*. 2001;32:1173-1175
66. Kissela BM, Sauerbeck L, Woo D, et al. Subarachnoid hemorrhage: A preventable disease with a heritable component. *Stroke*. 2002;33:1321-1326
67. Longstreth WT, Jr., Nelson LM, Koepsell TD, et al. Cigarette smoking, alcohol use, and subarachnoid hemorrhage. *Stroke*. 1992;23:1242-1249
68. Sundström J, Svennblad B, Söderholm M, Söderberg S, Eriksson M, Malfert M, Heller S, Engström G, Wiber B. Risk factors for subarachnoid hemorrhage. Abstract, accepted for American Heart Association Scientific Sessions November 2016.
69. Tartara F, Gaetani P, Tancioni F, et al. Alpha 1-antitrypsin activity in subarachnoid hemorrhage. *Life Sci*. 1996;59:15-20
70. Omvik P. How smoking affects blood pressure. *Blood Press*. 1996;5:71-77
71. Sankai T, Iso H, Shimamoto T, et al. Prospective study on alcohol intake and risk of subarachnoid hemorrhage among Japanese men and women. *Alcohol Clin Exp Res*. 2000;24:386-389
72. Feigin V, Parag V, Lawes CM, et al. Smoking and elevated blood pressure are the most important risk factors for subarachnoid hemorrhage in the Asia-Pacific region: An overview of 26 cohorts involving 306,620 participants. *Stroke*. 2005;36:1360-1365
73. Suh I, Jee SH, Kim HC, et al. Low serum cholesterol and haemorrhagic stroke in men: Korea medical insurance corporation study. *Lancet*. 2001;357:922-925
74. Lindbohm JV, Kaprio J, Korja M. Cholesterol as a risk factor for subarachnoid hemorrhage: A systematic review. *PLoS One*. 2016;11:e0152568
75. Adamson J, Humphries SE, Ostergaard JR, et al. Are cerebral aneurysms atherosclerotic? *Stroke*. 1994;25:963-966
76. Zhang Y, Tuomilehto J, Jousilahti P, et al. Total and high-density lipoprotein cholesterol and stroke risk. *Stroke*. 2012;43:1768-1774

77. Shiue I, Arima H, Hankey GJ, et al. Dietary intake of key nutrients and subarachnoid hemorrhage: A population-based case-control study in Australasia. *Cerebrovasc Dis.* 2011;31:464-470
78. Okamoto K. Habitual green tea consumption and risk of an aneurysmal rupture subarachnoid hemorrhage: A case-control study in Nagoya, Japan. *Eur J Epidemiol.* 2006;21:367-371
79. Anderson C, Ni Mhurchu C, Scott D, et al. Triggers of subarachnoid hemorrhage: Role of physical exertion, smoking, and alcohol in the Australasian Cooperative Research on subarachnoid hemorrhage study (ACROSS). *Stroke.* 2003;34:1771-1776
80. Hu G, Sarti C, Jousilahti P, et al. Leisure time, occupational, and commuting physical activity and the risk of stroke. *Stroke.* 2005;36:1994-1999
81. Chei CL, Yamagishi K, Kitamura A, et al. C-reactive protein levels and risk of stroke and its subtype in Japanese: The Circulatory Risk in Communities Study (CIRCS). *Atherosclerosis.* 2011;217:187-193
82. Sato S, Iso H, Noda H, et al. Plasma fibrinogen concentrations and risk of stroke and its subtypes among Japanese men and women. *Stroke.* 2006;37:2488-2492
83. Isakova T, Xie H, Yang W, et al. Fibroblast growth factor 23 and risks of mortality and end-stage renal disease in patients with chronic kidney disease. *JAMA.* 2011;305:2432-2439
84. Munoz Mendoza J, Isakova T, Ricardo AC, et al. Fibroblast growth factor 23 and inflammation in CKD. *Clin J Am Soc Nephrol.* 2012;7:1155-1162
85. Wright CB, Dong C, Stark M, et al. Plasma fgf23 and the risk of stroke: The Northern Manhattan Study (NOMAS). *Neurology.* 2014
86. Panwar B, Jenny NS, Howard VJ, et al. Fibroblast growth factor 23 and risk of incident stroke in community-living adults. *Stroke.* 2015;46:322-328
87. Yasuno K, Bilguvar K, Bijlenga P, et al. Genome-wide association study of intracranial aneurysm identifies three new risk loci. *Nat Genet.* 2010;42:420-425
88. Lim K, Lu TS, Molostvov G, et al. Vascular klotho deficiency potentiates the development of human artery calcification and mediates resistance to fibroblast growth factor 23. *Circulation.* 2012;125:2243-2255
89. Korja M, Kaprio J. Controversies in epidemiology of intracranial aneurysms and SAH. *Nat Rev Neurol.* 2016;12:50-55
90. Roos YB, de Haan RJ, Beenen LF, et al. Complications and outcome in patients with aneurysmal subarachnoid haemorrhage: A prospective hospital based cohort study in the Netherlands. *J Neurol Neurosurg Psychiatry.* 2000;68:337-341
91. Frontera JA, Ahmed W, Zach V, et al. Acute ischaemia after subarachnoid haemorrhage, relationship with early brain injury and impact on outcome: A prospective quantitative MRI study. *J Neurol Neurosurg Psychiatry.* 2015;86:71-78
92. Dorhout Mees SM, Rinkel GJ, Feigin VL, et al. Calcium antagonists for aneurysmal subarachnoid haemorrhage. *Cochrane Database Syst Rev.* 2007;CD000277
93. Germans MR, Coert BA, Vandertop WP, et al. Time intervals from subarachnoid hemorrhage to rebleed. *J Neurol.* 2014;261:1425-1431
94. Molyneux AJ, Kerr RS, Birks J, et al. Risk of recurrent subarachnoid haemorrhage, death, or dependence and standardised mortality ratios after clipping or coiling of an intracranial aneurysm in the International Subarachnoid Aneurysm trial (ISAT): Long-term follow-up. *Lancet Neurol.* 2009;8:427-433
95. Hillman J, Fridriksson S, Nilsson O, et al. Immediate administration of tranexamic acid and reduced incidence of early rebleeding after aneurysmal subarachnoid hemorrhage: A prospective randomized study. *J Neurosurg.* 2002;97:771-778

96. Baharoglu MI, Germans MR, Rinkel GJ, et al. Antifibrinolytic therapy for aneurysmal subarachnoid haemorrhage. *Cochrane Database Syst Rev*. 2013;CD001245
97. Germans MR, Post R, Coert BA, et al. Ultra-early tranexamic acid after subarachnoid hemorrhage (ULTRA): Study protocol for a randomized controlled trial. *Trials*. 2013;14:143
98. Bromberg JE, Rinkel GJ, Algra A, et al. Subarachnoid haemorrhage in first and second degree relatives of patients with subarachnoid haemorrhage. *BMJ*. 1995;311:288-289
99. Sundquist J, Li X, Sundquist K, et al. Risks of subarachnoid hemorrhage in siblings: A nationwide epidemiological study from sweden. *Neuroepidemiology*. 2007;29:178-184
100. Bor AS, Rinkel GJ, Adami J, et al. Risk of subarachnoid haemorrhage according to number of affected relatives: A population based case-control study. *Brain*. 2008;131:2662-2665
101. Gaist D, Vaeth M, Tsiropoulos I, et al. Risk of subarachnoid haemorrhage in first degree relatives of patients with subarachnoid haemorrhage: Follow up study based on national registries in Denmark. *BMJ*. 2000;320:141-145
102. Brown RD, Jr., Huston J, Hornung R, et al. Screening for brain aneurysm in the familial intracranial aneurysm study: Frequency and predictors of lesion detection. *J Neurosurg*. 2008;108:1132-1138
103. Rothman KJ. Six persistent research misconceptions. *J Gen Intern Med*. 2014;29:1060-1064
104. Bevan S, Traylor M, Adib-Samii P, et al. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. *Stroke*. 2012;43:3161-3167
105. Brass LM, Isaacssohn JL, Merikangas KR, et al. A study of twins and stroke. *Stroke*. 1992;23:221-223
106. Bak S, Gaist D, Sindrup SH, et al. Genetic liability in stroke: A long-term follow-up study of Danish twins. *Stroke*. 2002;33:769-774
107. Peden JF, Farrall M. Thirty-five common variants for coronary artery disease: The fruits of much collaborative labour. *Hum Mol Genet*. 2011;20:R198-205
108. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med*. 2000;343:78-85
109. MacArthur DG, Balasubramanian S, Frankish A, et al. A systematic survey of loss-of-function variants in human protein-coding genes. *Science*. 2012;335:823-828
110. Gianfagna F, Cugino D, Santimone I, et al. From candidate gene to genome-wide association studies in cardiovascular disease. *Thromb Res*. 2012;129:320-324
111. Alg VS, Sofat R, Houlden H, et al. Genetic risk factors for intracranial aneurysms: A meta-analysis in more than 116,000 individuals. *Neurology*. 2013;80:2154-2165
112. Yasuno K, Bakircioglu M, Low SK, et al. Common variant near the endothelin receptor type a (EDNRA) gene is associated with intracranial aneurysm risk. *Proc Natl Acad Sci U S A*. 2011;108:19707-19712
113. Olsson S, Csajbok LZ, Jood K, et al. Association between genetic variation on chromosome 9p21 and aneurysmal subarachnoid haemorrhage. *J Neurol Neurosurg Psychiatry*. 2011;82:384-388
114. Helgadottir A, Thorleifsson G, Magnusson KP, et al. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet*. 2008;40:217-224
115. Foroud T, Koller DL, Lai D, et al. Genome-wide association study of intracranial aneurysms confirms role of ANRIL and SOX17 in disease risk. *Stroke*. 2012;43:2846-2852

116. Woo D, Khoury J, Haverbusch MM, et al. Smoking and family history and risk of aneurysmal subarachnoid hemorrhage. *Neurology*. 2009;72:69-72
117. Bilguvar K, Yasuno K, Niemela M, et al. Susceptibility loci for intracranial aneurysm in European and Japanese populations. *Nat Genet*. 2008;40:1472-1477
118. Foroud T, Lai D, Koller D, et al. Genome-wide association study of intracranial aneurysm identifies a new association on chromosome 7. *Stroke*. 2014;45:3194-3199
119. Visel A, Zhu Y, May D, et al. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature*. 2010;464:409-412
120. Motterle A, Pu X, Wood H, et al. Functional analyses of coronary artery disease associated variation on chromosome 9p21 in vascular smooth muscle cells. *Hum Mol Genet*. 2012;21:4021-4029
121. Almontashiri NA, Antoine D, Zhou X, et al. 9p21.3 coronary artery disease risk variants disrupt TEAD transcription factor-dependent TGF-beta regulation of p16 expression in human aortic smooth muscle cells. *Circulation*. 2015;132:1969-1978
122. Leeper NJ, Raiesdana A, Kojima Y, et al. Loss of CDKN2B promotes p53-dependent smooth muscle cell apoptosis and aneurysm formation. *Arterioscler Thromb Vasc Biol*. 2013;33:e1-e10
123. Kondo S, Hashimoto N, Kikuchi H, et al. Apoptosis of medial smooth muscle cells in the development of saccular cerebral aneurysms in rats. *Stroke*. 1998;29:181-188.
124. Lee S, Kim IK, Ahn JS, et al. Deficiency of endothelium-specific transcription factor SOX17 induces intracranial aneurysm. *Circulation*. 2015;131:995-1005
125. Illumina, Ltd. Data sheet. DNA analysis in human exome beadchips. Available at: [http://res.illumina.com/documents/products/datasheets/datasheet\\_humanexome\\_beadchips.pdf](http://res.illumina.com/documents/products/datasheets/datasheet_humanexome_beadchips.pdf). Accessed: April 25 2015.
126. Ariesen MJ, Claus SP, Rinkel GJ, et al. Risk factors for intracerebral hemorrhage in the general population: A systematic review. *Stroke*. 2003;34:2060-2065
127. Zia E, Hedblad B, Pessah-Rasmussen H, et al. Blood pressure in relation to the incidence of cerebral infarction and intracerebral hemorrhage. Hypertensive hemorrhage: Debated nomenclature is still relevant. *Stroke*. 2007;38:2681-2685
128. Johnson LS, Juhlin T, Engström G, et al. Reduced forced expiratory volume is associated with increased incidence of atrial fibrillation: The Malmö Preventive Project. *Europace*. 2013
129. Engström G, Melander O, Hedblad B. Population-based study of lung function and incidence of heart failure hospitalisations. *Thorax*. 2010;65:633-638
130. Hole DJ, Watt GC, Davey-Smith G, et al. Impaired lung function and mortality risk in men and women: Findings from the Renfrew and Paisley prospective population study. *BMJ*. 1996;313:711-715; discussion 715-716
131. Lange P, Nyboe J, Appleyard M, et al. Spirometric findings and mortality in never-smokers. *J Clin Epidemiol*. 1990;43:867-873
132. Sin DD, Wu L, Man SF. The relationship between reduced lung function and cardiovascular mortality: A population-based study and a systematic review of the literature. *Chest*. 2005;127:1952-1959
133. Agnarsson U, Thorgeirsson G, Sigvaldason H, et al. Effects of leisure-time physical activity and ventilatory function on risk for stroke in men: The Reykjavik study. *Ann Intern Med*. 1999;130:987-990
134. Hozawa A, Billings JL, Shahar E, et al. Lung function and ischemic stroke incidence: The Atherosclerosis Risk in Communities Study. *Chest*. 2006;130:1642-1649



135. Truelsen T, Prescott E, Lange P, et al. Lung function and risk of fatal and non-fatal stroke. The Copenhagen city Heart study. *Int J Epidemiol.* 2001;30:145-151
136. Wannamethee SG, Shaper AG, Ebrahim S. Respiratory function and risk of stroke. *Stroke.* 1995;26:2004-2010
137. Engström G, Hedblad B, Valind S, et al. Increased incidence of myocardial infarction and stroke in hypertensive men with reduced lung function. *J Hypertens.* 2001;19:295-301
138. Schanen JG, Iribarren C, Shahar E, et al. Asthma and incident cardiovascular disease: The Atherosclerosis risk in Communities study. *Thorax.* 2005;60:633-638
139. Buist AS, McBurnie MA, Vollmer WM, et al. International variation in the prevalence of COPD (the BOLD study): A population-based prevalence study. *Lancet.* 2007;370:741-750
140. Lundbäck B, Lindberg A, Lindström M, et al. Not 15 but 50% of smokers develop COPD?--report from the obstructive lung disease in northern Sweden studies. *Respir Med.* 2003;97:115-122
141. Wurtz ET, Schlunssen V, Malling TH, et al. Occupational COPD among Danish never-smokers: A population-based study. *Occup Environ Med.* 2015;72:456-459
142. Salvi SS, Barnes PJ. Chronic obstructive pulmonary disease in non-smokers. *Lancet.* 2009;374:733-743
143. Lindberg A, Bjerg A, Rönmark E, et al. Prevalence and underdiagnosis of COPD by disease severity and the attributable fraction of smoking report from the obstructive lung disease in northern Sweden studies. *Respir Med.* 2006;100:264-272
144. Lindberg A, Jonsson AC, Rönmark E, et al. Prevalence of chronic obstructive pulmonary disease according to bts, ers, gold and ats criteria in relation to doctor's diagnosis, symptoms, age, gender, and smoking habits. *Respiration.* 2005;72:471-479
145. Sin DD, Man SF. Chronic obstructive pulmonary disease as a risk factor for cardiovascular morbidity and mortality. *Proc Am Thorac Soc.* 2005;2:8-11
146. Mapel DW, Dedrick D, Davis K. Trends and cardiovascular co-morbidities of COPD patients in the veterans administration medical system, 1991-1999. *COPD.* 2005;2:35-41
147. Curkendall SM, DeLuise C, Jones JK, et al. Cardiovascular disease in patients with chronic obstructive pulmonary disease, Saskatchewan Canada cardiovascular disease in COPD patients. *Ann Epidemiol.* 2006;16:63-70
148. Sidney S, Sorel M, Quesenberry CP, Jr., et al. COPD and incident cardiovascular disease hospitalizations and mortality: Kaiser permanente medical care program. *Chest.* 2005;128:2068-2075
149. Finkelstein J, Cha E, Scharf SM. Chronic obstructive pulmonary disease as an independent risk factor for cardiovascular morbidity. *Int J Chron Obstruct Pulm Dis.* 2009;4:337-349
150. Feary JR, Rodrigues LC, Smith CJ, et al. Prevalence of major comorbidities in subjects with copd and incidence of myocardial infarction and stroke: A comprehensive analysis using data from primary care. *Thorax.* 2010;65:956-962
151. Schneider C, Bothner U, Jick SS, et al. Chronic obstructive pulmonary disease and the risk of cardiovascular diseases. *Eur J Epidemiol.* 2010;25:253-260
152. Huiart L, Ernst P, Suissa S. Cardiovascular morbidity and mortality in COPD. *Chest.* 2005;128:2640-2646
153. Trell E. Community-based preventive medical department for individual risk factor assessment and intervention in an urban population. *Prev Med.* 1983;12:397-402
154. Berglund G, Nilsson P, Eriksson KF, et al. Long-term outcome of the Malmö Preventive Project: Mortality and cardiovascular morbidity. *J Intern Med.* 2000;247:19-29

155. Berglund G, Elmståhl S, Janzon L, et al. The Malmö Diet and Cancer study. Design and feasibility. *J Intern Med*. 1993;233:45-51
156. Manjer J, Carlsson S, Elmstahl S, et al. The Malmö Diet and Cancer study: Representativity, cancer incidence and mortality in participants and non-participants. *Eur J Cancer Prev*. 2001;10:489-499
157. Manjer J, Elmståhl S, Janzon L, et al. Invitation to a population-based cohort study: Differences between subjects recruited using various strategies. *Scand J Public Health*. 2002;30:103-112
158. Pero RW, Olsson A, Bryngelsson C, et al. Quality control program for storage of biologically banked blood specimens in the Malmö Diet and Cancer study. *Cancer Epidemiol Biomarkers Prev*. 1998;7:803-808
159. Riboli E, Elmståhl S, Saracci R, et al. The Malmö food study: Validity of two dietary assessment methods for measuring nutrient intake. *Int J Epidemiol*. 1997;26:161-173
160. Rosvall M, Janzon L, Berglund G, et al. Incidence of stroke is related to carotid IMT even in the absence of plaque. *Atherosclerosis*. 2005;179:325-331
161. Ludvigsson JF, Andersson E, Ekbom A, et al. External review and validation of the Swedish National Inpatient Register. *BMC Public Health*. 2011;11:450
162. Inghammar M, Engström G, Löfdahl CG, et al. Validation of a COPD diagnosis from the Swedish Inpatient Registry. *Scand J Public Health*. 2012;40:773-776
163. Ekberg-Aronsson M, Pehrsson K, Nilsson JA, et al. Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. *Respir Res*. 2005;6:98
164. Kristenson H, Trell E. Indicators of alcohol consumption: Comparisons between a questionnaire (MM-Mast), interviews and serum gamma-glutamyl transferase (GGT) in a health survey of middle-aged males. *Br J Addict*. 1982;77:297-304
165. Assarsson E, Lundberg M, Holmquist G, et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS One*. 2014;9:e95192
166. Olink Bioscience. Data package Proseek Multiplex CVD. <http://www.olink.com/proseek-multiplex/cvd/> Accessed: Aug 21, 2016.
167. Pessah-Rasmussen H, Engström G, Jerntorp I, et al. Increasing stroke incidence and decreasing case fatality, 1989-1998: A study from the Stroke Register in Malmö, Sweden. *Stroke*. 2003;34:913-918
168. Rothman K. *Epidemiology. An introduction*. New York: Oxford University Press; 2002.
169. Kleinbaum DGK, M. *Survival analysis. A self-learning text*. New York: Springer Science+Business Media, Inc.; 2005.
170. Rich JT, Neely JG, Paniello RC, et al. A practical guide to understanding Kaplan-Meier curves. *Otolaryngol Head Neck Surg*. 2010;143:331-336
171. Cox D. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*. 1972;34:187-220
172. Sperrin M, Buchan I. Modelling time to event with observations made at arbitrary times. *Stat Med*. 2013;32:99-109
173. Schemper M, Wakounig S, Heinze G. The estimation of average hazard ratios by weighted Cox regression. *Stat Med*. 2009;28:2473-2489
174. Andersson T, Alfredsson L, Källberg H, et al. Calculating measures of biological interaction. *Eur J Epidemiol*. 2005;20:575-579
175. Rothman K, Greenland S, Lash T. *Modern epidemiology*. Lippincott Williams & Wilkins; 2008:139-142.

176. Langholz B, Richardson D. Are nested case-control studies biased? *Epidemiology*. 2009;20:321-329
177. Rodrigues L, Kirkwood BR. Case-control designs in the study of common diseases: Updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. *Int J Epidemiol*. 1990;19:205-213
178. Pencina MJ, D'Agostino RB. Overall c as a measure of discrimination in survival analysis: Model specific population value and confidence interval estimation. *Stat Med*. 2004;23:2109-2123
179. Lee S, Emond MJ, Bamshad MJ, et al. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet*. 2012;91:224-237
181. Costanza MC. Matching. *Prev Med*. 1995;24:425-433
182. Vandembroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational studies in Epidemiology (STROBE): Explanation and elaboration. *Epidemiology*. 2007;18:805-835
183. Sjölander A, Greenland S. Ignoring the matching variables in cohort studies - when is it valid and why? *Stat Med*. 2013;0
184. Delgado-Rodriguez M, Llorca J. Bias. *J Epidemiol Community Health*. 2004;58:635-641
185. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615-625
186. Berry SD, Ngo L, Samelson EJ, et al. Competing risk of death: An important consideration in studies of older adults. *J Am Geriatr Soc*. 2010;58:783-787
187. Andersen PK, Geskus RB, de Witte T, et al. Competing risks in epidemiology: Possibilities and pitfalls. *Int J Epidemiol*. 2012;41:861-870
188. MacMahon S, Peto R, Cutler J, et al. Blood pressure, stroke, and coronary heart disease. Part 1, prolonged differences in blood pressure: Prospective observational studies corrected for the regression dilution bias. *Lancet*. 1990;335:765-774
189. American Thoracic Society. Snowbird workshop on standardization of spirometry. *Am Rev Respir Dis*. 1979;119:831-838
190. Engström G, Segelstorm N, Ekberg-Aronsson M, et al. Plasma markers of inflammation and incidence of hospitalisations for COPD: Results from a population-based cohort study. *Thorax*. 2009;64:211-215
191. Montnemery P, Adelroth E, Heuman K, et al. Prevalence of obstructive lung diseases and respiratory symptoms in southern Sweden. *Respir Med*. 1998;92:1337-1345
192. Montnemery R, Nihlen U, Andersson M, et al. Obstructive airways diseases, smoking and use of inhaled corticosteroids in southern Sweden in 1992 and 2000. *Int J Tuberc Lung Dis*. 2006;10:490-498
193. Hallström B, Jönsson AC, Nerbrand C, et al. Lund stroke register: Hospitalization pattern and yield of different screening methods for first-ever stroke. *Acta Neurol Scand*. 2007;115:49-54
194. Stegmayr B, Asplund K. Measuring stroke in the population: Quality of routine statistics in comparison with a population-based stroke registry. *Neuroepidemiology*. 1992;11:204-213
195. Merlo J, Lindblad U, Pessah-Rasmussen H, et al. Comparison of different procedures to identify probable cases of myocardial infarction and stroke in two Swedish prospective cohort studies using local and national routine registers. *Eur J Epidemiol*. 2000;16:235-243
196. Köster M, Asplund K, Johansson A, et al. Refinement of Swedish administrative registers to monitor stroke events on the national level. *Neuroepidemiology*. 2013;40:240-246

197. Appelros P, Terent A. Validation of the Swedish Inpatient and Cause-of-death registers in the context of stroke. *Acta Neurol Scand.* 2011;123:289-293
198. McNamee R. Regression modelling and other methods to control confounding. *Occup Environ Med.* 2005;62:500-506, 472
199. Greenland S. Modeling and variable selection in epidemiologic analysis. *Am J Public Health.* 1989;79:340-349
200. Engström G, Gerhardsson de Verdier M, Dahlbäck M, et al. Blood pressure variability and cardiovascular autonomic function in relation to forced expiratory volume: A population-based study. *Chest.* 2009;136:177-183
201. Lind P, Engström G, Stavenow L, et al. Risk of myocardial infarction and stroke in smokers is related to plasma levels of inflammation-sensitive proteins. *Arterioscler Thromb Vasc Biol.* 2004;24:577-582
202. Engström G, Site-Flondell D, Lindblad B, et al. Risk of treatment of peripheral arterial disease is related to inflammation-sensitive plasma proteins: A prospective cohort study. *J Vasc Surg.* 2004;40:1101-1105
203. Engström G, Borner G, Lindblad B, et al. Incidence of fatal or repaired abdominal aortic aneurysm in relation to inflammation-sensitive plasma proteins. *Arterioscler Thromb Vasc Biol.* 2004;24:337-341
204. Schnabel E, Nowak D, Brasche S, et al. Association between lung function, hypertension and blood pressure medication. *Respir Med.* 2011;105:727-733
205. Engström G, Wollmer P, Valind S, et al. Blood pressure increase between 55 and 68 years of age is inversely related to lung function: Longitudinal results from the cohort study 'Men born in 1914'. *J Hypertens.* 2001;19:1203-1208
206. Portegies ML, Lahousse L, Joos GF, et al. Chronic obstructive pulmonary disease and the risk of stroke. The Rotterdam study. *Am J Respir Crit Care Med.* 2016;193:251-258
207. Lindberg A, Larsson LG, Muellerova H, et al. Up-to-date on mortality in COPD - report from the OLIN COPD study. *BMC Pulm Med.* 2012;12:1
208. Löfdahl CG, Tilling B, Ekstrom T, et al. COPD health care in Sweden - a study in primary and secondary care. *Respir Med.* 2010;104:404-411
209. National board of health and welfare, Sweden. Tobaksvanor och tobaksrelaterade sjukdomar. Public health report 2009.; 2009:291-310.
210. Lindberg A, Larsson LG, Rönmark E, et al. Co-morbidity in mild-to-moderate COPD: Comparison to normal and restrictive lung function. *COPD.* 2011;8:421-428
211. Ford ES, Wheaton AG, Mannino DM, et al. Elevated cardiovascular risk among adults with obstructive and restrictive airway functioning in the united states: A cross-sectional study of the National Health and Nutrition Examination Survey from 2007-2010. *Respir Res.* 2012;13:115
212. Donaldson GC, Hurst JR, Smith CJ, et al. Increased risk of myocardial infarction and stroke following exacerbation of COPD. *Chest.* 2010;137:1091-1097
213. Windsor C, Herrett E, Smeeth L, et al. No association between exacerbation frequency and stroke in patients with COPD. *Int J Chron Obstruct Pulm Dis.* 2016;11:217-225
214. McColgan P, Thant KZ, Sharma P. The genetics of sporadic ruptured and unruptured intracranial aneurysms: A genetic meta-analysis of 8 genes and 13 polymorphisms in approximately 20,000 individuals. *J Neurosurg.* 2010;112:714-721
215. Baizabal Carvallo JF, Cantu Brito C, Estanol B, et al. Subarachnoid hemorrhage as a complication of systemic lupus erythematosus. *Cerebrovasc Dis.* 2007;24:301-304

