



# LUND UNIVERSITY

## Population-based prediction of atrial fibrillation

Smith, Gustav

2012

[Link to publication](#)

*Citation for published version (APA):*

Smith, G. (2012). *Population-based prediction of atrial fibrillation*. [Doctoral Thesis (compilation), Cardiology]. Lund University.

*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

# Population-based prediction of atrial fibrillation

J. Gustav Smith, MD

Doctoral Thesis

which, with due permission from the Faculty of Medicine, Lund University, will be publicly defended 09.00, May 4<sup>th</sup>, 2012  
Segerfalksalen, Biomedicinskt Centrum hus A



**LUND**  
UNIVERSITY

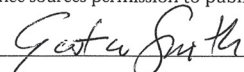
**Faculty opponent**

Professor Stefan Käab, MD, PhD  
Ludwig Maximilian University of Munich, Germany

Organization LUND UNIVERSITY		Document name DOCTORAL DISSERTATION	
Department of Cardiology Clinical Sciences, Lund Faculty of Medicine, Lund University Lund, Sweden		Date of issue May 4, 2012	
Author(s) J. Gustav Smith		Sponsoring organization	
Title and subtitle Population-based prediction of atrial fibrillation			
Abstract Background: Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia and increases in prevalence. As AF confers increased risk of morbidity, stroke, dementia and mortality, a major objective of cardiovascular epidemiology today is the development of tools for prediction and prevention of AF and its consequences. Although AF has traditionally been considered a non-heritable disease, a familial component to AF has recently been established and genetic polymorphisms associated with AF risk. The aim of this thesis was therefore to evaluate the ability of genetic polymorphisms, conventional cardiovascular risk factors and blood biomarkers reflecting diverse pathophysiological pathways to predict onset of AF in the general population. Methods: Individuals with incident AF were identified from a large, population-based cohort study (Malmö Diet and Cancer, n=30 447) using national registers during up to 17.8 years follow-up. Genetic polymorphisms on chromosomes 4q25, 16q22 and in the KCNH2 gene reproducibly associated with AF were genotyped in the entire cohort. Six blood biomarkers (MR-proANP, MR-proADM, NT-proBNP, copeptin, CRP and cystatin c) were measured in a random subcohort (n=6104). Results: Case validity of AF diagnoses in national registers was estimated to be high, 95-97%. A substantial proportion of population risk was conferred by conventional risk factors (hypertension, 34-38%; obesity, 10-11%). Conventional risk factors predicted AF with reasonable accuracy (C-statistic 0.732), although age and sex considered alone were only modestly less accurate. Addition of blood MR-proANP and CRP improved predictive accuracy modestly (C-statistic 0.753). Polymorphisms on chromosomes 4q25 and 16q22 but not in KCNH2 were associated with AF independently of conventional risk factors, but did not significantly improve predictive accuracy. In a meta-analysis of published studies including up to 150 000 individuals, the association of polymorphisms with AF was robust across diverse study designs, but with widely varying risk estimates. Genetic polymorphisms but not conventional risk factors were associated with AF in heart failure (HF) patients. A SNP on chromosome 16q22 was more strongly associated with AF in this context (p for interaction = 7x10 <sup>-4</sup> ), suggestive of a pathophysiological interaction of the gene affected by this SNP and HF. Conclusion: The results of this thesis have implications for population-based prediction of AF and provides novel information on risk factors for AF in HF patients.			
Key words: atrial fibrillation, prediction, epidemiology, genetics, natriuretic peptide, heart failure			
Classification system and/or index terms (if any):			
Supplementary bibliographical information:		Language English	
ISSN and key title: 1652-8220		ISBN 978-91-86871-96-3	
Recipient's notes		Number of pages	Price
		Security classification	

Distribution by (name and address)

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

Signature 

Date March 20, 2012

# Population-based prediction of atrial fibrillation

J. Gustav Smith, MD



**LUND**  
UNIVERSITY

Doctoral Thesis  
2012

Department of Cardiology  
Faculty of Medicine  
Lund University, Sweden

“Certis rebus certa signa praecurrunt”  
Cicero (106 BC - 43BC)

“Prediction is very difficult, especially about the future”  
Niels Bohr (1885-1962)

“I never think of the future – it comes soon enough”  
Albert Einstein (1879-1955)

Copyright © 2012 J. Gustav Smith

Lund University, Faculty of Medicine Doctoral Dissertation Series 2012:34  
ISBN 978-91-86871-96-3  
ISSN 1652-8220

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

# Contents

List of papers	7
Abbreviations	8
Chapter 1: Introduction	11
1.1 Epidemiological transitions in the spectrum of cardiovascular disease	12
1.2 Atrial fibrillation	13
1.3 Prediction and prevention of cardiovascular disease	15
1.4 Genetic causes of cardiovascular disease	16
1.4.1 The genetic architecture of complex disease	16
1.4.2 Atrial fibrillation	19
Chapter 2: Aims of the thesis	23
Chapter 3: Methods	25
3.1 Prospective cohort study	26
3.1.1 Population and cohort descriptions	26
3.1.2 Procedures and definitions	27
3.1.3 Laboratory methods	27
3.2 Follow-up with national registers	28
3.3 Prediction of atrial fibrillation	30
3.3.1 Transformations	30
3.3.2 Association tests	30
3.3.3 Attributable risk	31
3.3.4 Discrimination	31
3.3.5 Risk reclassification	31
3.4 Literature-based meta-analysis	32
3.4.1 Systematic review	32
3.4.2 Meta-analysis	32
3.5 Risk factors for AF in HF	33

Chapter 4: Results	35
4.1 Diagnostic validity of register diagnosis	36
4.2 Conventional risk factors	37
4.3 Blood biomarkers	39
4.3.1 Blood biomarkers associated with AF	39
4.3.2 Predictive accuracy	41
4.4 Genetic polymorphisms	43
4.4.1 Association of polymorphisms with AF	43
4.4.2 Discrimination	44
4.5 Literature-based meta-analysis	46
4.5.1 Systematic review	46
4.5.2 Meta-analyses	48
4.6 Risk factors for AF in HF	51
4.6.1 Conventional risk factors	52
4.6.2 Genetic polymorphisms and AF in HF patients	53
4.6.3 Interaction analyses in the entire MDCS cohort	56
Chapter 5: Discussion	57
5.1 Diagnostic validity and generalizability of findings	57
5.2 Conventional risk factors	59
5.3 Blood biomarkers	60
5.4 Genetic polymorphisms	61
5.5 Risk factors for AF in HF	62
Chapter 6: Perspectives	65
6.1 Prediction and prevention of AF	65
6.2 The genetic architecture of AF and implications for prediction	66
6.3 Heterogeneity of AF	68
Chapter 7: Conclusions	69
Summary in Swedish (Sammanfattning på svenska)	71
Acknowledgements	73
References	75
Appendix	

# List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals. Reprints of the papers are appended at the end of the thesis with permission from publishers.

- I. Smith JG, Platonov PG, Hedblad B, Engström G, Melander O. Atrial fibrillation in the Malmö Diet and Cancer study: a study of occurrence, risk factors and diagnostic validity. *Eur J Epidemiol* 2010; 25 (2): 95-102.
- II. Smith JG, Newton-Cheh C, Almgren P, Struck J, Morgenthaler NG, Bergmann A, Platonov PG, Hedblad B, Engström G, Wang TJ, Melander O. Assessment of conventional cardiovascular risk factors and multiple biomarkers for the prediction of incident heart failure and atrial fibrillation. *J Am Coll Cardiol* 2010; 56 (21): 1712-9.
- III. Smith JG, Newton-Cheh C, Almgren P, Melander O, Platonov PG. Genetic polymorphisms for estimating risk of atrial fibrillation in the general population: a prospective study. *Arch Int Med* 2012: in press.
- IV. Smith JG, Almgren P, Platonov PG, Newton-Cheh C, Melander O. Genetic polymorphisms for estimating risk of atrial fibrillation: literature-based meta-analysis. *J Intern Med* 2012: accepted with minor revisions.
- V. Smith JG, Melander O, Sjögren M, Hedblad B, Engström G, Newton-Cheh C, Platonov PG. Genetic polymorphisms confer risk of atrial fibrillation in patients with heart failure. (submitted)

(I) Copyright © 2010, European Society of Epidemiology. Published by Springer. (II) Copyright © 2010, American College of Cardiology Foundation. Published by Elsevier Inc. (III) Copyright © 2012, American Medical Association. All rights reserved. (IV) Copyright © 2012, The Association for the Publication of the Journal of Internal Medicine. Published by Blackwell Publishing.



# Abbreviations

95% CI	95 % confidence interval
AF	atrial fibrillation
ANP	atrial natriuretic peptide
BMI	body mass index
BNP	b-type natriuretic peptide
CAD	coronary artery disease
CDR	the swedish cause of death register
CRP	c-reactive protein
CVD	cardiovascular disease
DBP	diastolic blood pressure
ECG	electrocardiogram
FHS	framingham heart study
GWAS	genome-wide association study
HDR	the swedish hospital discharge register
HF	heart failure
HR	hazard ratio
ICD	the international classification of diseases
IDI	integrated discrimination improvement
MDCS	malmö diet and cancer study
MDC-CC	the cardiovascular cohort of mdcs
MDC-DNA	mdcs participants with dna
MDC-HF	the hf subcohort of mdcs
MR-proANP	midregional pro-atrial natriuretic peptide

NRI	net reclassification improvement
NT-proBNP	n-terminal pro-b-type natriuretic peptide
OR	odds ratio
PAR	population attributable risk
SBP	systolic blood pressure
SD	standard deviation



# Chapter 1: Introduction

The human heart is a muscular organ which contracts rhythmically to generate pressure to drive blood through the circulatory system. Contractions are precisely coordinated by electrical impulses which are initiated and propagated through the heart in a highly specific way, along its route triggering sequential initiation of contraction. Once blood has been ejected from the heart it is prevented from reversing its flow back into the cardiac chambers by a valve apparatus. The energy-demanding work in cardiac muscle cells (collectively referred to as the myocardium) is fueled by oxygen and nutrients supplied by the cardiac vasculature, termed the coronary arteries because they radiate around the top of the heart resembling a crown (*lat.* corona) and send smaller branches downwards across the heart. The heart is also an endocrine organ, as it releases at least two types of peptide hormones into the bloodstream in response to increased wall tension: Atrial natriuretic peptide (ANP) and B-type natriuretic peptide (BNP), which both induce renal sodium excretion.

Diseases affecting the heart and blood vessels (cardiovascular diseases) are common and the leading cause of death worldwide. The major heart diseases in the population are diseases that blood limit flow in the cardiac vasculature (coronary artery disease, CAD), diseases that result in reduced function of cardiac muscle (myocardial disease or heart failure, HF), diseases that impair function of the cardiac valves, and diseases that affect the cardiac structures involved in generation and propagation of electrical impulses (arrhythmias), but many diseases impact a combination of these structures.

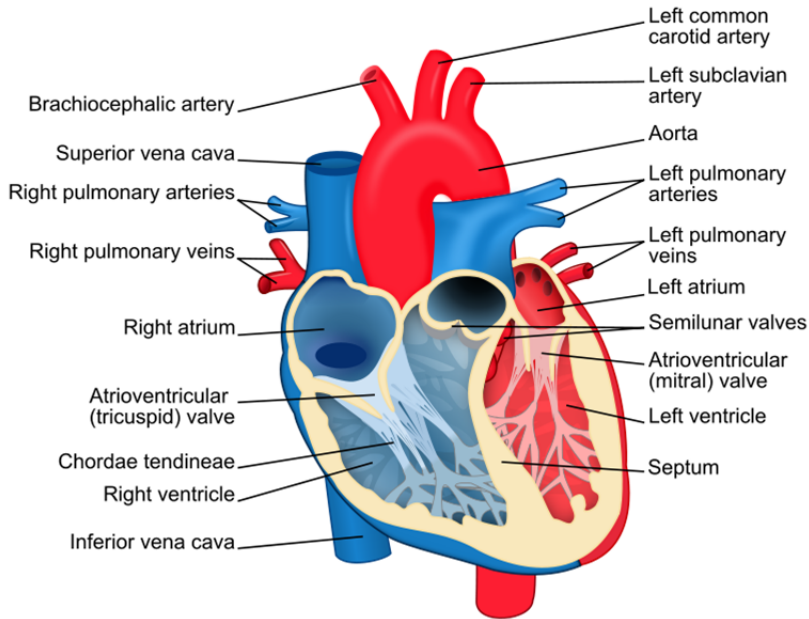


Figure 1.1: Schematic diagram of the human heart. Anterior view.

## 1.1 Epidemiological transitions in the spectrum of cardiovascular disease

Epidemiology is the study of the distribution and determinants of disease in populations and the application of this study to control health problems. Following the success of John Snow, often considered the father of modern epidemiology, who used epidemiological methods to track the origin of cholera to water pumps in the London epidemic of 1854 (1) epidemiological methods were utilized to monitor infectious diseases. During the second half of the 20<sup>th</sup> century, epidemiological efforts have instead increasingly focused on cardiovascular disease and cancer, the leading causes of death following the large decline in deaths attributable to famine and infectious disease in the first half of the 20<sup>th</sup> century (2). Early progress in cancer and cardiovascular epidemiology was made with the British Doctors Study which strongly linked lung cancer and myocardial infarction to tobacco smoking (3) and the Framingham Heart Study (4,5) and Seven Countries Study (6) which helped confirm cholesterol, hypertension and obesity as key risk factors for coronary artery disease.

The last three decades have seen a shift in the spectrum of cardiovascular disease in high-income countries. Specifically, the incidence and mortality in myocardial infarction, the acute manifestation of coronary artery disease, has declined in parallel with improvements in health technology and adoption of healthier lifestyles (7-11) while the incidence and prevalence of atrial fibrillation and heart failure have increased. This transition from famine and infectious disease to metabolic disease presenting in middle age and now recently to degenerative disease presenting in old age has been termed ‘the epidemiological transition of cardiovascular disease’ (7) and is outlined in **Table 1.1**.

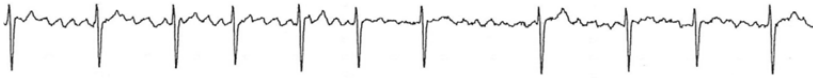
Stage	Description	Predominant type of CVD
Pestilence and famine	Low mean life expectancy, high rate of infant and child mortality	Rheumatic heart disease, cardiomyopathy caused by infection and malnutrition
Receding pandemics	Improvements in nutrition and public health in the 19th-20th centuries	Rheumatic valvular disease, CAD, stroke
Degenerative disease	Increased caloric intake and decreased physical activity led to emergence of hypertension and atherosclerosis in middle age. Noncommunicable disease exceeds mortality from malnutrition and infectious disease	CAD, stroke
Delayed degenerative disease	Improved treatment and prevention of cardiovascular disease in the late 20th century. Decline in age-adjusted CVD mortality	CAD, stroke, heart failure, atrial fibrillation, degenerative valvular disease

**Table 1.1:** Stages of the epidemiological transition and predominant types of cardiovascular disease (CVD). Adapted from (7).

## 1.2 Atrial fibrillation

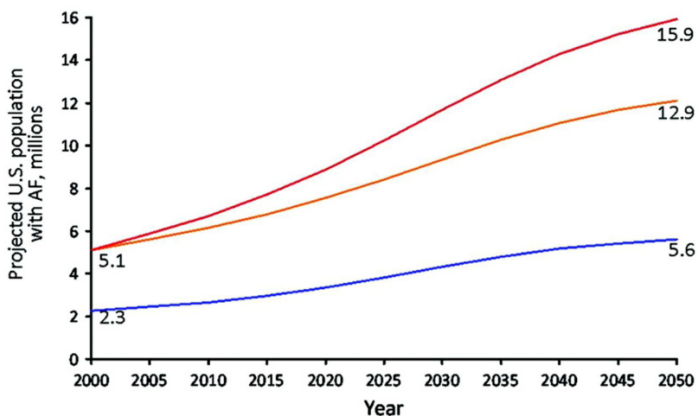
Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia. AF is characterized by an irregular heart rhythm, loss of P-waves on the electrocardiogram (ECG) and loss of atrial contraction resulting in an increased propensity to atrial clotting and thromboembolism. AF is diagnosed from the ECG (12), as it has been since the electrocardiographic representation of the ‘pulsus irregularis perpetuus’ was first described (13) only shortly after the invention of the ECG by Willem Einthoven in the

early 20<sup>th</sup> century (14). Symptoms of AF include palpitations, fatigue, dyspnea on exertion and chest pain, but AF can also be asymptomatic and discovered incidentally from an ECG or when a patient develops thromboembolic disease.



**Figure 1.2:** Electrocardiographic representation of AF.

Previous studies have estimated the prevalence of recognized AF to 1-2% in the general population of Europe and North America (15-18), and that number is expected to increase at least twofold in Europe and Northern America over the next 50 years as the population ages (19,20) as shown in **Figure 1.3**. The lifetime risk of middle-aged individuals to develop AF has been estimated to 25% (21) but might be even higher, as unrecognized, clinically silent AF episodes might not be uncommon (22).



**Figure 1.3:** Projections for AF prevalence in the United States. The bottom line shows projections from the ATRIA study (15). The line in the middle shows projections based on observed AF rates between 1980 and 2000 in Olmsted County, Minnesota assuming no further increase in age-adjusted AF incidence (19). The upper line shows projections based on rates in Olmsted County, assuming a continued increase in incidence rate. Reprinted from a published review article (23) with permission from the publisher.

AF confers an approximately 4- to 5-fold increased risk of stroke (24,25), twofold increased risk of asymptomatic cerebral infarction (26-28), twofold increased risk of cognitive dysfunction (29,30), 4- to 6-fold increased risk of thromboembolism in internal organs (31) and twofold increase in mortality (16,25,32-34). At least 20% of stroke patients have AF and, compared with stroke of other etiologies, stroke in the context of AF generally has greater severity, tends to be more disabling, and is associated with increased mortality (35). Silent AF has also been associated with increased stroke risk (22). Substantial reductions in risk of thromboembolism and stroke, in the order of 50-70% relative risk reduction, can be achieved with oral anticoagulation (36-38), which could also improve survival (39,40). Such therapy is also likely to prevent silent infarctions and cognitive dysfunction, although interventional studies are lacking (41).

AF is closely associated with HF (25). Onset of AF in HF patients is associated with increased morbidity and mortality, potentially mediated by a rapid ventricular rate, loss of atrial contraction, irregular ventricular filling time, or thromboembolism (42-44). However, pathophysiological mechanisms linking the two diseases remain incompletely understood (43,45,46). The prevalence of AF rises steeply with the severity of HF, such that patients with severe HF (New York Heart Association class IV) have an AF prevalence of up to 50% (42,43,47,48). In patients with diastolic dysfunction, AF risk is proportional to ventricular filling pressures, as measured echocardiographically (49). These observations and animal models of HF support the assumption that atrial stretch resulting from increased filling pressures could alter atrial electrophysiological properties, facilitating AF (46). Other studies have suggested that neurohormonal activation or atrial ionic channel remodeling occurring in patients with HF may play a role in facilitating AF risk (46). The association of AF with HF could also result from common risk factors acting independently on both ventricles and atria (43). Finally, AF with a rapid ventricular rate can result in ventricular dysfunction, which could also explain part of the association (46).

### 1.3 Prediction and prevention of cardiovascular disease

Much research during the second half of the 20<sup>th</sup> century has focused on determining the causes of CAD and on the development of preventive medications. For CAD, lifestyle interventions and preventive medications including cholesterol-lowering agents, antihypertensives and platelet inhibitors have been shown to reduce risk of first or recurrent acute coronary events (50-52). However, whereas lifestyle interventions targeting smoking cessation and possibly weight reduction have great benefits for populations, the benefits of treating entire populations with preventive medications are small, why costs and risks might outnumber potential gains (53). Current guidelines for



the use of medications in primary prevention of CAD therefore focus on the identification of individuals at high risk. To identify such individuals, guidelines recommend the use of statistical prediction models incorporating information on multiple risk factors for estimation of individual absolute risk (53). These risk prediction models are usually based on relative risk estimates for individual risk factors derived from multivariable regression models in population-based cohort studies, which are applied to an estimate of the average population risk to produce an estimate of absolute risk. Such models were pioneered in the Framingham Heart Study for CAD, and changes to these models have gradually been suggested by the FHS investigators, incorporating the key risk factors (53-56). Prediction models for other cardiovascular diseases have also been developed by FHS investigators, including stroke (57), AF (58,59) and HF (60). However, studies have shown that FHS models overestimate risk in populations with lower disease rates than Framingham and underestimate risk in populations with higher disease rate, which has spurred development of other models, including a European prediction model for CAD – the SCORE model – which is widely used throughout Europe today (61).

For the epidemics of HF and AF, efforts for prediction and prevention are still in an early phase with promising results for inhibitors of the renin-angiotensin-aldosterone system for prevention of HF (62) and perhaps also for AF (63), and interesting results have been reported with statins for prevention of AF (64,65). Importantly, as noted above it is now well established that anticoagulants reduce risk of thromboembolic events with 50-70% in patients with established AF, and could also impact mortality and risk of dementia. It is possible that individuals at very high risk of AF could benefit from anticoagulation for prevention of thromboembolism or other interventions to reduce AF risk. To identify such individuals, investigators in the FHS and the Atherosclerosis Risk in Communities (ARIC) study have proposed prediction models (58,66), but the consistency of risk estimates has only been the subject of limited research (59).

## 1.4 Genetic causes of cardiovascular disease

### 1.4.1 The genetic architecture of complex disease

Variations in the genetic sequence between any two individuals constitutes approximately 0.1% of the genome and can take different forms, including nucleotide substitutions, deletions, insertions, duplications or inversions. Regional and global (67,68) resequencing studies have found smaller sequence variants to be the most common type, with single-base substitutions predominating. Large-scale discovery projects of single-base substitutions with high population frequencies, called single nucleotide polymorphisms

(SNPs), have identified more than 10 million SNPs in human populations, available in online catalogs such as dbSNP ([www.ncbi.nlm.nih.gov/projects/SNP](http://www.ncbi.nlm.nih.gov/projects/SNP)) (69).

It has long been known that a family history of disease increases an individual's risk of contracting that very same disease. Indeed, most common diseases are thought to have at least some level of genetic determination. In recent years, progress in the field of genetics has made possible the identification of causal genetic variants in systematic studies of whole human genomes. Rare diseases segregating in families with obvious inheritance patterns, of which the congenital Long QT syndrome, hypertrophic cardiomyopathy and familial hypercholesterolemia are perhaps the best known examples affecting the heart, have been found to result from rare mutations of strong effect. Variants with such strong effect on disease susceptibility remain rare because of negative selection (70). Some variants of strong effect may be more common in certain populations owing to founder effects or positive selection in certain environments, as exemplified by sickle cell anemia, caused by variants in the beta-globin gene present at high frequency in regions where malaria is endemic as it confers resistance to Malaria (71). More than 2000 diseases have been found to result from rare mutations of strong effect, with population frequencies typically below 1%. However, the contribution of rare variants to complex human diseases and traits is unknown. These traits have been postulated to result from interactions between multiple genetic variants, both common and rare, and environmental factors, and because of late onset have typically escaped negative selection (70).

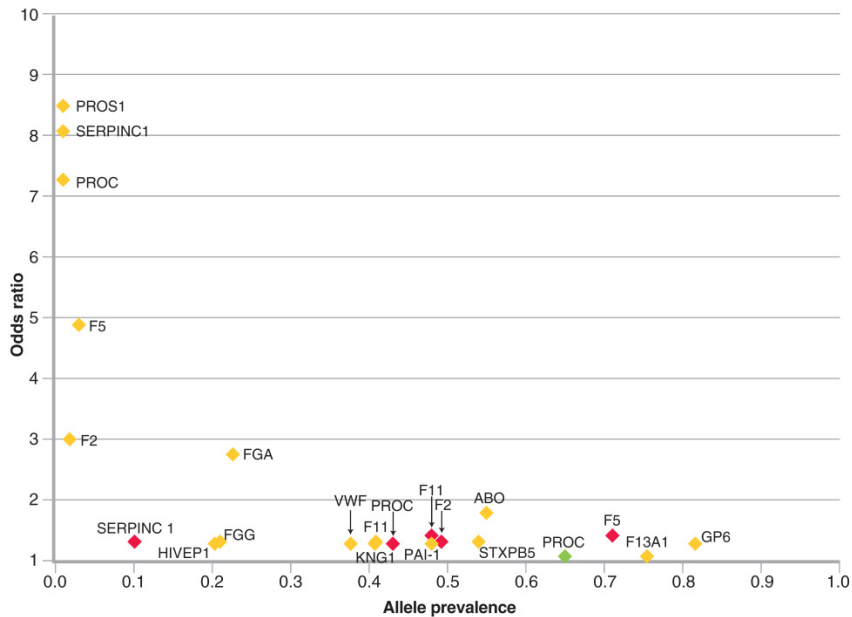
In steadily increasing numbers, some genes have been found to harbor both common and rare variants that are associated with human diseases and traits. For example, rare variants in the low-density lipoprotein (LDL) receptor gene that cause familial hypercholesterolemia with very high plasma concentrations of LDL, leading to accelerated atherosclerosis and myocardial infarction if untreated (72). Common variants in the same gene have also been found to incrementally increase LDL cholesterol and risk of myocardial infarction in the general population (73). The effects of common variants are typically smaller, with relative risk estimates  $<2$ , but can translate into much higher population-attributable risks than mutations underlying monogenic diseases, owing to the comparatively high population frequency of polymorphisms. The causal genetic variant for monogenic diseases can be identified using either whole-genome linkage analysis, which test for the joint transmission of chromosomal segments and diseases in families, or more recently with whole-exome resequencing, in which all known coding regions throughout the genome are examined for mutations with the potential to cause deleterious effects on the gene product.

For common diseases, such as AF, HF and CAD, the method of choice to identify genetic variants that are common in the general population is association analysis, which simply tests for differences in allele frequencies between cases and controls. Historically,

association studies have been able to examine polymorphisms in only those candidate genes known or proposed to play a role in the pathophysiology of disease. The study of the genetic basis of complex traits was set back by the failure of candidate gene association findings reported in the 1990s and early 2000s to be replicated, which arose from use of inappropriately permissive *P*-value thresholds and small sample sizes in replication samples (74). This resulted in wide-spread adoption of recommendations for strict significance thresholds and replication in independent samples in the field (75).

More recently, it has become possible to examine large numbers of polymorphisms, in the order of 100 000 – 1 000 000, throughout the genome using highly parallel genotyping arrays. These genome-wide association (GWA) studies systematically examine common variation throughout the human genome regardless of putative biologic function. Owing to the fact that common polymorphisms are correlated, a smaller number of polymorphisms can be chosen to serve as proxies for the majority of common sequence variations. The GWA study is hypothesis generating in the sense that it can identify polymorphisms near genes without a recognized pathophysiological link to disease. Given the large number of tests, strict significance thresholds and independent replication is usually required (76).

**Figure 1.3** illustrates genotypic effects on risk of venous thromboembolism, for which genotyping is often performed in a clinical context, compared to population allele frequency. Note how rare mutations in the genes encoding Protein S (*PROS1*), Protein C (*PROC*) and Antithrombin (*SERPINC1*) private to certain families (risk allele frequency in the population <1%) confer very high risk of thromboembolism. In contrast, the more common polymorphisms in Factor V (the so-called Leiden polymorphism) and Factor II (prothrombin) with population frequencies of 0.03 and 0.02 in many European populations, confer intermediate risks which are additive to other risk factors such as surgery, immobilization, cancer and pregnancy. Shown are also a large number of common variants (frequency 10-90%) in other genes, identified using association analyses, which confer very modest risks.



**Figure 1.4:** Genetic variants associated with risk of venous thromboembolism. Inverse relation of allele prevalence to odds ratio. Reprinted with permission from the publisher from Blondon et al, *Curr Cardiovasc Risk Rep* 2011; 5: 525-32.

### 1.4.2 Atrial fibrillation

Familial forms of AF were described 70 years ago (77,78) but only in the 21<sup>st</sup> century was it shown that a significant proportion of population risk was attributable to a family history, particularly in early-onset cases (79-84). More recently, actual genetic variants underlying such familial clustering have been identified using linkage analyses, resequencing and association studies.

Rare genetic mutations have been described in genes encoding several cardiac potassium channels, subunits of the major cardiac sodium channel, Atrial Natriuretic Peptide and Nucleoporin, as shown in **Table 1.2**. However, these mutations have thus far been private to the family in which they were described, and no mutations have been observed in exonic resequencing of consecutive clinical series of patients with familial AF for several of these genes (85,86). Therefore, current guidelines from international arrhythmologist societies do not currently recommend genotyping for atrial fibrillation (87).

Gene	Gene product	Functional effect of mutation	Family/proband characteristics	Reference
ABCC9	K <sub>ATP</sub> subunit	Defect channel gating in adrenergic state	Caucasian proband	(88)
KCNA5	K <sub>v</sub> 1.5 channel	Loss-of-function: reduced $I_{Kur}$	Caucasian proband	(89)
KCNQ1	$\alpha$ -subunit of I <sub>Ks</sub> channel	Gain-of-function: increased $I_{Ks}$	Three separate mutations in three families	(90-92)
KCNE2	$\beta$ -subunit of I <sub>Ks</sub> channel	Gain-of-function: increased $I_{Ks}$	Two Chinese AF kindreds	(93)
KCNE5	$\beta$ -subunit of I <sub>Ks</sub> channel	Gain-of-function: increased $I_{Ks}$	Isolated caucasian case	(94)
KCNJ2	Kir 2.1 channel	Gain-of-function: increased $I_{K1}$	Chinese AF kindred	(95)
NPPA	Atrial natriuretic peptide	Elevated levels of mutant ANP	Caucasian family	(96)
NUP155	Nucleoporin	Reduced nuclear membrane permeability	Family from Uruguay	(97)
SCN5A	$\alpha$ -subunit of sodium channel	Different	Separate mutations in different families	(98-104)
SCN1B	$\beta$ -subunit of sodium channel	Loss of function: reduced current	Two isolated cases	(105)
SCN2B	$\beta$ -subunit of sodium channel	Loss of function: reduced current	Two isolated cases	(105)

**Table 1.2:** Rare genetic variants associated with AF. Modified from a published review article (106).

Common genetic variants at six loci have been reproducibly associated with AF. Current guidelines from arrhythmologist societies do not recommend genotyping of such variants (87), but direct-to-consumer genotyping is available from commercial companies (107). Of a large number of candidate gene studies (108), only three studies have reported independent replication of results, with association to SNPs in *KCNH2*, *IL6R* and *GJA5* (109-111). GWA studies have also identified SNPs at three genetic loci reproducibly associated with AF; one locus on chromosome 4q25 upstream of the gene *PITX2* with a relatively large effect in the initial GWA report (112), a more common SNP on chromosome 16q22 intronic to the gene *ZFHX3* with a smaller effect (113,114), and a SNP intronic to the gene *KCNN3* on 1p21 that was associated with lone AF (115). The SNPs on chromosomes 4q25 and 16q22 have also been associated with ischemic stroke (113,116). The six loci are summarized in **Table 1.3**. A GWA study of electrocardiographic PR interval also identified five SNPs at five loci which subsequently

showed weak association (odds ratios 1.07-1.13) with AF (in the genes *SCN5A*, *SCN10A*, *NKX2-5*, *CAVI/CAV2*, *SOX5*), although independent replication has not been reported (117).

Locus	SNP	Description	MAF	Reference
1p21	rs13376333	Intronic to <i>KCNN3</i> , encoding a cardiac potassium channel. Associated with lone AF in GWAS.	34%	(115)
4q25	rs2200733	Upstream of <i>PITX2</i> , encoding a cardiac transcription factor. GWAS.	12%	(112)
16q22	rs2106261	Intronic to <i>ZFHX3</i> , encoding a cardiac transcription factor. GWAS.	18%	(113,114)
GJA5	rs10465885	Promoter SNP in the gene encoding a gap junction protein. Candidate gene.	49%	(111)
IL6R	rs4845625	Proxy for missense SNP in the gene encoding the Interleukin-6 receptor. Candidate gene.	46%	(110)
KCNH2	rs1805123	Missense SNP in the gene encoding a cardiac potassium channel. Candidate gene.	24%	(109)

**Table 1.3:** Genetic polymorphisms associated with AF. MAF, Minor Allele Frequency. Allele frequencies refer to populations of European ancestry (CEU) in HapMap phase III.



# Chapter 2: Aims of the thesis

The objectives of the present thesis were to evaluate:

- the feasibility of identifying AF cases from national health registers
- the utility of conventional cardiovascular risk factors, blood biomarkers reflecting diverse pathophysiological pathways, and genetic polymorphisms reproducibly associated with AF, for prediction of AF in the general population.
- the consistency of risk estimates for genetic polymorphisms associated with AF.
- the association of conventional cardiovascular risk factors and genetic polymorphisms with AF in individuals with HF.





# Chapter 3: Methods

The impact of conventional risk factors, blood biomarkers and polymorphisms on AF risk (Study I-III) was evaluated in a large, prospective cohort study from Malmö, Sweden – the Malmö Diet and Cancer Study (MDCS). The impact of conventional risk factors was studied in the entire cohort (Study I), and in individuals in the cardiovascular cohort (MDC-CC). The impact of blood biomarkers was studied in the MDC-CC (Study II) and polymorphisms were studied in individuals who contributed DNA (MDC-DNA, Study III). Analyses were performed using cross-sectional or prospective study designs. Cohort participants diagnosed with AF at any Swedish hospital before the baseline examination or during follow-up were identified from nation-wide health registers. Date of emigration or death were also ascertained from national registers, with censoring at the time of such events in prospective analyses. In Study IV, results from MDCS were combined with estimates from the previous literature in a meta-analysis. In Study V, all MDCS participants hospitalized for a primary diagnosis of HF were identified (MDC-HF) from national registers and risk factors for AF were evaluated.

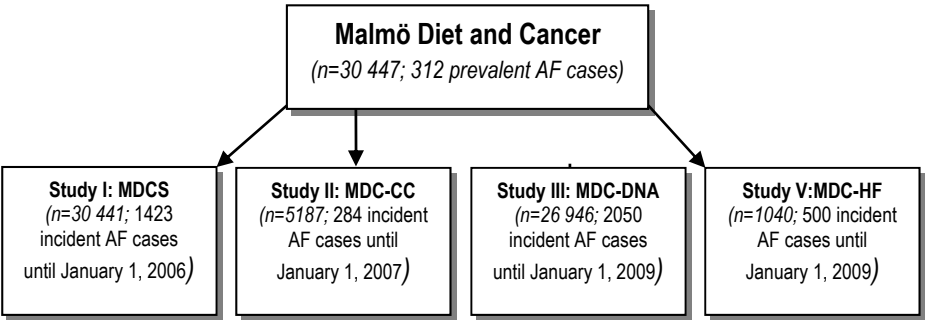


Figure 3.1: Description of study samples in the present thesis.

## 3.1 Prospective cohort study

### 3.1.1 Population and cohort descriptions

The MDCS was designed in the 1980s as an investigation into the association of diet and certain forms of cancer (particularly cancer of the breast, colon, rectum, pancreas, ovary, endometrium and prostate) based on discussions between the International Agency for Research on Cancer (IARC), the Swedish Cancer Society and the Faculty of Medicine at Lund University (118).

At the time the study was instigated, the Malmö municipality had a total population of about 230 000 inhabitants, of whom about 20% were born or had both parents born outside of Sweden, most in Denmark across the strait of Öresund. In the 21<sup>st</sup> century, the population has expanded with increasing migration from the Balkan peninsula and the Middle East to 300 000 in 2010, 40 % born or with both parents born outside of Sweden, according to official statistics from Malmö city (<http://www.malmo.se/statistik>) and Statistics Sweden (119). Thus, although the study is not representative for the present population of Malmö, it represents better the endogeneous Swedish population of Malmö and is thus less likely to include significant population stratification, a potential source of errors in genetic association studies (69).

Malmö is served by one hospital (Malmö University Hospital), which was merged with the tertiary care hospital in the neighboring city of Lund (Lund University Hospital) in 2010 to form Skåne University Hospital.

All men born between 1923 and 1945 and women born between 1923 and 1950 from Malmö were invited to participate in the MDCS. Young women were oversampled to improve the sample size for studies of breast cancer in premenopausal women. In total, 74 138 individuals were invited by letter, advertisements in local newspapers and in public places. A total of 30 447 individuals participated and attended a baseline examination between January 1, 1991 and September 25, 1996 (41% attendance rate) (120). During baseline exams between 1991 and 1994, a randomly selected subset of 6103 individuals also participated in a study of the epidemiology of carotid artery disease, the MDCS Cardiovascular Cohort (MDC-CC), and underwent additional blood tests and an ultrasonographic examination of the carotid arteries (121).

Informed consent was obtained from all participants, and the study was approved by the Ethics committee of Lund University, Lund, Sweden. The study protocol is consistent with the principles of the Declaration of Helsinki.

### 3.1.2 Procedures and definitions

At the baseline examination, participants underwent anthropometric measurements, measurement of blood pressure, filled out a questionnaire including a dietary assessment, and blood samples were collected. The questionnaire was handed out at the baseline exam and was collected and checked for missing values at a second visit, typically two weeks later. A total of 1998 individuals failed to complete either the questionnaire, anthropometric measurements or the dietary assessment, leaving 28 449 individuals with complete data, which were used in several previous MDCS analyses. For the present studies, we did not include information on dietary assessments and most questionnaire items, so opted to include all participants with data on the variables of interest.

The questionnaire included information on education, occupation, physical activity, social network, use of alcohol and tobacco, current health, medical history, current medications and disease in close relatives. Women were also asked about reproductive history. Dietary assessment was performed according to a validated dietary history method, in which frequency and intake of foods for 7 consecutive days was documented and was discussed in detail and completed at the second visit.

At the baseline exam, blood pressure was measured using a mercury-column sphygmomanometer after 10 minutes of rest in the supine position. Baseline hypertension was defined as blood pressure  $\geq 140/90$  mm Hg or use of antihypertensive medications. Participants also underwent measurement of height, weight and body composition by an impedance method (118).

Blood samples of 45 ml blood were drawn from a peripheral vein under standardized fasting conditions. Samples were immediately separated and stored in a biological bank at  $-80^{\circ}$  C or  $-140^{\circ}$  C, according to blood component fraction. In the MDC-CC, blood samples were collected from 5543 participants who underwent several blood tests at the baseline exam, including blood glucose, HbA1c, insulin, and cholesterol (total cholesterol, HDL, triglycerides) according to standard procedures at the Department of Clinical Chemistry, University Hospital Malmö (121). LDL cholesterol was calculated using the Friedewald formula. Baseline diabetes mellitus was defined as self-reported history of a physician's diagnosis of diabetes or use of antidiabetic medications. In analyses of MDC-CC, participants with fasting blood glucose  $>6.0$  mmol/l were also defined as having diabetes.

### 3.1.3 Laboratory methods

Stored plasma samples from the MDC-CC recently underwent measurement of a number of peptides (121,122), including biomarkers reflecting *hemodynamic stress* (the midregional fragment of the pro-atrial natriuretic peptide [MR-proANP], the amino-

terminal fragment of the pro-B-type natriuretic peptide [NT-proBNP], the midregional fragment of pro-adrenomedullin [MR-proADM]), *plasma volume and osmolarity* (copeptin), *inflammation* (C-reactive protein [CRP]) and *renal function* (cystatin C). Blood biomarkers were measured from plasma samples that had been frozen at  $-80^{\circ}\text{C}$  immediately after sampling and that had not previously been thawed. Copeptin and the midregional fragments of pro-atrial natriuretic peptide and pro-adrenomedullin were measured using immunoluminometric sandwich assays (BRAHMS, Berlin, Germany). NT-proBNP was measured using the automated Dimension Vista Intelligent Lab System method (Siemens Healthcare Diagnostics Inc., Deerfield, Illinois). CRP was measured by a high-sensitivity assay (Roche Diagnostics, Basel, Switzerland). Cystatin C was measured using a particle-enhanced immuno-nephelometric assay (N Latex Cystatin, Siemens Diagnostics).

DNA extracted from peripheral blood cells was assigned to batches without regard to AF status or personal identity. The batches were genotyped with the same set of reagents using real-time polymerase chain reaction (rtPCR) with 2.5 ng of DNA as the rtPCD template for allelic discrimination (ABI 7900HT; Life Technologies). Genotype calls were obtained using SDS version 2.3 software (Life Technologies) and fluorescence intensity plots curated manually. We selected for genotyping the single nucleotide polymorphism (SNP) with the strongest association at each of the two genetic regions associated with AF in the general population in GWA studies (112,114), on chromosomes 4q25 (rs2200733) and 16q22 (rs2106261). We also included the only SNP at the time that had been associated with AF in candidate gene studies and subsequently replicated, a missense SNP (rs1805123) in the *KCNH2* gene (108,109).

### 3.2 Follow-up with national registers

Sweden has kept a population census since 1749 and is today internationally well known to epidemiologists for holding nation-wide health registers of high quality (123). The use of such registers in medical research is greatly facilitated by the ten-digit Swedish personal identification numbers that are unique for each permanent Swedish resident, maintained by the National Tax Board (124), which help to identify patients across registers. Furthermore, the structure of the Swedish health system, with good access to health care across the country, low patient charges to ensure equal access for all citizens to the public health care and almost no private inpatient treatment, further facilitates population-based epidemiological studies.

For the present studies, we used personal identification numbers to link data from the study baseline examination to information on date of death or emigration from the Total Population Register, maintained by Statistics Sweden (<http://www.scb.se>), and to information from two national health registers: the Hospital Discharge Register (HDR)

and the Cause of Death Register (CDR). The register linkage was performed by the National Board of Health and Welfare (<http://www.socialstyrelse.se/english>), which maintain the HDR and CDR. Diagnoses in the HDR and CDR are coded using the International Classification of Disease (ICD). The 8<sup>th</sup> edition (ICD-8) was used until the end of 1986, the 9<sup>th</sup> edition (ICD-9) between 1987 and 1996 and the 10<sup>th</sup> edition (ICD-10) from 1997 until present.

The HDR was established in 1964 and includes dates of admission and discharge as well as information on primary and contributory diagnoses upon hospital discharge from all public hospitals in Sweden, as described recently (125). Reporting to the HDR has been compulsory since 1987 but the only hospital in Malmö has reported since 1969.

The CDR includes diagnoses from death certificates since 1952, regardless if death occurred outside of Sweden, as described previously (126). The register includes information on underlying and contributing causes of death.

High validity has been described for primary diagnoses of heart failure in the HDR (127,128) and for diagnoses of myocardial infarction in the HDR and CDR (128-134), but the validity of AF diagnoses has not been studied. For the present studies, patients with MI and HF were identified using definitions from previous validation studies: a first primary or contributing diagnosis of MI in the HDR or CDR as code 410 for ICD-8 and I21 for ICD-10, a first primary diagnosis of HF as code 427.00, 427.10, 428.99 for ICD-8, 428 for ICD-9 and I50 and I11 for ICD-10.

AF patients were identified in the HDR and CDR as a first primary or contributing diagnosis of atrial fibrillation or atrial flutter using diagnosis codes 427.92 for ICD-8, 427D for ICD-9 and I48 for ICD-10. As in previous studies, we elected to include both diagnoses of atrial fibrillation and flutter, given the close interrelationship of these diseases (135). To examine the validity of AF diagnoses, we randomly selected 100 patients from the MDCS diagnosed with AF before baseline or during follow-up. ECGs were retrieved from the ECG database (GE MUSE, GE Healthcare) of the Scania region in southern Sweden and were all reviewed by a trained arrhythmologist (Associate Professor Pyotr G. Platonov) and a physician in training at the cardiology department (Dr J. Gustav Smith). Atrial fibrillation was defined as lack of consistent P-waves preceding the QRS complex and irregular R-R intervals and atrial flutter was defined as presence of flutter waves (12). Patient records and any ECGs from the local ECG database of the hospital in the neighboring city, Lund University Hospital (Siemens Megacare), were reviewed for individuals where ECGs were unavailable or with inconclusive ECG findings.

Register linkages were updated consecutively during work with the present thesis, such that follow-up extended until: January 1, 2006 in Study I; January 1, 2007 in Study II; and January 1, 2009 in Study III and V.

## 3.3 Prediction of atrial fibrillation

### 3.3.1 Transformations

Variables with right-skewed distributions underwent natural logarithmic transformation for regression analyses. For ease of comparison, biomarkers were scaled to an SD of 1 and continuous risk factors were scaled to natural quantities: age per 10 years and BMI per 5 units. Genetic polymorphisms were primarily evaluated using additive genetic models, with risk estimates estimated per risk allele, but dominant and recessive models were also explored in the meta-analysis (Study IV). All analyses were performed using SPSS version 16 (SPSS Inc, Chicago, IL), SAS 9.2 (SAS Institute, Cary, NC) or STATA 11.1 (Statacorp, College Station, Texas).

### 3.3.2 Association tests

Association of risk factors with AF was evaluated using Wald tests or likelihood-ratio tests of coefficients from regression analyses: unconditional logistic regression (136) for cross-sectional analyses in Study III and for non-time dependent analyses in Study V, and Cox proportional hazards regression (137) for prospective analyses in Study I, II, III and Study V. Logistic regression was also explored for prospective analyses in Study I, to compare results with Cox regression given the uncertainty in diagnosis dates relative to disease onset. Regression models were adjusted for potential confounders, and sex-specific models were explored in Study I. In Study I and II, conventional risk factors independently associated with AF were determined using regression analysis with backward elimination including age, sex, systolic blood pressure, diastolic blood pressure, use of antihypertensive treatment, body mass index, low-density lipoprotein, high-density lipoprotein, current smoking, history of diabetes mellitus, history of myocardial infarction and history of heart failure. Age- and sex-adjusted models were also assessed for comparison. Survival analyses were performed in individuals free from AF at baseline, and with censoring at death, emigration or end of follow-up. The proportionality of hazards assumption was confirmed by inspection of log-cumulative hazard plots and Schoenfeld's global test. Adequate risk calibration across quantiles was confirmed using the Grønnesby and Borgan test (138) for prospective analyses, as implemented in the STATA package `stcoxgof`, and the Hosmer and Lemeshow test for cross-sectional analyses (139).

Absolute risk of AF was estimated per genotype in Study III using the Kaplan-Meier estimator. Kaplan-Meier plots were constructed to visualize cumulative incidence during follow-up per sex and 5-year age category at baseline in Study I and across

quartiles of a risk score created from biomarkers associated with AF by summing individual z-scores weighted by the beta estimate per SD for each biomarker, in Study II.

### 3.3.3 Attributable risk

Population attributable risk (PAR) estimates (140) in Study I and V were calculated using the standard formula where  $P_e$  is the population risk factor prevalence and RR is the relative risk, which is equivalent to the hazard ratio here:

$$PAR = \frac{Pe [RR - 1]}{1 + Pe [RR - 1]}$$

Adjustment for age was performed using a weighted sum approach (141) in which PARs were calculated per 5-year age group and summed weighted for the proportion of cases in each 5-year age group ( $w_i$ ):

$$PAR = \sum_{k=1}^6 (w_i \times PAR_i)$$

### 3.3.4 Discrimination

Conventional risk factors, blood biomarkers and genetic variants associated with AF were assessed for predictive discrimination using Harrells concordance (C) statistic (142), a generalization of the area under the receiver-operating characteristic (ROC) curve, with confidence interval estimates using a jackknife resampling method as implemented in the STATA package somersd (Study II and III).

### 3.3.5 Risk reclassification

Measures of risk reclassification have been proposed as a more clinically relevant method than discrimination for evaluation of the usefulness of novel risk markers (143). The most commonly used such measure is the Net Reclassification Index (NRI), the proportion of individuals correctly reclassified across risk categories minus the proportion of individuals incorrectly reclassified (144). The NRI for blood biomarkers when added to conventional risk factors was therefore evaluated, using risk category thresholds of <5%, 5% to <15%, and  $\geq 15\%$  as proposed by the Framingham investigators (58). We also analyzed the integrated discrimination improvement (IDI), which is a more continuous measure of reclassification defined as the difference in Yates slopes between models, where the Yates slope is the mean difference in predicted probabilities between



cases and controls (144). However, although risk reclassification analyses are clinically relevant for CAD, where a risk threshold (20% absolute risk) translates into a treatment recommendation, i.e. statin treatment for primary prevention (53), no such thresholds exist for AF. Further, the Framingham investigators have used different thresholds in different publications (58,82). For these reasons, such analyses were explored for biomarkers but not pursued for genetic polymorphisms in Study III.

## 3.4 Literature-based meta-analysis

### 3.4.1 Systematic review

A systematic literature search was performed in PubMed for studies testing the association of AF with single nucleotide polymorphisms (SNPs) in the four genetic regions for which reproducible association had been reported at the time: chromosome 4q25 (rs2200733), chromosome 16q22 (rs2106261) and chromosome 1q21 (rs13376333) identified in genome-wide association studies (112-115), and a missense SNP (rs1805123) in the *KCNH2* gene identified in candidate gene studies (109). The following search criterion was used: “atrial fibrillation” AND (“4q25” OR “16q22” OR “1q21” OR “PITX2” OR “ZFHX3” OR “KCNN3” OR “KCNH2” OR “rs2200733” OR “rs2106261” OR “rs1337633” OR “rs1805123”). Reference lists of included articles were hand searched. We excluded studies reported in other languages than English, using samples of primarily non-European ancestry or reporting only on interventional subgroups, such as patients undergoing coronary artery bypass graft surgery or cardiac catheter ablation procedures. Samples reporting results in more than one study were only included once, but included samples could contribute to both cross-sectional and prospective analyses.

### 3.4.2 Meta-analysis

For each genetic polymorphism, effect estimates were combined across all studies and per study design (case-control, cross-sectional, prospective cohort) using random-effects meta-analysis as described by DerSimonian and Laird (145). Heterogeneity was assessed using Cochran’s Q test for heterogeneity, computed as the sum of the squared deviations of each study’s effect from the weighted mean over the study variance, and the  $I^2$  test, the percentage of total variation across studies that is due to heterogeneity rather than chance ( $I^2 = [Q - df] / Q$ ) (146,147). To assess the potential role of study design or sample age as the source of observed heterogeneity, additional heterogeneity analyses were performed by study design and sample mean age was meta-regressed on point estimates for cross-

sectional and prospective cohort studies. Meta-regression was not performed for case-control studies as most studies had substantial differences in age distribution between cases and controls.

### 3.5 Risk factors for AF in HF

To identify risk factors for AF in HF patients, all patients hospitalized for a primary diagnosis of HF were identified from the HDR. As the exact onset of AF and HF can be difficult to determine and the temporal sequence of diagnoses of AF and HF can vary (42), primary analyses were performed under non-time dependent assumptions. Genetic polymorphisms and conventional risk factors (hypertension, BMI, history of myocardial infarction, diabetes mellitus and current smoking) were tested for association with AF in HF patients using unconditional logistic regression analysis adjusted for baseline age and sex. Genetic polymorphisms were primarily tested using additive genetic models as in previous studies (112-114), and population attributable risk was estimated from these models (140), but genotype-specific risk models were also evaluated. In a first sensitivity analysis, the association of genetic polymorphisms with AF was tested prospectively in patients diagnosed with AF only after HF, using Cox proportional hazards regression with censoring at death or emigration. Results were compared to logistic regression analyses, given the potential differences between onset date and diagnosis date. In a second sensitivity analysis, individuals diagnosed with AF or HF prior to baseline were excluded. Finally, in an analysis of all participants in the entire cohort with DNA, risk estimates for AF in HF patients were compared to estimates in the general population obtained from logistic regression analysis by including multiplicative interaction terms for genotype with HF. Interaction of genetic polymorphisms with age was also explored. All analyses were performed using SAS 9.2 (SAS Institute, Cary, NC).

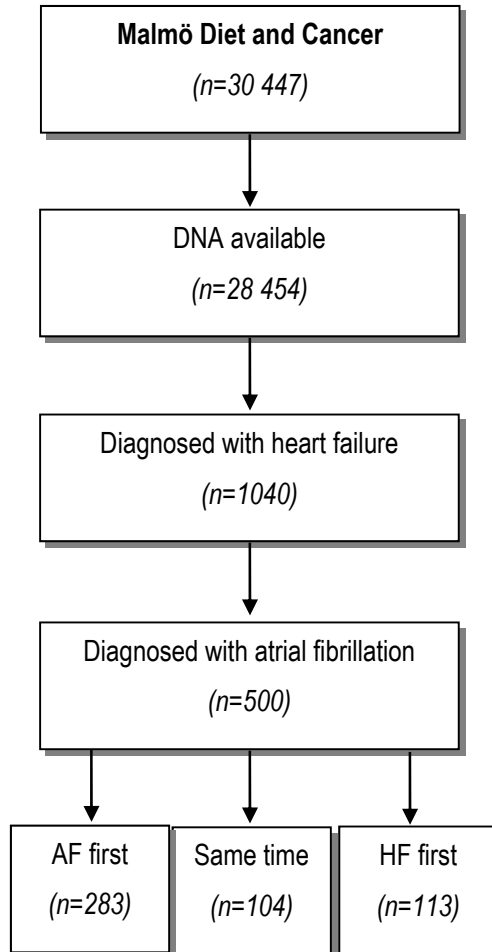


Figure 3.2: Outline of the MDC-HF cohort.

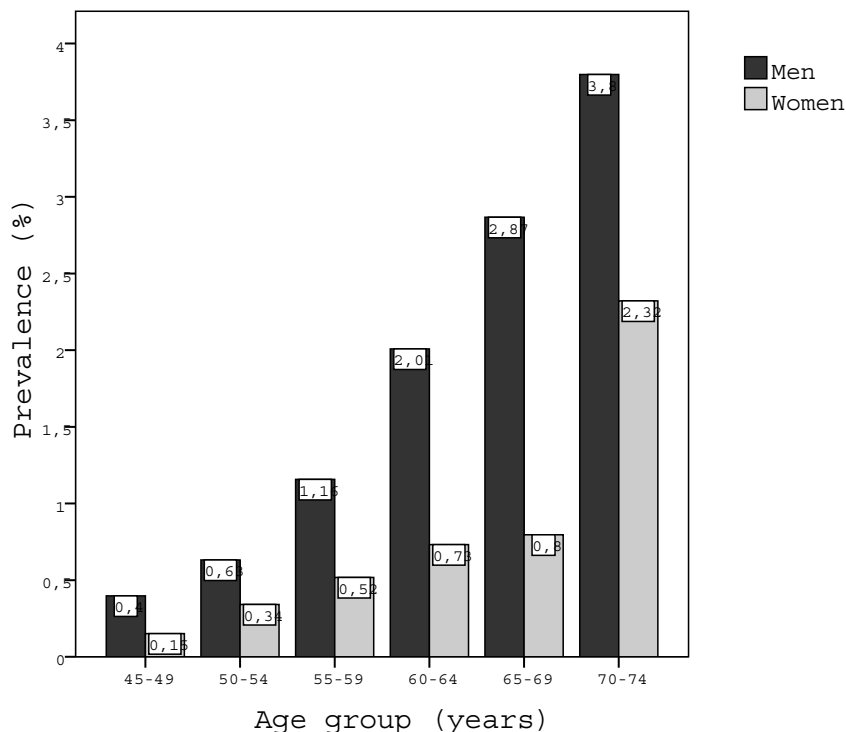
# Chapter 4: Results

From all studies, we excluded individuals with missing date of baseline visit ( $n=6$ ), leaving 18 323 women and 12 118 men aged 44-73 years who were examined at the baseline visit. Baseline characteristics for the entire cohort are shown in **Table 4.1**.

	Men	Women	Total
Age (years)	59.1 (7.0)	57.3 (7.9)	58.0 (7.6)
BMI	26.3 (3.6)	25.5 (4.3)	25.8 (4.0)
SBP (mm Hg)	144.0 (19.5)	139.2 (20.2)	141.1 (20.1)
DBP (mm Hg)	88.0 (9.9)	84.0 (9.8)	85.6 (10.0)
Antihypertensive treatment (%)	19.5	15.9	17.3
Current smoking (%)	28.7	28.1	28.3
History of diabetes (%)	4.2	2.5	3.1
History of heart failure (%)	0.5	0.2	0.3
History of myocardial infarction (%)	4.0	0.7	2.0

**Table 4.1:** Baseline characteristics of MDCS. Continuous variables are presented as mean (SD). Information on BMI, blood pressure and antihypertensive treatment at the baseline visit was available in most individuals ( $n=30\ 352$ ) but information from the questionnaire was missing for a substantial proportion on current smoking ( $n=28\ 559$ ) and diabetes ( $n=28\ 500$ ).

A total of 312 individuals developed AF before the baseline visit, resulting in a prevalence of 10 per 1000 individuals and higher in men (16 per 1000) than in women (6 per 1000). Prevalence increased markedly with age from 2 per 1000 in ages 45-49 to 29 per 1000 in ages 70-74 as shown in **Figure 4.1**. Most AF events were identified from the HDR, and only a few events from the CDR (<0.5%).



**Figure 4.1:** Prevalence of AF by gender and age group at baseline. The prevalence of AF cases in each age group is shown in boxes. The sample size was larger in younger age groups: 5875 (45-49 years), 6360 (50-54), 5593 (55-59), 6152 (60-64), 3649 (65-69) and 2812 (70-74).

## 4.1 Diagnostic validity of register diagnosis

The hundred cases that were randomly selected for validation of AF diagnoses in national registers had a similar baseline age- and sex-distribution as the full sample of AF cases. Using the Scania ECG database we were able to confirm ECGs of 94 individuals as definitive atrial fibrillation or flutter. For the remaining 6 individuals paper ECGs were retrieved from patient records which made validation of 1 additional individual as definitive atrial fibrillation possible. However, 3 individuals did not have atrial fibrillation or flutter; 1 individual instead had sinus tachycardia and pulmonary embolism and 2 individuals had different supraventricular tachycardias of whom one subsequently underwent a successful catheter ablation for concealed Wolff-Parkinson-White syndrome and the other was waiting for ablation at the time of record review. We were unable to retrieve ECGs for two individuals but according to records both had been seen by several physicians including consultant physicians who clearly stated diagnoses of

AF. Both were admitted primarily for other causes, specifically pneumonia and heart failure, and atrial fibrillation was coded as contributory rather than primary.

## 4.2 Conventional risk factors

In the entire cohort (Study I), sex (HR 2.0, 95% CI=1.8-2.2) was associated with incident AF during a mean follow-up of 11.2 years (until January 1, 2006; N=1423 incident AF cases). All risk factors except smoking in women were also associated with AF in sex-specific, age-adjusted models as shown in **Table 4.2**. Results from Cox regression analyses were virtually identical to results from logistic regression analyses in the prospective setting. In stepwise multiple regression models, all risk factors remained significantly associated with incident AF except diabetes in men (HR 1.1, 95% CI=0.98-1.44) and smoking (HR 1.2, 95% CI=0.98-1.44), diabetes (HR 1.4, 95% CI=0.95-2.05) and myocardial infarction (HR 1.01, 95% CI=0.47-2.20) in women. The highest risk estimates were observed for heart failure and myocardial infarction, but given the low prevalence of these diseases in this middle-aged cohort, PAR estimates were substantially lower than for obesity and hypertension.

Variable	Men Prevalence (%)	HR	PAR (%)	Women Prevalence (%)	HR	PAR (%)
Age	-	1.11 (1.10-1.12)	-	-	1.14 (1.13-1.15)	-
BMI >30	13.4	1.88 (1.59-2.23)	11	14.2	1.66 (1.38-2.00)	10
Hypertension, ≥140/90 mm Hg	68.1	1.78 (1.48-2.14)	38	56.3	1.74 (1.42-2.13)	34
Hypertension, ≥160/95 mm Hg	42.8	1.85 (1.60-2.13)	29	33.0	1.69 (1.43-1.99)	25
Current smoking	28.8	1.20 (1.02-1.42)	5	26.6	1.12 (0.92-1.35)	-
History of diabetes	4.1	1.39 (1.02-1.90)	2	2.5	1.67 (1.15-2.43)	2.5
History of heart failure	0.3	4.53 (2.34-8.75)	2	0.1	8.70 (3.60-20.94)	1.4
History of myocardial infarction	3.8	2.03 (1.65-2.49)	5	0.6	1.84 (1.26-2.69)	0.6

**Table 4.2:** Risk estimates and population attributable risks (PAR) for AF in the entire MDCS cohort. Risk estimates are presented as hazard ratios (HR) with 95% confidence intervals from Cox proportional hazards models adjusted for age. PARs were age-adjusted using a weighted sum approach. Prevalent AF cases were excluded. Two definitions of hypertension were compared, based on use of antihypertensives and either blood pressure  $\geq 140/90$  or  $\geq 160/95$  mm Hg.

## 4.3 Blood biomarkers

The MDC-CC included 6103 individuals, of whom blood samples and all conventional risk factors were available in 5187 individuals. 47 individuals (0.9%) were diagnosed with AF before the baseline visit and were excluded from analyses. Baseline characteristics are shown in **Table 4.3**. During a median follow-up of 13.8 years (until January 1, 2007; Study II), 284 individuals were diagnosed with AF.

Variable	
Age (years)	57.6 (5.9)
Male	2094 (41 %)
Systolic blood pressure (mm Hg)	141.4 (19.0)
Diastolic blood pressure (mm Hg)	86.9 (9.4)
Antihypertensive treatment	850 (17 %)
BMI	25.7 (3.9)
Low density lipoprotein (mmol/L)	4.2 (1.0)
High density lipoprotein (mmol/L)	1.4 (0.4)
Diabetes mellitus	399 (8 %)
Current smoking	1379 (27 %)
History of myocardial infarction	75 (2 %)
MP-proANP (pmol/L, n=4,880)	66.1 (50.9-85.9)
NT-proBNP (pg/mL, n=4,778)	61.0 (34.0-111.0)
Midregional pro-adrenomedullin (nmol/L, n=4,879)	0.5 (0.1)
CRP (mg/L, n=4,922)	1.4 (0.7-2.8)
Cystatin C (mg/dL, n=4,777)	0.8 (0.1)
Copeptin (pmol/L, n=4,873)	5.1 (3.2-8.1)

**Table 4.3:** Baseline characteristics of MDC-CC participants free of AF and HF and with information on all conventional risk factors. For continuous variables, mean (standard deviation) are presented for normally distributed variables, and median (interquartile range) for right-skewed variables. For categorical variables, n (%) are presented.

### 4.3.1 Blood biomarkers associated with AF

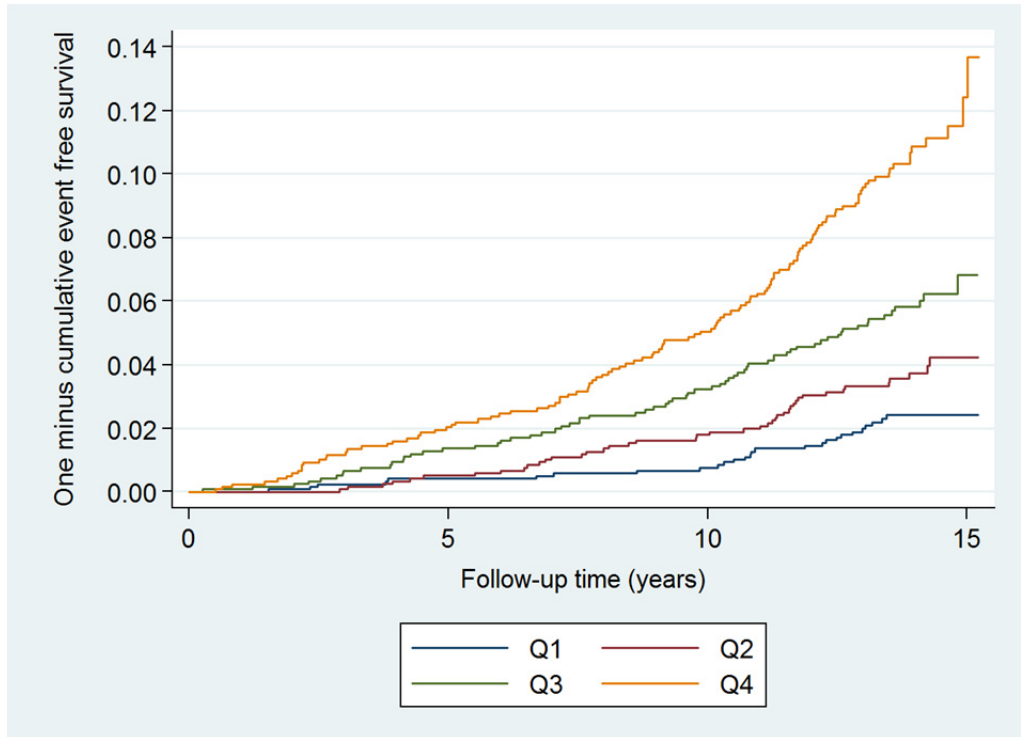
Independently significant conventional risk factors used in final models are shown in **Table 4.4** with corresponding hazard ratios (model 1). In models adjusting for conventional risk factors, MR-proANP, NT-proBNP, MR-proADM, and CRP were individually associated with AF after adjustment for conventional risk factors as shown in **Table 4.4** (model 1). In stepwise regression models with backward elimination including all significant biomarkers and conventional risk factors, MR-proANP and CRP remained



significant as also shown in **Table 4.4** (model 2). In a secondary analysis with censoring at myocardial infarction or heart failure, both MR-proANP and CRP remained significant predictors independently of conventional risk factors.

	AF model 1	AF model 2
Age	1.11 (1.08-1.14)	1.09 (1.06-1.12)
Female sex	0.61 (0.48-0.78)	0.55 (0.42-0.70)
BMI	1.05 (1.02-1.08)	1.05 (1.01-1.08)
SBP	1.01 (1.00-1.02)	1.01 (1.00-1.02)
DBP	0.98 (0.96-1.00)	0.99 (0.98-1.01)
Antihypertensive treatment	1.38 (1.04-1.84)	1.21 (0.89-1.64)
LDL cholesterol	1.05 (1.02-1.08)	0.90 (0.78-1.02)
History of MI	2.64 (1.57-4.44)	1.81 (1.05-3.13)
History of Diabetes	-	-
Current smoking	-	-
lnMR-proANP	1.67 (1.47-1.89)	1.62 (1.42-1.84)
lnNT-proBNP	1.45 (1.28-1.65)	-
MR-proADM	1.26 (1.12-1.41)	-
lnCRP	1.17 (1.04-1.33)	1.18 (1.03-1.34)
CystC	1.11 (1.00-1.24)	-
lnCopeptin	1.09 (0.95-1.26)	-

**Table 4.4:** Risk estimates for AF with conventional risk factors and blood biomarkers in the MDC-CC. Shown are hazards ratios for conventional risk factors and biomarkers per standard deviation with 95 % confidence intervals from Cox proportional hazards models. Results are shown for a model with only conventional risk factors and for single biomarkers with adjustment for conventional risk factors (model 1) and models with all individually significant biomarkers adjusted for traditional risk factors and backward elimination at  $p \geq 0.05$  (model 2).



**Figure 4.2:** Cumulative incidence of AF across quartiles of biomarker scores in individuals free of AF at baseline. Risk score distribution expressed as median (minimum, maximum) was -2.2 (-12.7, -1.3) for the first quartile, -0.7 (-1.3, -0.3) for the second quartile, 0.5 (-0.3, 1.2) for the third quartile, and 2.2 (1.2, 11.2) for the fourth quartile. P-value for trend across quartiles <0.001.

In paper II we also describe the predictive accuracy of conventional biomarkers and blood biomarkers for HF. Although these results are beyond the scope of the present thesis, it is interesting to note that the same independent blood biomarkers were associated with both diseases, with the exception that Nt-proBNP is the natriuretic peptide that was associated with HF whereas MR-proANP was associated with AF.

### 4.3.2 Predictive accuracy

Conventional risk factors predicted incident AF with reasonable accuracy (c-statistic 0.732). Discrimination improved slightly with addition of MR-proANP (c 0.750) to conventional risk factors but only minimally with the addition of CRP (c 0.734) as shown in Table 4.5.

	<b>C-statistic</b>	<b>IDI</b>	<b>NRI</b>
Conventional risk factors	0.732	-	-
MR-proANP	0.750	0.02 (p<0.001)	8 % (p=0.04)
CRP	0.734	0.001 (p=0.25)	2 % (p=0.32)
Both	0.753	0.02 (p<0.001)	7 % (p=0.06)

**Table 4.5:** Discrimination and risk category reclassification of AF with conventional risk factors and blood biomarkers. Shown are measures of discrimination and risk reclassification of AF for models with conventional risk factors only and models with the addition of biomarkers to conventional risk factors for heart failure. For IDI, the IDI statistics and p-values are shown. For NRI, the proportion of individuals correctly reclassified over the proportion of individuals incorrectly reclassified and p-values are shown.

The NRI improved with 8% when including MR-proANP to conventional risk factors (p=0.04). The addition of CRP resulted in a slightly lower improvement compared to conventional risk factors (7%), not reaching statistical significance (p=0.06). The improvement in NRI was mostly due to upward reclassification (19%) of individuals who were diagnosed with AF during follow-up to a higher risk category as shown in **Table 4.6**. Significant improvement of IDI was also observed with biomarkers for both diseases (p<0.001).

<b>Model with conventional risk factors</b>	<b>Model with biomarkers</b>			<b>Total</b>
	<b>&lt;5%</b>	<b>5 to 15%</b>	<b>≥15%</b>	
<5%	57 / 2640	15 / 221	0 /	72 / 2861
5 to 15%	24 / 388	92 / 1228	34 / 137	150 / 1753
≥15%	0 / 0	13 / 64	21 / 105	34 / 169
Total	81 / 3028	120 / 1513	55 / 242	256 / 4783

**Table 4.6:** Reclassification across risk categories for AF with addition of blood biomarkers MR-proANP and CRP to conventional risk factors. Numbers in the cells refer to numbers of events during follow-up over total number of individuals.

## 4.4 Genetic polymorphisms

Clinical data and DNA was available in 26 946 MDCS participants (Study III). In this cohort, 287 individuals had been diagnosed with AF at baseline (prevalence 1.0%). During a follow-up of up to 17.8 years (median 14.1, IQR 12.9-15.7) 2050 individuals (7.3%) developed AF. The Kaplan-Meier estimate of AF risk was 11.9% during the entire follow-up. The call rate was higher than 95% for all three SNPs. Minor allele frequencies, shown in Table 4.7, were similar to previous studies and the European panel (CEU) of the HapMap project.

	Men (n=11,213)	Women (n=17,260)	Total (n=28,473)
Age (years)	59.2 (7.0)	57.4 (7.9)	58.1 (7.6)
BMI (kg/m <sup>2</sup> )	26.3 (3.5)	25.4 (4.3)	25.8 (4.0)
Diabetes (%)	4.5 %	2.1 %	3.1 %
Hypertension (%)	68.9 %	56.8 %	61.5 %
History of myocardial infarction (%)	3.9 %	0.6 %	1.9 %
History of heart failure (%)	0.5 %	0.2 %	0.3 %
History of atrial fibrillation (%)	1.6 %	0.6 %	1.0 %
4q25 (rs2200733, C/T)	10.1 %	10.1 %	10.1 %
16q22 (rs2106261, G/A)	17.6 %	17.4 %	17.4 %
KCNH2 (rs1805123, T/G)	21.2 %	21.2 %	21.2 %

**Table 4.7:** Baseline characteristics and minor allele frequency of genetic polymorphisms. Age and BMI are shown as mean (standard deviation). Categorical variables are shown as percentages. For genetic polymorphisms, major/minor alleles are shown within parentheses followed by minor allele frequencies.

### 4.4.1 Association of polymorphisms with AF

Both SNPs from GWA studies were associated with both incident and prevalent AF with independently of and with similar risk estimates to most conventional risk factors, as shown in Table 4.8. The SNP in *KCNH2* identified through candidate gene studies was not associated with AF. The Kaplan-Meier estimate of AF risk was 27.7% (95% CI=11.6%-57.4%) for homozygotes of the risk allele T of rs2200733 and 11.4% (95% CI=9.9%-13.1%) for C allele homozygotes. For rs2106261, the Kaplan-Meier estimate of AF risk was 18.4% (95% CI=12.8%-26.0%) for homozygotes of the risk allele T and 11.6% (95% CI=9.9%-13.5%) for C allele homozygotes. Few individuals were

homozygous for risk alleles of both SNPs (n=11) but had a high prevalence (9%, n=1) and incidence (45%, n=5) of AF.

Risk factor	Cross-sectional		Prospective	
	OR (95 % CI)	p-value	HR (95 % CI)	p-value
Age	2.12 (1.75-2.57)	<0.001	2.77 (2.57-2.98)	<0.001
Male sex	1.94 (1.48-2.54)	<0.001	1.79 (1.63-1.97)	<0.001
BMI	1.21 (1.03-1.42)	0.07	1.29 (1.22-1.37)	<0.001
Hypertension	2.91 (1.89-4.49)	<0.001	1.46 (1.29-1.65)	<0.001
History of Diabetes	1.80 (1.13-2.87)	0.02	1.25 (1.01-1.54)	0.04
History of MI	1.59 (0.95-2.67)	0.04	1.63 (1.31-2.02)	<0.001
History of HF	18.55 (9.86-34.91)	<0.001	3.22 (1.88-5.53)	<0.001
4q25 (rs2200733)	2.15 (1.69-2.74)	<0.001	1.47 (1.33-1.62)	<0.001
16q22 (rs2106261)	1.28 (1.02-1.61)	0.03	1.13 (1.04-1.12)	0.003
KCNH2 (rs1805123)	0.86 (0.68-1.09)	0.22	1.08 (1.00-1.17)	0.06

**Table 4.8:** Association of conventional risk factors and genetic polymorphisms with AF. Results are presented as effect estimates with 95% confidence intervals per risk factor from multivariable models including conventional risk factors and genetic polymorphisms. Cross-sectional results refer to logistic regression models of prevalent cases at baseline and prospective results refer to Cox proportional hazards models of incident cases during follow-up. Effect estimates for genetic polymorphisms are shown per risk allele, for age per 10 years and for body mass index per 5 units.

#### 4.4.2 Discrimination

Discrimination of AF with genotypes and conventional risk factors is shown in **Table 4.9**. Age and sex at baseline showed a high discrimination for prevalent (C-statistic 0.737) and incident AF (0.738). Small but nonsignificant improvements in discrimination were observed with the addition of single conventional risk factors or genetic polymorphisms (all  $p > 0.05$ ). The addition of all conventional risk factors to age and sex improved C-statistics modestly for prevalent AF (C 0.776) and for incident AF (C 0.750). The addition of genetic polymorphisms further improved the C-statistics for prevalent AF (C 0.785) and for incident AF (C 0.755), although this improvement was not significant ( $p=0.73$  and  $p=0.39$ ).

<b>Model</b>	<b>Cross-sectional C-statistic</b>	<b>Calibration</b>	<b>Prospective C-statistic</b>	<b>Calibration</b>
<b>Basic model</b>				
Age, sex	0.737 (0.711-0.763)	14.5 (p=0.07)	0.738 (0.728-0.748)	15.2 (p=0.08)
<b>Basic model and genetic polymorphisms</b>				
rs22700733	0.751 (0.724-0.779)	11.7 (p=0.17)	0.742 (0.732-0.753)	15.9 (p=0.07)
rs2106261	0.740 (0.713-0.767)	9.5 (p=0.30)	0.739 (0.729-0.750)	18.8 (p=0.03)
Both polymorphisms	0.751 (0.724-0.779)	5.8 (p=0.67)	0.743 (0.733-0.754)	14.4 (p=0.11)
<b>Basic model and conventional risk factors</b>				
Hypertension	0.755 (0.730-0.779)	4.1 (p=0.85)	0.743 (0.733-0.753)	19.9 (p=0.02)
BMI	0.745 (0.719-0.771)	6.0 (p=0.65)	0.747 (0.737-0.756)	8.4 (p=0.49)
Diabetes	0.738 (0.711-0.765)	13.0 (p=0.11)	0.738 (0.727-0.748)	17.4 (p=0.04)
History of MI	0.743 (0.717-0.769)	16.6 (p=0.03)	0.740 (0.730-0.750)	22.7 (p=0.007)
History of HF	0.750 (0.724-0.777)	14.5 (p=0.07)	0.739 (0.730-0.749)	17.2 (p=0.05)
All conventional risk factors	0.776 (0.750-0.802)	2.1 (p=0.98)	0.750 (0.741-0.762)	10.5 (p=0.31)
<b>Basic model, conventional risk factors and genetic polymorphisms</b>				
rs22700733	0.784 (0.757-0.812)	3.2 (p=0.92)	0.754 (0.743-0.765)	10.4 (p=0.32)
rs2106261	0.776 (0.749-0.804)	8.1 (p=0.42)	0.751 (0.741-0.762)	4.0 (p=0.91)
Both polymorphisms	0.785 (0.757-0.813)	5.2 (p=0.73)	0.755 (0.744-0.766)	9.6 (p=0.39)

**Table 4.9:** Prediction of AF with conventional risk factors and genetic polymorphisms. Results are presented as C-statistics with 95% confidence intervals and calibration statistics with p-value for each model. Cross-sectional results refer to logistic regression models of prevalent cases at baseline and prospective results refer to Cox proportional hazards models of incident cases during follow-up. Calibration refers to Hosmer-Lemeshow tests for cross-sectional analyses and Grønnesby-Borgan likelihood-ratio tests for prospective analyses.

## 4.5 Literature-based meta-analysis

### 4.5.1 Systematic review

The literature review identified five publications including five cross-sectional, seven case-control, and six prospective European samples with a total of 12,100 cases and 115,702 controls, testing the association of the SNP on chromosome 4q25 and AF (**Figure 4.3**), including Study III in the present thesis (112,114,148,149). For the SNP on chromosome 16q22, three publications were identified, including five cross-sectional, five case-control, and six prospective samples with a total of 12,694 cases and 132,602 controls (**Figure 4.4**), including Study III in the present thesis (113,114). For the SNP in *KCNH2*, three publications including one cross-sectional, two case-control, and one prospective European samples with 5,272 cases and 59,725 controls were identified (**Figure 4.5**) including Study III in the present thesis (108,109). For the SNPs on chromosome 1q21, *IL6R* and *GJA5*, we only identified the original publications in which the association were initially described (110,111,115), why no meta-analyses were performed for these SNPs.

Characteristics for samples in previous studies included in meta-analyses are shown in **Table 4.10**. All six population-based samples reported distribution of sex and age at baseline.

Study	Sample size (cases/controls)	Case age (mean, SD)	Control age (mean, SD)	Case sex (% male)	Control sex (% male)	Genomic region	Reference
<b>Case-control studies</b>							
Iceland	550 / 4,476	72.5 (11.0)	61.5 (15.8)	67.3	49.2	4q25	(112)
Iceland	2,251 / 13,238	70.5 (13.0)	61.9 (18.4)	59.8	42.7	4q25	(112)
Sweden	143 / 738	74.4 (8.7)	43.1 (12.3)	46.2	59.7	4q25	(112)
USA	636 / 804	NA	67.4 (12.3)	NA	50.9	4q25	(112)
Germany*	2,145 / 4,073	61.0 (11.6)	49.2 (13.9)	72.9	49.2	4q25, 16q22, KCNH2	(114) (108)
Iceland	2381 / 33,723	72.9 (12.0)	NA	59.2	41.4	16q22	(113)
Iceland	970 / 1,939	67.0 (13.5)	NA	66.8	56.1	16q22	(113)
Norway	722 / 711	NA	NA	NA	NA	16q22	(113)
USA	735 / 729	NA	NA	NA	NA	16q22	(113)
Nashville	556 / 598	54.2 (13.7)	56.6 (14.2)	67.8	66.6	4q25	(148)
Italy	78 / 348	64.0 (12.3)	35	NA	NA	4q25	(149)
Boston	790/1,330	63.4 (14.6)	66.4 (12.8)	69.2	53.5	KCNH2	(108)
<b>Cross-sectional studies</b>		<b>Age (mean, SD)</b>		<b>Sex (% male)</b>			
AGES	241 / 2,718	76.5 (5.5)		39.0		4q25, 16q22	(114)
CHS	66 / 3,201	72.3 (5.4)		39.1		4q25, 16q22	(114)
FHS	280 / 4,184	65.6 (12.7)		44.9		4q25, 16q22	(114)
MDCS	287 / 26,659	58.1 (7.6)		39.4		4q25, 16q22, <i>KCNH2</i>	Study III
RS	309 / 5,665	69.4 (9.1)		40.6		4q25, 16q22	(114)
<b>Prospective cohort studies</b>							
AGES	138 / 2,580	76.3 (5.5)		37.2		4q25, 16q22	(114)
ARIC	731 / 7,355	57.0 (5.7)		47.2		4q25, 16q22	(114)
CHS	763 / 2,438	72.2 (5.3)		38.8		4q25, 16q22	(114)
FHS	343 / 3,841	64.7 (12.6)		43.7		4q25, 16q22	(114)
MDCS	2,050 / 24,609	58.1 (7.6)		39.4		4q25, 16q22, <i>KCNH2</i>	Study III
RS	542 / 5,123	69.1 (9.0)		40.3		4q25, 16q22	(114)

**Table 4.10:** Baseline characteristics of samples included in meta-analyses. Characteristics at diagnosis for cases are presented separately from characteristics at recruitment for controls, for case-control studies. Case-control samples are referred to based on geographic region of origin. Baseline characteristics are presented for the entire cohort for cross-sectional and prospective studies. AGES, Age, Gene/Environment Susceptibility study. ARIC, Atherosclerosis Risk in Communities study. CHS, Cardiovascular Health Study. FHS, Framingham Heart Study. NA, not available. RS, Rotterdam Study. \* Sample characteristics retrieved from Sinner et al 2011 (108)

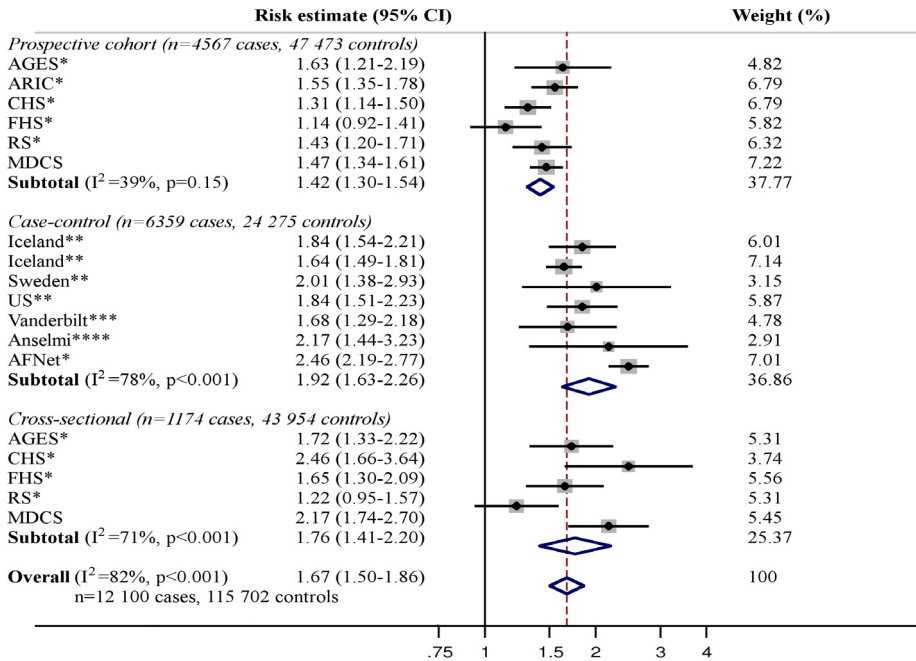


Population-based samples had a similar proportion of male participants but varied widely in distribution of baseline age, ranging from a sample mean of 58.1 (SD 7.6) in the MDCS to 76.5 (SD 5.5) in the Age, Gene/Environment Susceptibility study (AGES). Of the twelve case-control samples, two did not report distribution of sex and age at diagnosis for participants, and four additional samples did not provide characteristics for both cases and controls. The distribution of age and sex varied widely between cases and controls in most case-control studies. The age at diagnosis in cases ranged from 54.2 (13.7) in a case-control study from Nashville to 74.4 (8.7) in a study from Sweden.

#### 4.5.2 Meta-analyses

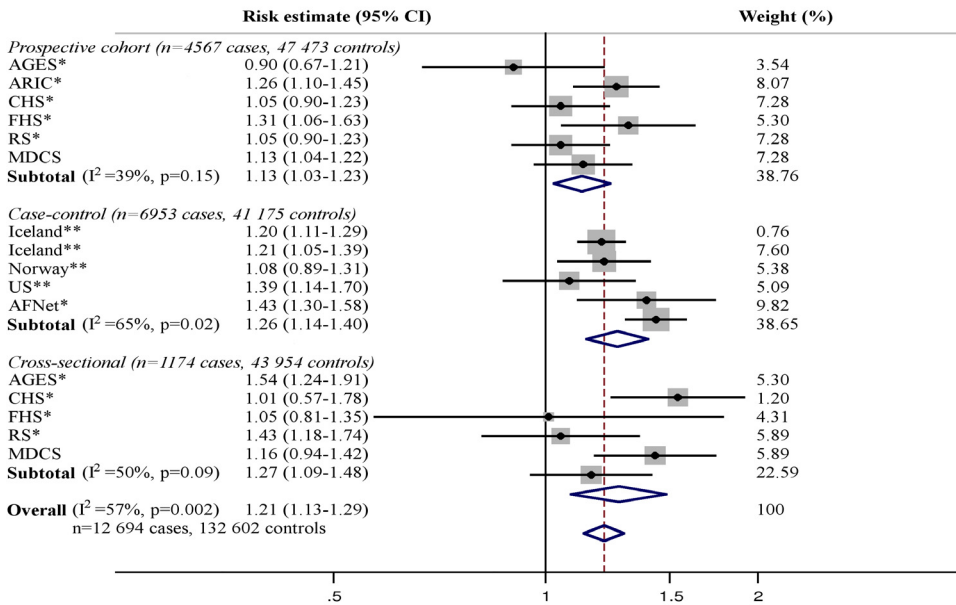
The association of 4q25 and 16q22 with AF was robust when analyzed in random-effects meta-analyses both overall (4q25,  $p=2 \times 10^{-21}$ ; 16q22,  $p=1 \times 10^{-8}$ ) and by study design as shown in **Figure 4.3** and **4.4**. No association was observed in a random-effects meta-analysis of *KCNH2* with AF ( $p=0.15$ ) as shown in **Figure 4.5**.

Significant heterogeneity of relative risk estimates across samples was observed for all three SNPs ( $I^2=57-82\%$ ). When assessed by study design, heterogeneity for 4q25 was observed in case-control ( $I^2=71\%$ ) and cross-sectional samples ( $I^2=78\%$ ) but not prospective samples ( $I^2=39\%$ ,  $p=0.15$ ). Heterogeneity for 16q22 was observed for case-control samples ( $I^2=65\%$ ). When meta-regressed on risk estimates for cross-sectional samples, no association was observed for mean sample age for either of the two SNPs ( $p>0.05$  for both polymorphisms).



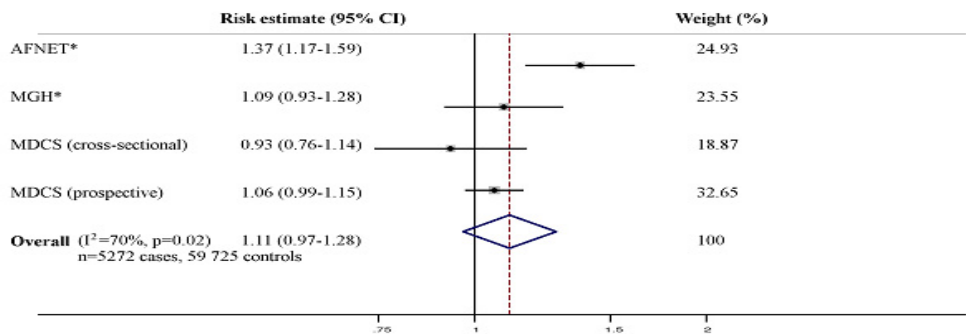
**Figure 4.3:** Meta-analysis of SNPs on chromosome 4q25 and AF.

All results are presented as relative risk estimates (95% confidence intervals) from unadjusted, additive genetic models (adjusted for age and sex in the study by Kääb et al (148) and MDCS). Heterogeneity analysis and random-effects meta-analysis was performed per study design and across all studies. P-values in heterogeneity analyses refer to Cochran's Q test. Weights refer to DerSimonian-Laird weights. \* Benjamin et al 2009 (114). A proxy SNP (rs17042171) was used in this study ( $r^2=1.0$  in HapMap CEU). \*\* Gudbjartsson et al 2007 (112). \*\*\* Kääb et al 2009 (148). \*\*\*\* Viviani Anselmi et al 2009 (149).



**Figure 4.4:** Meta-analysis of SNPs on chromosome 16q22 and AF.

All results are presented as relative risk estimates (95% confidence intervals) from unadjusted, additive genetic models (adjusted for age and sex in MDCS). Heterogeneity analysis and random-effects meta-analysis was performed per study design and across all studies. P-values in heterogeneity analyses refer to Cochran's Q test. Weights refer to DerSimonian-Laird weights. \* Benjamin et al 2009 (114). \*\* Gudbjartsson et al 2009 (113). A proxy SNP (rs7193343) was used in this study ( $r^2=0.78$  in HapMap CEU).



**Figure 4.5:** Meta-analysis of a missense SNP in *KCNH2* and AF.

All results are presented as relative risk estimates (95% confidence intervals) from additive genetic models adjusted for age and sex in MDCS, and additionally for hypertension in AFNET and MGH. Heterogeneity analysis and random-effects meta-analysis was performed per study design and across all studies. P-values in heterogeneity analyses refer to Cochran's Q test. P-values in heterogeneity analyses refer to Cochran's Q test. Weights refer to DerSimonian-Laird weights. \* Sinner et al 2011 (108).

## 4.6 Risk factors for AF in HF

Baseline characteristics for the HF cohort and the entire MDCS cohort are shown in Table 4.11. In contrast to the entire cohort, HF patients were predominantly male (58.0%), with an average age at diagnosis of 72.2 years. A large proportion of HF patients had a history of myocardial infarction at HF diagnosis, 29.4%.

	MDCS (n=28,454)	Heart failure patients (n=1040)
Baseline age, years	58.0 (7.7)	63.5 (6.7)
Age at HF diagnosis, years	-	72.2 (8.1)
Men, %	39.7%	58.0%
BMI, kg/m <sup>2</sup>	25.8 (4.0)	27.8 (4.7)
History of hypertension, %	61.3%	86.1%
History of myocardial infarction, %	2.0%	29.4%
History of diabetes, %	3.1%	11.6%
Current smoking, %	28.8%	32.7%
4q25 (rs2200733, C/T)	10.1 %	11.8 %
16q22 (rs2106261, G/A)	17.5 %	16.2 %

**Table 4.11:** Baseline characteristics. Mean and standard deviation are given for quantitative variables. For history of myocardial infarction, numbers refer to proportion of patients diagnosed prior to baseline or HF diagnosis, respectively. For genetic polymorphisms, major/minor alleles are shown within parentheses followed by minor allele frequencies.

#### 4.6.1 Conventional risk factors

Of clinical risk factors, age, sex, BMI, history of hypertension, myocardial infarction and diabetes were associated with AF in the entire cohort (Table 4.12). In HF patients, increased age at diagnosis (OR=1.22 per ten years, 95% CI=1.04-1.43, p=0.02) but not history of hypertension, BMI, sex, smoking or diabetes (p≥0.05) was associated with increased risk of AF. A history of myocardial infarction was marginally associated with a decreased risk of atrial fibrillation (OR=0.75, 95%CI=0.57-0.99, p=0.04) independently of age. In the first sensitivity analysis, excluding individuals diagnosed with AF prior or simultaneously to HF, no risk factor was associated with AF in HF patients (all p>0.05). In the second sensitivity analysis, excluding individuals diagnosed with HF or AF prior to baseline, only age remained associated with AF (OR=1.44 per ten years, 95% CI=1.17-1.78, p=6x10<sup>-4</sup>).

Risk factor	MDCS		Heart failure patients	
	OR (95% CI)	P-value	OR (95% CI)	P-value
Age, 10 years	2.33 (2.19-2.49)	<0.001	1.22 (1.04-1.43)	0.02
Male sex	1.77 (1.62-1.94)	<0.001	1.30 (1.00-1.68)	0.05
BMI, kg/m <sup>2</sup>	1.06 (1.04-1.07)	<0.001	1.01 (0.99-1.04)	0.33
History of hypertension	1.54 (1.37-1.72)	<0.001	1.15 (0.79-1.68)	0.48
History of myocardial infarction	1.82 (1.48-2.25)	<0.001	0.75 (0.57-0.99)	0.04
History of diabetes	1.31 (1.08-1.59)	0.007	0.77 (0.51-1.15)	0.20
Current smoking	1.09 (0.99-1.20)	0.10	0.85 (0.65-1.11)	0.24

**Table 4.12:** Association of clinical risk factors with AF. Relative risk estimates from multiple regression analyses are presented with corresponding confidence interval and p-value for the entire cohort and HF patients. Risk estimates are presented per 10 years of baseline age for the entire cohort and per 10 years of age at diagnosis in HF patients. Risk estimates for myocardial infarction are presented for prevalent myocardial infarction at the baseline visit for the entire cohort and for prevalent myocardial infarction at the time of HF diagnosis in HF patients. Body mass index and history of hypertension, diabetes and smoking refer to the baseline visit.

#### 4.6.2 Genetic polymorphisms and AF in HF patients

Minor allele frequencies in the HF cohort were similar to in the entire cohort (Table 4.11), and similar to previous AF studies (4q25: 0.17-0.23, 16q22: 0.17-0.19) (112,114) and the European sample (CEU) in the third phase of the HapMap Project (4q25: 0.12, 16q22: 0.16) (150).

Both SNPs were associated with increased risk of AF in HF patients in additive models (16q22 OR 1.75 per risk allele, 95% CI=1.35-2.26,  $p=2 \times 10^{-5}$ ; 4q25 OR 1.57 per risk allele, 95% CI=1.18-2.09,  $p=0.002$ ). Risk of AF by genotype and in additive models is shown in Table 4.13. Adjustment for established risk factors shown in Table 4.12 did not attenuate the SNP associations (data not shown). We observed no significant interaction of polymorphisms with age at diagnosis ( $p$  for interaction  $>0.05$ ). The AF risks in HF patients attributable to polymorphisms were 19% and 12% for SNPs on 16q22 and 4q25, respectively.

In the first sensitivity analysis, individuals diagnosed with AF prior or simultaneously to HF diagnosis were excluded, leaving 653 individuals with HF. During a median follow-up of 1.7 years, 113 individuals were diagnosed with AF. In this sample, the SNP on 16q22 remained strongly associated with AF risk (HR 1.96 per risk allele, 95% CI=1.40-2.73,  $p=8 \times 10^{-5}$ ) but not 4q25 (HR 1.14 per risk allele, 95% CI=0.75-1.75,  $p=0.53$ ). In prospective analyses, risk estimates were similar when obtained from Cox proportional hazards regression and logistic regression models. In a second sensitivity analysis, individuals diagnosed with HF ( $n=87$ ) or AF ( $n=69$ ) prior to baseline

were excluded. Both SNPs remained associated with AF risk (4q25 OR 1.40 per risk allele, 95% CI=1.03-1.91,  $p=0.03$ ; 16q22 OR 1.76 per risk allele, 95% CI=1.35-2.30,  $p=3 \times 10^{-5}$ ).

Compared to HF patients with no risk alleles ( $n=547$ ), AF risk increased in a dose-dependent fashion in patients with one risk allele ( $n=345$ , OR=1.46, 95% CI=1.13-1.89,  $p=0.004$ ) or two risk alleles ( $n=89$ , OR=2.04, 95% CI=1.30-3.22,  $p=0.002$ ). Individuals with three ( $n=5$ ) or four ( $n=1$ ) risk alleles were all diagnosed with AF (Fisher's exact  $p=0.03$  and  $p=0.48$ , respectively).

Genotype	MDCS		P-value	HF patients		P-value	Participants without HF		P-value
	N (%)	OR (95% CI)		N (%)	OR (95% CI)		N (%)	OR (95% CI)	
4q25 (rs2200733)									
Major homozygote (CC)	21,927 (80%)	1.00	-	772 (78%)	1.00	-	21,155 (81%)	1.00	-
Heterozygote (CT)	4982 (19%)	1.62 (1.46-1.81)	2x10 <sup>-19</sup>	213 (21%)	1.59 (1.17-2.17)	0.003	4769 (18%)	1.61 (1.44-1.81)	7x10 <sup>-16</sup>
Minor homozygote (TT)	265 (1%)	2.13 (1.48-3.08)	6x10 <sup>-5</sup>	11 (1%)	2.00 (0.58-6.91)	0.003	254 (1%)	2.18 (1.46-3.27)	1x10 <sup>-4</sup>
Additive	27,174	1.58 (1.44-1.73)	6x10 <sup>-22</sup>	996	1.57 (1.18-2.09)	0.28	26,178	1.58 (1.42-1.75)	3x10 <sup>-18</sup>
HF interaction			0.83			0.002			
16q22 (rs2106261)									
Major homozygote (GG)	19,214 (68%)	1.00	-	724 (70%)	1.00	-	18,490 (68%)	1.00	-
Heterozygote (AG)	8147 (29%)	1.11 (1.01-1.22)	0.03	280 (27%)	1.60 (1.21-2.12)	0.001	7867 (29%)	1.08 (0.98-1.21)	0.14
Minor homozygote (AA)	849 (3%)	1.40 (1.12-1.77)	0.004	27 (3%)	2.86 (1.23-6.66)	0.01	822 (3%)	1.39 (1.08-1.79)	0.01
Additive	28,120	1.14 (1.05-1.23)	0.001	1031	1.75 (1.35-2.26)	2x10 <sup>-5</sup>	27,179	1.12 (1.03-1.22)	0.01
HF interaction			7x10 <sup>-4</sup>						

**Table 4.13:** Association of genetic polymorphisms with AF. Relative risk estimates are presented per genotype and per risk allele in an additive genetic model, with corresponding confidence interval and p-value. Risk estimates are presented in the entire cohort, in HF patients and in participants without HF. P-values are also presented for interaction of linear genetic model with heart failure status in the entire cohort.



### 4.6.3 Interaction analyses in the entire MDCS cohort

In the entire cohort, AF risk was higher in HF patients (OR 9.24, 95% CI=8.03-10.64,  $p=1 \times 10^{-210}$ ) and in risk allele carriers for SNPs on both 4q25 (OR 1.58, 95% CI=1.44-1.73,  $p=6 \times 10^{-22}$ ) and 16q22 (OR 1.14, 95% CI=1.05-1.23,  $p=0.001$ ) as shown in **Table 4.13**. AF risk estimates were similar to those reported in previous population-based studies for SNPs on 4q25 (**Figure 4.3**) and 16q22 (**Figure 4.4**) (114). AF risk by genotype was significantly magnified in the presence of HF for 16q22 ( $p$  for interaction =  $7 \times 10^{-4}$ ) but not 4q25 ( $p$  for interaction = 0.83), as detailed in **Table 4.13**. Interaction terms of SNPs with age at HF diagnosis were not significant ( $p>0.05$ )

# Chapter 5: Discussion

Although previous studies have identified many risk factors for AF (151), only one study had evaluated the population attributable risk of conventional cardiovascular risk factors for AF before work on this thesis began (152). Furthermore, although previous studies have associated increased levels of natriuretic peptides and CRP with AF (153,154) the predictive accuracy had not been studied. Genetic polymorphisms at three genetic loci (4q25, 16q22 and in *KCNH2*) had been reproducibly associated with AF (109,112-114) but the predictive accuracy and effect consistency of these polymorphisms had not been studied. We studied these questions in a large, population-based cohort of middle-aged participants who were followed for up to 17.8 years. Finally, we are not aware of any studies that have evaluated risk factors other than NYHA class for AF in HF patients, and explored this in the subcohort of patients diagnosed with HF before baseline or during follow-up.

## 5.1 Diagnostic validity and generalizability of findings

We observed high case validity for AF in patients receiving a hospital diagnosis of AF, with 95% of patients classified as ‘definite AF’ and additionally 2% classified as ‘likely AF’, indicating minimal impact of case misclassification bias. The case validity was thus identical to that in a study which validated AF diagnoses in the Danish National Patient Registry in which 112 of 116 (97%) diagnoses could be verified (155). As expected, the case validity of AF diagnoses was higher than for HF and MI, which have been reported previously as 82-87 (127,128) and 86-100% (128-134) respectively, and similar to a primary diagnosis of HF, 95% (127). An American study in a cohort including elderly participants also estimated the case validity of diagnoses in hospital registers to 99% (156). However, in elderly individuals, the rate of other supraventricular tachyarrhythmias, which may be misinterpreted as AF, is likely to be lower relative to AF. These findings of high case validity reflect that the AF diagnosis is based entirely on ECG findings, while HF and MI diagnoses are based on a combination of ECG findings, blood biomarkers, symptoms and clinical findings.

Our use of national hospital discharge and cause of death registers reduces the impact of attendance bias on endpoint ascertainment, but also means that AF cases that

were diagnosed outside hospitals, i.e. in primary care, escape detection, potentially leading to underestimates of disease rates. Unfortunately, information on the sensitivity of AF diagnoses in national hospital registers is currently not available, as national databases with diagnoses from primary care do not exist, but efforts to collect such information using local primary care databases are ongoing. However, the clinical experience of the investigators for the present studies is that in the relatively densely populated southern part of Sweden, with short distances to the nearest hospital, most new-onset AF patients are referred to the hospital, with the exception of elderly and immobilized patients. As patients in the present study are middle-aged and followed for a median of 11-14 years, we believe our endpoint should have acceptable sensitivity.

As we did not regularly evaluate participants with resting ECGs or Holter registrations, we also did not detect undiagnosed and asymptomatic cases. Although a study including elderly participants reported finding a number of additional AF cases when screening with resting ECGs (156), that number is likely to be lower in our middle-aged cohort. In the same study, additional screening with 24-hour Holter monitoring did not contribute to identification of additional cases (157).

The sensitivity of our endpoint is therefore likely to be somewhat lower than in some of the other population-based studies that have studied AF, such as the Framingham Heart Study where 12-lead ECGs are recorded at biennial examinations, but is similar to the Danish Diet, Cancer and Health study (155), and the Renfrew/Paisley study with the exception that ECGs were not recorded at baseline in MDCS (158).

The total prevalence observed in our study, about 1%, is in agreement with estimates from previous population-based American studies (15,159). The age category-specific prevalence estimates were similar to those observed in a screening with 24-h ambulatory ECG registrations in men from Malmö (160). Age category-specific estimates were also similar to those observed in the general American population but with lower estimates in the 70-74 years category at baseline (15), which could reflect recruitment bias towards healthy individuals in MDCS. Furthermore, previous studies used different sets of exclusions, such as AF after cardiac surgery or association with hyperthyroidism, which further complicates comparisons. As in previous studies, we found AF to be a disease mainly affecting older adults, with 70% of cases >60 years at diagnosis (159). The prevalence of AF below 50 years of age was low, <0.5%, and increased gradually across middle age. Like in previous studies the prevalence was higher in men across all age categories.

As in previous studies, we did not discriminate between atrial fibrillation and flutter given the close interrelationship of these diseases (135) or between permanent, persistent and paroxysmal AF which means that findings can only be generalized to the same broad definition of AF. Finally, findings in our middle-aged population with long-term follow-

up might not be directly generalizable to other populations, especially not young patients with lone AF but this group is quite small (161).

## 5.2 Conventional risk factors

Conventional cardiovascular risk factors are actively screened for and monitored in primary care and often treated, primarily to prevent CAD. Despite reduced population burden of conventional cardiovascular risk factors and declining incidence of CAD in recent years, the prevalence of AF is increasing. Our results confirm previous studies which link conventional risk factors to AF and suggest that interventions targeting conventional modifiable risk factors for CAD could further reduce the population burden of AF, with high population attributable risk (PAR) estimates for risk factors such as hypertension and obesity (Study I). The large PAR estimates for hypertension and obesity are supported by recent findings in two American cohorts, the Framingham Heart Study (FHS) and the Atherosclerosis Risk in Communities (ARIC) study (152,162). In the somewhat younger ARIC cohort, where hypertension was less common but obesity was more common than in the MDCS, the PAR estimate for obesity was similar to MDCS (13%) and for hypertension was lower than MDCS (22%). In the older FHS cohort, the PAR estimate for hypertension was the highest for any risk factor and was lower (14%) than in MDCS, while the prevalence of manifest cardiac disease was higher and explained a large proportion of AF risk. However, as causality cannot be established by observational studies which can be subject to unmeasured confounders, hypotheses about reduced AF burden from risk factor treatment need to be tested in interventional studies.

We also found that conventional cardiovascular risk factors predict AF with reasonable accuracy (Study II). However, in a subsequent analysis (Study III) we noted that the contribution of cardiovascular risk factors to predictive accuracy was modest when added to age and sex. These findings are difficult to interpret, as age both appears to be an independent risk factor and is a measure of combined exposure time for other measured risk factors, and potentially for exposure time for unmeasured risk factors. A large number of other independent risk factors for AF have also been reported (151). Interestingly, age was recently shown not to be associated with atrial fibrosis, which was however associated with AF, in autopsy series (163).

## 5.3 Blood biomarkers

In an attempt to identify novel blood biomarkers for AF that could provide incremental information to conventional risk factors, we evaluated a set of six blood biomarkers reflecting diverse pathophysiological pathways including hemodynamic stress, plasma volume and osmolarity, inflammation and renal function (Study II). Interestingly, although all biomarkers except copeptin were associated with AF independently of conventional risk factors, only MR-proANP and CRP were associated with AF in models including all biomarkers. Our data also suggest that the association of biomarkers with AF is not mediated by interim MI and HF. A modest improvement in predictive accuracy was observed with MR-proANP but not with CRP. The independent association with MR-proANP was in contrast to findings for HF in Study II, where NT-proBNP and CRP were independently associated with HF. ANP is secreted from the cardiac atria while BNP is produced by the ventricles, both in response to increased wall tension in the corresponding cavity. ANP is also produced from the cardiac ventricles in individuals with dysfunction of the ventricular myocardium (164). Our results suggest that MR-proANP, a mid-regional fragment of the propeptide to ANP, might better than NT-proBNP reflect atrial stress leading to atrial fibrillation. These findings are supported by findings in the FHS that N-terminal pro-atrial natriuretic peptide predicted AF slightly more strongly than NT-proBNP (154).

The association of CRP with incident AF and HF has been reported previously, and has been postulated to reflect active inflammatory processes in the myocardium which, following an initial precipitating event, progressively drive remodeling and dilatation of cardiac chambers (153,165). A recent large population-based study did not find association of genetic elevations in CRP with AF (166), arguing against a direct causal role of CRP for AF. In contrast, a polymorphism in the gene encoding the receptor for Interleukin-6 (*IL6R*), the central mediator of the acute phase response, was recently associated with increased risk of AF, providing additional evidence of a role of inflammation in AF (110). However, further data are warranted to establish the underlying mechanisms, and the improvement in discrimination with CRP was modest. Our findings that natriuretic peptides modestly improve predictive accuracy, while CRP although associated with AF does not contribute to predictive accuracy, were also recently confirmed in the FHS (167). More recently, studies have linked other inflammatory markers to AF as well (168-170). It is possible that better markers reflecting myocardial remodeling and other pathways exist and could improve risk prediction beyond natriuretic peptides.

## 5.4 Genetic polymorphisms

In meta-analyses including up to 150 000 individuals of European ancestry (Study IV), we observed robust association of genetic polymorphisms on chromosomes 4q25 and 16q22 with AF, both across all studies and within each individual study design. Relative risk estimates varied widely in magnitude across published studies and across case-control studies and cross-sectional studies, with 50-78% of variation due to heterogeneity rather than chance, but not across prospective cohort studies. Overall relative risk estimates were modest, 1.67 per risk allele for the polymorphism on chromosome 4q25 and 1.21 for the SNP on 16q22. The very high odds ratios reported in some studies were confined to individual case-control or cross-sectional samples.

The origin of the observed heterogeneity remains uncertain. Although a stronger association with a family history of AF has been reported for AF presenting at a young age (79,80), we did not detect any effect interaction with age in meta-regression analyses. A particularly strong association with family history has been reported in individuals presenting with lone AF (81) and a previous study supports a larger effect of at least the polymorphism on chromosome 4q25 with lone AF (115). In the light of these observations, it was not entirely unexpected that case-control studies varied widely in effect estimates, as they vary greatly in exclusion and inclusion criteria and ascertainment method. For example, cases with concomitant heart failure, valvular disease or hyperthyroidism have been excluded from some case-control studies. The observation of heterogeneity across cross-sectional studies was more unexpected. Cohorts differ in age distribution but while age-dependent genotypic effects have been described for genetic polymorphisms associated with other traits (171), meta-regression analyses did not support any association between sample age and effect size. Effect heterogeneity across populations could also be due to varying linkage disequilibrium between genotyped polymorphisms with an ungenotyped, causal variant. However, a recent fine-mapping study of the 4q25 region did not find any stronger association than with the polymorphism genotyped here (172). More likely, the observed heterogeneity across cross-sectional studies is explained by differences in population characteristics based on differences in recruitment, exclusion criteria and endpoint ascertainment. For example, participants with prevalent heart failure, valvular disease or coronary artery disease were excluded from genotyping in the Cardiovascular Health Study. These observations have implications for the interpretation of risk estimates from genetic studies of AF and other complex diseases.

When tested for predictive accuracy in a prospective setting in the MDCS (Study III), risk estimates for the SNP on chromosome 4q25 was found to be similar to many of the conventional risk factors when considered individually, but to contribute minimally and non-significantly to predictive accuracy as measured by the C-statistic. It has

previously been observed that the C-statistic scales with the magnitude of risk estimates (173). Our findings thus do not support the utility of clinical genotyping for AF prediction with these SNPs. It seems unlikely that other population-based studies using different AF ascertainment methods such as regular ECG recordings would provide different results, as effect estimates were similar across population-based cohort studies in the meta-analysis in Study IV.

Recently, additional SNPs have been associated with AF (Table 1.3). However, for SNPs on chromosome 1p21 and in *GJA5* and *IL6R*, only the initial discovery studies have been reported to date. Future studies will be needed to evaluate the robustness of the reported effect estimates and the association with AF in the general population. Our large, population-based study and meta-analysis did not support the association of AF with the K897T missense variant in *KCNH2*. Although replication was reported in the initial report, the association did not replicate in a recent large case-control sample with similar age distribution and clinical characteristics to the sample in which the association was first identified (108). The present results do not support the large effect described in the initial report (109), but cannot rule out a smaller effect.

Future studies will need to address the effect of these SNPs in populations of other ancestries than European. When our meta-analysis was performed, studies in other ancestries were largely lacking. However, we note that recent studies have reported association of 4q25 and 16q22 in individuals of Asian (112,174-176) and African American ancestry (110,177).

## 5.5 Risk factors for AF in HF

In Study I, HF was by far the strongest risk factor for incident AF, but the prevalence was low in our middle-aged cohort resulting in small PAR estimates. In Study V, we evaluated risk factors for AF in HF patients and did not observe association with the conventional risk factors for AF in the general population – hypertension, BMI, smoking and diabetes. However, associations were observed with the two genetic polymorphisms on chromosomes 4q25 and 16q22. The SNP on chromosome 16q22 but not 4q25 remained robustly associated with AF in sensitivity analyses restricted to patients with AF diagnosed after HF. Moreover, our findings suggest context dependency on AF risk for the chromosome 16q22 polymorphism, with a substantially larger relative risk of AF in HF patients, 75% per copy, than in the general population, 14% per copy. We did not explore the association of blood biomarkers with AF as these were only available in the smaller subcohort MDC-CC.

Our findings might provide new clues on mechanisms linking HF to AF. Although certain HF etiologies such as infiltrative, infectious or deposition diseases could

potentially act on both atria and ventricles, the lack of association of the major AF risk factors with AF in HF patients in the present study argue against independent effects of these factors on atria and ventricles, and suggest that association between cardiovascular risk factors and AF could instead be mediated by ventricular dysfunction resulting in increased atrial stretch, neurohormonal activation and ion channel remodeling (46). In that context, measures of HF severity, ventricular filling pressures and timeliness of treatment would be expected to be more important determinants of AF.

Our findings also demonstrate the relevance to AF in HF patients of genetic polymorphisms on chromosomes 4q25 and particularly 16q22 found in previous studies to be associated with AF in the general population, providing initial evidence of a genetic component to AF in the context of HF. Furthermore, the observed interaction of the SNP on chromosome 16q22 with HF suggests that the gene influenced by this SNP might play a more important role in the pathophysiology of HF. Unfortunately, to date no published studies have addressed pathophysiological mechanisms linking the association of this SNP with AF.

The major strength of our study was the very large size of the source population sample, which allowed detection of a large, population-based cohort of HF patients with AF (**Figure 3.2**). The use of a source population cohort also makes our study more representative of the general HF population than previous studies of AF in HF patients which predominantly included younger, male participants of randomized controlled trials (43). Our cohort of hospitalized HF patients was similar to previous population-based studies of HF (178,179) with a high age at diagnosis (mean 72 years), a slight majority of patients of male sex (58%) and a history of MI in about a third of HF patients.

Our study has several limitations which merit consideration. First, only patients hospitalized for a primary diagnosis of HF were detected. In contrast to diagnoses in the outpatient setting and patients hospitalized with a contributing diagnosis of HF, such diagnoses have shown low inter-reader variability (127), but are likely to bias our sample towards more severe cases as evidenced by the high rate of AF. Although generalizability to outpatients seems likely, additional studies need to address this question. Second, our study design did not allow information on clinical variables such as left ventricular ejection fraction (LVEF), NYHA class, left ventricular filling pressures, specific etiologies, use of medications or details on circumstances surrounding AF onset. Third, undetected AF events may have resulted in underestimates of AF risk, but is not likely to have caused false positive associations. Fourth, as study participants did not attend follow-up examinations, it is possible that regression dilutions bias may have attenuated risk estimates for conventional risk factors. Thus, replication of our results in a more carefully clinically characterized HF cohort seems warranted.





# Chapter 6: Perspectives

## 6.1 Prediction and prevention of AF

Despite the large PARs for conventional cardiovascular risk factors observed in MDCS, ARIC and FHS, at least half of population risk remains unexplained. This is in contrast to findings for CAD, where the Interheart study suggested that known and modifiable risk factors explain up to 90-94% of risk of myocardial infarction (180). Furthermore, risk scores for easy use to predict AF have recently been created, using estimates for conventional risk factors from American population-based cohort studies, but have shown modest predictive accuracy (58,59,66). Such scores could guide interventions to prevent AF and, more importantly, potential consequences of AF including stroke, dementia and mortality. However, AF is often viewed as a more complex disease than CAD. Specifically, population-based studies did not take into account the large number of non-cardiac diseases which have been associated with AF, including endocrine diseases (181-184), pulmonary diseases (156,185,186), renal diseases (187), diseases of the nervous system (188), autoimmune diseases (189,190) and infectious diseases (191), as well as AF complicating surgical procedures (192-195), binge drinking (196), athletic training (197) or use of certain medications (198). Although all these parameters may not be directly causal and individually only contribute modestly to the population burden of AF, solid prediction seems unlikely without inclusion of a large number of parameters with potentially complex interactions. Substantially more precise prediction models will need to be developed and subjected to randomized controlled trials if models are to guide use of medications presently used to prevent AF episodes (antiarrhythmics) or anticoagulants as such medications confer substantial risks (12). However, it is important to note that the predictive accuracy with current models including conventional cardiovascular risk factors is surprisingly good and much could potentially be gained by targeting modifiable risk factors by lifestyle interventions or pharmacologic lowering of blood pressure. Such interventions have not been convincingly shown to reduce risk of AF, but have been shown to reduce the risk of CAD and stroke and are associated with limited adverse effects. It could therefore be argued that the clinical role of AF prediction models today might simply be to confirm a high burden of conventional risk factors, as potential motivation to achieve risk factor control, in addition to results from models for prediction of CAD and stroke. The author of this thesis would however be careful with using the AF prediction models even in this

context. Because of the strong emphasis on age in these models, the absolute risk of AF will still be low in middle-aged individuals with a high burden of risk factors as shown in the following two examples.

**Example 1.** A 60-year old male, 175 cm tall, with overweight (95 kg), current smoking and a sedentary lifestyle but not using any medications, has no history of CVD or diabetes and no murmur or signs or symptoms of HF. Examinations show hypertension (170/100 mm Hg), a PR duration of 160 msec, no electrocardiographic evidence for left ventricular hypertrophy, total cholesterol of 7.0 mmol/L and HDL cholesterol of 0.8. Absolute 10-year risks according to FHS models (53,57,58): AF risk, 7%. CAD risk,  $\geq 30\%$ . Stroke risk, 15%.

**Example 2.** A 80-year old male, 175 cm tall, with normal weight (70 kg) who has never smoked, leads an active life, uses no medications and has no history of CVD or diabetes, no cardiac murmur and no signs or symptoms of HF. Examinations show borderline hypertension (140/90 mm Hg), a PR duration of 160 msec, no electrocardiographic evidence for left ventricular hypertrophy, total cholesterol of 5.0 mmol/L and HDL cholesterol of 1.4. Absolute 10-year risks according to FHS models (53,57,58): AF risk, 15%. CAD risk, 16%. Stroke risk, 15%.

## 6.2 The genetic architecture of AF and implications for prediction

A genetic contribution to AF has been established in several studies (79-84), but in a recent study family history, when evaluated using several different definitions including premature AF, only provided a minimal improvement in predictive accuracy (82). However, genetic polymorphisms could still provide clinically meaningful improvements, as genetic polymorphisms in that study was independent of family history (82), a finding also observed in studies of CAD (199,200). Our findings of only minimal and non-significant improvement in predictive accuracy do not support utility of including these first identified genetic polymorphisms into risk prediction scores. We note that a number of additional AF loci have been identified after genotyping for the present studies was initiated as shown in **Table 1.3**. However, these loci conferred smaller risks and are unlikely to alter our conclusions. Further studies are needed to evaluate the collective impact of common genetic variants on AF risk more comprehensively, e.g. using whole-genome SNP data. Furthermore, the hypothesis that multiple rare variants of large effect could explain a substantial proportion of AF risk remains to be tested, using e.g. large-scale genotyping for the recently developed exome-chip or resequencing. Such studies are ongoing and results are expected in the near term.

If one should still be interested in genotyping for the polymorphisms studied in this thesis, as currently offered directly to consumers by commercial companies with provision of absolute risk estimates (107), our findings of substantial heterogeneity highlight that GWA studies constitute a discovery tool and the importance of interpreting risk estimates in the context of the original study. These findings also have implications for genetic testing of polymorphisms for many other diseases, as few studies have evaluated the robustness of risk estimates across study designs. Although between-study heterogeneity might be considered unsurprising, our findings are in contrast to findings for CAD where no significant heterogeneity was detected for genetic polymorphisms in large-scale meta-analyses (199,201). In our opinion, much additional work is needed before genotypic risks could be reliably computed and it remains a question of debate whether such genotyping could at all be motivated, even if improved predictive accuracy should be achieved, in the absence of preventive treatments.

Finally, whether genetics contribute to predictive accuracy or not, the identification of novel loci and genes associated with AF is likely to provide important novel insights into the pathophysiology of AF (202), as evidenced by the findings of SNPs on chromosomes 4q25 and 16q22, where none of the proximal genes have previously been implicated in AF pathophysiology. Recent studies have begun to dissect mechanisms linking polymorphisms on 4q25 to AF, but such work is still only in an early phase. Specifically, the most proximal gene to the SNP, the *PITX2* gene encoding a transcription factor expressed in the developing heart during embryonic cardiogenesis, has been found in animal models to play a role in the formation of the pulmonary vein myocardium (203), where ectopic impulse generation triggering AF onset has been reported in humans (204), and to inhibit left-sided atrial pacemaker specification, such that haploinsufficient mice showed increased expression of sinoatrial node-specific genes in the left atrium and increased propensity to atrial arrhythmia (205,206). Genes near the SNP on chromosome 16q22 have yet to be studied in relation to AF. However, it is interesting to note that the SNP is intronic to the gene Zinc Finger Homeobox 3 (*ZFHX3*), encoding another transcription factor with cardiac expression. Little is known about the function of *ZFHX3* but it has been shown to interact strongly with the Protein Inhibitor of Activated Stat3 (PIAS3), an inhibitor of STAT3 (207), which in turn is a regulator of paracrine circuits in the heart essential for interstitial matrix deposition balance and capillary vasculature maintenance (208). Increased expression of STAT3 has been observed in animal models of AF and proposed to contribute to atrial matrix deposition (209,210).

## 6.3 Heterogeneity of AF

In everyday clinical practice, AF is currently subdivided based on duration (paroxysmal, persistent, long-standing persistent or permanent) and by the extent of symptoms (12). This subclassifications helps inform therapeutic decisions and provides some indirect information about the atrial substrate. Recent authors have argued that a more pathophysiologically oriented classification could facilitate both therapeutic decisions and research on AF (151,211). A working group of the European Heart Rhythm Association and the Atrial Fibrillation competence network (151) recently suggested subdivision of AF into several groups including a group characterized by “focal AF”, characterized by atrial ectopy and many short episodes of self-terminating AF in the absence of structural heart disease, which might be more amenable for rhythm control strategies such as catheter ablation than the larger group of AF cases with more complex substrates, but could develop into more complex AF with time.

It is therefore interesting to note very large effects of the genetic variant on chromosome 4q25 has been described in cohorts with lone AF (115), which might correspond to a phenotype similar to that for “focal AF”, and could potentially explain at least part of the observed heterogeneity for this SNP. Importantly, a recent study found that risk allele carriers on 4q25 had a greatly increased risk of AF recurrence after catheter ablation, but this finding has yet to be replicated (212). In Study V, we noted a similarly interesting finding, that the SNP on chromosome 16q22 was associated with a significantly higher risk of AF in patients with HF compared to the general population, suggestive of a role in the pathophysiology of this more complex AF phenotype. Further studies are needed to clarify the mechanistic links between these SNPs and the pathophysiology of AF, but the current findings offer great promise for such studies to yield novel insights into the pathophysiology of both the proposed “focal” and “complex” AF subphenotypes.

# Chapter 7: Conclusions

From a series of analyses in a large, population-based cohort and a meta-analysis with the previous literature a number of important conclusions can be drawn:

- AF can be identified from nation-wide health registers with high case validity for use in epidemiological studies.
- Potentially modifiable cardiovascular risk factors, particularly hypertension and obesity, confer substantial population attributable risk. Improved control of such risk factors could potentially reduce the population burden of AF or delay its onset.
- Age and sex predict onset of AF in the general population with reasonable accuracy.
- Although conventional risk factors and blood MR-proANP significantly improve predictive accuracy beyond age and sex the improvement is modest.
- Although genetic polymorphisms on chromosomes 4q25 and 16q22 are robustly associated with AF across cross-sectional studies, case-control studies and prospective cohort studies, there is substantial heterogeneity of risk estimates across studies. Genetic polymorphisms do not significantly improve predictive accuracy.
- The missense variant (K897T) in KCNH2 could not be replicated in MDCS and in a meta-analysis of published studies, suggestive that the effect on general AF, if any, is smaller than in the initial report.
- Genetic factors but not conventional risk factors confer risk of AF development in HF patients. Further, a SNP on chromosome 16q22 shows a multiplicative interaction with HF on AF risk, suggestive of a distinct pathophysiology in the context of HF which remains to be clarified.



# Summary in Swedish (Sammanfattning på svenska)

På senare år har sjukdomspanoramata för hjärtsjukdomar förändrats, med sjunkande insjuknande och dödlighet i hjärtinfarkt men stigande insjuknande i förmaksflimmer och hjärtsvikt. Förmaksflimmer är den vanligaste hjärtrytmrubbningen och medför ökad risk för uppkomst av blodproppar i hjärtat som kan resultera i stroke, demens och förtida död. Symptom på förmaksflimmer är hjärtklappning, trötthet, andfåddhet och bröstsmärta, men förmaksflimmer kan också vara asymptomatiskt och upptäckas av en slump vid EKG-undersökning eller om patienten insjuknar i stroke. Diagnosen ställs med hjälp av EKG-undersökning. Förmaksflimmer är också starkt kopplat till hjärtsvikt, sannolikt på grund av den ökade belastning på förmaken som orsakas av den nedsatta pumpförmågan vid hjärtsvikt vilken får till konsekvens att blodet stockar sig bakåt till förmaken. Vid svår hjärtsvikt drabbas upp till 50% av samtidigt förmaksflimmer.

Blodförtunnande läkemedel såsom waran är kraftfulla redskap för att förebygga insjuknande i stroke hos patienter med förmaksflimmer, och minskar sannolikt också risken för demens och förtida död. Det är dock önskvärt att förhindra insjuknande i förmaksflimmer innan det brutit ut, bland annat då stroke ofta kan vara det första symptomet på förmaksflimmer. För att kunna förhindra förmaksflimmer och dess konsekvenser krävs dels möjligheten att identifiera individer i riskzonen och dels effektiva förebyggande åtgärder. Som ett första steg utvärderades därför i detta arbete ett antal möjliga redskap för att identifiera individer med ökad risk för att insjukna i förmaksflimmer, inklusive kända risk faktorer för hjärtsjukdom som visat god förmåga att förutsäga insjuknande i hjärtinfarkt, mätning av ämnen i blodet och genetiska varianter som nyligen kopplats till förmaksflimmer med hjälp av moderna genetiska metoder.

Invånare i Malmö som valts ut slumpmässigt och deltog i studien 'Malmö Kost och Cancer' (totalt 30 447 stycken) lämnade blodprover, genetiskt prov (DNA), besvarade ett formulär och genomgick mätning av längd, vikt och blodtryck mellan 1990 och 1996. Ett antal blodprover togs, som speglar olika patofysiologiska processer med möjlig koppling till förmaksflimmer. Genetiska varianter på kromosom 4, 7 och 16 som kopplats till förmaksflimmer analyserades. Med hjälp av nationella register (Patientregistret och Dödsorsaksregistret) kunde vi sedan studera vilka individer som



insjuknade i förmaksflimmer. Totalt insjuknade 2050 individer i förmaksflimmer under en uppföljningstid på upp till 17 år. Genom att granska journaler och EKG-undersökningar som gjordes i samband med insjuknande fann vi att diagnoser av förmaksflimmer i nationella register hade hög tillförlitlighet.

Vi fann att välkända risk faktorer för hjärtsjukdom, särskilt högt blodtryck och övervikt, svarade för en betydande andel av insjuknanden i förmaksflimmer. Undersökning av dessa risk faktorer medförde att förmaksflimmer kunde förutsägas med någorlunda god precision. Mätning av ett protein som frisätts från förmaket (ANP) vid ökad belastning på förmaksväggen visade sig ytterligare förbättra precisionen.

Genetiska varianter på kromosom 4 och 16 var visserligen kopplade till insjuknande i förmaksflimmer både i vår studie och i andra studier, men i en stor kombinerad analys av våra fynd och tidigare studier (innefattande totalt upp till 150 000 individer) fann vi stora skillnader i vilken risk dessa medför. I vissa studier medförde de genetiska varianterna hög risk för förmaksflimmer men i vår studie av den allmänna befolkningen i Malmö medförde de relativt små risker och ingen förbättrad precision för att förutsäga förmaksflimmer. Sannolikt speglar detta att förmaksflimmer egentligen inte är en enskild sjukdomsentitet utan ett mönster av symptom och EKG-fynd som kan orsakas av olika sjukdomsprocesser, och att olika studier inkluderat olika patientgrupper.

Slutligen studerade vi risk faktorer för insjuknande i förmaksflimmer hos patienter med hjärtsvikt. Dessa utgör en särskild grupp, där det inte är känt varför hälften av patienter med svår hjärtsvikt inte utvecklar förmaksflimmer. Vi fann att vanliga risk faktorer för hjärtsjukdom, inklusive högt blodtryck, övervikt, rökning och diabetes, inte var kopplade till risk för förmaksflimmer hos patienter med hjärtsvikt. Däremot var genetiska varianter starkt kopplade till förmaksflimmer i denna grupp, särskilt en genetisk variant på kromosom 16. Sannolikt speglar det protein som kodas av genen som påverkas av denna genetiska variant en viktig roll för utveckling av förmaksflimmer hos hjärtsviktpatienter, vilket kan ge nya kunskaper om mekanismer hos denna grupp och kanske leda till nya behandlingsstrategier.

Sammanfattningsvis har vi funnit att vanliga risk faktorer för hjärtsjukdom kan förutsäga insjuknande i förmaksflimmer med relativt god precision, att ANP men inte genetiska varianter medför förbättrad precision och att förmaksflimmer vid hjärtsvikt utgör en särskild grupp där en genetisk variant på kromosom 16, men inte vanliga risk faktorer, verkar vara av betydelse.

# Acknowledgements

At times our own light goes out and is rekindled by a spark from another person. Each of us has cause to think with deep gratitude of those who have lighted the flame within us

Albert Schweitzer (1875-1965)

This thesis would not have been possible without the contribution and support of many individuals to whom my gratitude cannot be expressed enough:

My three brilliant mentors: Olle Melander, for his enthusiastic and cheerful mentorship yet sharp mind, and for always being supportive. Chris Newton-Cheh, who took me into his group when I knew nothing except a bunch of wild ideas and managed to turn me into a decent genetic epidemiologist. Always a smile and an entertaining story, yet with an incredibly fast analytical mind and immensely literate within genetic epidemiology. Pyotr Platonov, for unselfishly sharing his extensive knowledge of atrial fibrillation and for always having an open door for me.

David Erlinge, head of the Department of Cardiology, Arne Lindgren and Bosse Norrving at the Department of Neurology, who taught me the joys and rigors of science when I first became interested in research as a medical student, and for support and great camaraderie throughout the years. Pelle Alm, former professor of Pathology, for sparking my interest in research.

Bo Hedblad and Gunnar Engström, for enjoyable collaboration and providing a stimulating working environment in the Cardiovascular Epidemiology community. Tommy Wang for enthusiastic and insightful input on the blood biomarker study.

Marketa Sjögren, Peter Almgren, Anders Dahlin and Håkan Andersson in Malmö for lots of competent help and cheerfulness along the way.

My brother-in-arms, Sasha Koul, who guided and still guides me through the world of clinical cardiology, and all my great friends at the department of Cardiology for

creating a stimulating environment and introducing me into other exciting fields of scientific cardiology (including how to handle both pigs, mice and monocytes...): Jesper van der Pals, Oscar Braun, Mathias Götberg, Henrik Wagner, Göran Olivecrona, Göran Arstad, Jonas Carlson, Fredrik Holmqvist, Stefan Jovinge and Patrik Andersson.

Anders Roijer and Björn Ekmehag who gave me a chance. The staff on the former ward 19 (current ward 18/19) for friendliness and patience with an inexperienced physician. A very special thanks to Lena Lindén and Monica Magnusson, without whom nothing would have been possible.

All my friends who have supported me along the way and made life be about more than cardiovascular epidemiology. No one mentioned and no one forgotten.

All participants of the Malmö Diet and Cancer study, without whose contributions this thesis would not have been possible.

My grand-parents Greta and Erik Olsson, for providing such great examples and showing me the way to medicine.

The family Hans, Vivianne, Kristina, Björn and Martin Olsson for always being supportive and interested in my progress.

A very special thanks to Gun Persson.

Finally, nothing would have been possible without my family – Eva, Jan and Jan Erik Smith – for everlasting support, for providing me with an intellectually stimulating environment in our home and for inspiring me to do great things.

# References

1. Snow J. Snow on Cholera, Reprint of two papers. London: Oxford Press, 1936.
2. Dawber TR, Meadors GF, Moore FE, Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health* 1951;41:279-81.
3. Doll R, Hill AB. The mortality of doctors in relation to their smoking habits; a preliminary report. *Br Med J* 1954;1:1451-5.
4. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J, 3rd. Factors of risk in the development of coronary heart disease--six year follow-up experience. The Framingham Study. *Ann Intern Med* 1961;55:33-50.
5. Kannel WB, LeBauer EJ, Dawber TR, McNamara PM. Relation of body weight to development of coronary heart disease. The Framingham study. *Circulation* 1967;35:734-44.
6. Keys A. Seven countries: a multivariate analysis of death and coronary heart disease. Cambridge, MA, USA and London, England: Harvard University Press, 1980.
7. Gaziano TA, Gaziano JM. Chapter 1: Global Burden of Cardiovascular Disease. In: Bonow RO, Mann DL, Zipes DP, Libby P, eds. *Braunwald's Heart Disease*. 9th edition ed. Philadelphia, PA, USA: Saunders, 2011.
8. Bjorck L, Rosengren A, Bennett K, Lappas G, Capewell S. Modelling the decreasing coronary heart disease mortality in Sweden between 1986 and 2002. *Eur Heart J* 2009;30:1046-56.
9. Wijeyesundera HC, Machado M, Farahati F, et al. Association of temporal trends in risk factors and treatment uptake with coronary heart disease mortality, 1994-2005. *JAMA* 2010;303:1841-7.
10. Braunwald E. Shattuck lecture--cardiovascular medicine at the turn of the millennium: triumphs, concerns, and opportunities. *N Engl J Med* 1997;337:1360-9.
11. Yeh RW, Sidney S, Chandra M, Sorel M, Selby JV, Go AS. Population trends in the incidence and outcomes of acute myocardial infarction. *N Engl J Med* 2010;362:2155-65.
12. Camm AJ, Kirchhof P, Lip GY, et al. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). *Eur Heart J* 2010;31:2369-429.
13. Hering HE. Das Elektrokardiogramm des Pulsus Irregularis perpetuus. *Deutsches Archiv für Physiologie* 1908;94:205-8.
14. Einthoven W. Le telecardiogramme. *Arch Int de Physiol* 1906;4:132-64.

15. Go AS, Hylek EM, Phillips KA, et al. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 2001;285:2370-5.
16. Stewart S, Hart CL, Hole DJ, McMurray JJ. A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/Paisley study. *Am J Med* 2002;113:359-64.
17. Heeringa J, van der Kuip DA, Hofman A, et al. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study. *Eur Heart J* 2006;27:949-53.
18. Naccarelli GV, Varker H, Lin J, Schulman KL. Increasing prevalence of atrial fibrillation and flutter in the United States. *Am J Cardiol* 2009;104:1534-9.
19. Miyasaka Y, Barnes ME, Gersh BJ, et al. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. *Circulation* 2006;114:119-25.
20. Stefansdottir H, Aspelund T, Gudnason V, Arnar DO. Trends in the incidence and prevalence of atrial fibrillation in Iceland and future projections. *Europace* 2011;13:1110-7.
21. Lloyd-Jones DM, Wang TJ, Leip EP, et al. Lifetime risk for development of atrial fibrillation: the Framingham Heart Study. *Circulation* 2004;110:1042-6.
22. Healey JS, Connolly SJ, Gold MR, et al. Subclinical atrial fibrillation and the risk of stroke. *N Engl J Med* 2012;366:120-9.
23. Magnani JW, Rienstra M, Lin H, et al. Atrial fibrillation: current knowledge and future directions in epidemiology and genomics. *Circulation* 2011;124:1982-93.
24. Wolf PA, Dawber TR, Thomas HE, Jr., Kannel WB. Epidemiologic assessment of chronic atrial fibrillation and risk of stroke: the Framingham study. *Neurology* 1978;28:973-7.
25. Krahn AD, Manfreda J, Tate RB, Mathewson FA, Cuddy TE. The natural history of atrial fibrillation: incidence, risk factors, and prognosis in the Manitoba Follow-Up Study. *Am J Med* 1995;98:476-84.
26. Feinberg WM, Seeger JF, Carmody RF, Anderson DC, Hart RG, Pearce LA. Epidemiologic features of asymptomatic cerebral infarction in patients with nonvalvular atrial fibrillation. *Arch Intern Med* 1990;150:2340-4.
27. Ezekowitz MD, James KE, Nazarian SM, et al. Silent cerebral infarction in patients with nonrheumatic atrial fibrillation. The Veterans Affairs Stroke Prevention in Nonrheumatic Atrial Fibrillation Investigators. *Circulation* 1995;92:2178-82.
28. Das RR, Seshadri S, Beiser AS, et al. Prevalence and correlates of silent cerebral infarcts in the Framingham offspring study. *Stroke* 2008;39:2929-35.
29. Ott A, Breteler MM, de Bruyne MC, van Harskamp F, Grobbee DE, Hofman A. Atrial fibrillation and dementia in a population-based study. The Rotterdam Study. *Stroke* 1997;28:316-21.
30. Miyasaka Y, Barnes ME, Petersen RC, et al. Risk of dementia in stroke-free patients diagnosed with atrial fibrillation: data from a community-based cohort. *Eur Heart J* 2007;28:1962-7.

31. Frost L, Engholm G, Johnsen S, Moller H, Henneberg EW, Husted S. Incident thromboembolism in the aorta and the renal, mesenteric, pelvic, and extremity arteries after discharge from the hospital with a diagnosis of atrial fibrillation. *Arch Intern Med* 2001;161:272-6.
32. Frost L, Andersen LV, Johnsen SP, Mortensen LS. Lost life years attributable to stroke among patients with nonvalvular atrial fibrillation: a nationwide population-based follow-up study. *Neuroepidemiology* 2007;29:59-65.
33. Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. *Circulation* 1998;98:946-52.
34. Conen D, Chae CU, Glynn RJ, et al. Risk of death and cardiovascular events in initially healthy women with new-onset atrial fibrillation. *JAMA* 2011;305:2080-7.
35. Lin HJ, Wolf PA, Kelly-Hayes M, et al. Stroke severity in atrial fibrillation. The Framingham Study. *Stroke* 1996;27:1760-4.
36. Andersen LV, Vestergaard P, Deichgraeber P, Lindholt JS, Mortensen LS, Frost L. Warfarin for the prevention of systemic embolism in patients with non-valvular atrial fibrillation: a meta-analysis. *Heart* 2008;94:1607-13.
37. Hart RG, Benavente O, McBride R, Pearce LA. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. *Ann Intern Med* 1999;131:492-501.
38. van Walraven C, Hart RG, Singer DE, et al. Oral anticoagulants vs aspirin in nonvalvular atrial fibrillation: an individual patient meta-analysis. *JAMA* 2002;288:2441-8.
39. Go AS, Hylek EM, Chang Y, et al. Anticoagulation therapy for stroke prevention in atrial fibrillation: how well do randomized trials translate into clinical practice? *JAMA* 2003;290:2685-92.
40. Hylek EM, Go AS, Chang Y, et al. Effect of intensity of oral anticoagulation on stroke severity and mortality in atrial fibrillation. *N Engl J Med* 2003;349:1019-26.
41. Flaker GC, Pogue J, Yusuf S, et al. Cognitive function and anticoagulation control in patients with atrial fibrillation. *Circ Cardiovasc Qual Outcomes* 2010;3:277-83.
42. Wang TJ, Larson MG, Levy D, et al. Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham Heart Study. *Circulation* 2003;107:2920-5.
43. Maisel WH, Stevenson LW. Atrial fibrillation in heart failure: epidemiology, pathophysiology, and rationale for therapy. *Am J Cardiol* 2003;91:2D-8D.
44. Mamas MA, Caldwell JC, Chacko S, Garratt CJ, Fath-Ordoubadi F, Neyses L. A meta-analysis of the prognostic significance of atrial fibrillation in chronic heart failure. *Eur J Heart Fail* 2009;11:676-83.
45. Platonov PG. Substrate for development of atrial fibrillation in patients with congestive heart failure: are we close to the answer? *Heart Rhythm* 2009;6:452-3.
46. Cha YM, Redfield MM, Shen WK, Gersh BJ. Atrial fibrillation and ventricular dysfunction: a vicious electromechanical cycle. *Circulation* 2004;109:2839-43.

47. Caldwell JC, Contractor H, Petkar S, et al. Atrial fibrillation is under-recognized in chronic heart failure: insights from a heart failure cohort treated with cardiac resynchronization therapy. *Europace* 2009;11:1295-300.
48. Ehrlich JR, Nattel S, Hohnloser SH. Atrial fibrillation and congestive heart failure: specific considerations at the intersection of two common and important cardiac disease sets. *J Cardiovasc Electrophysiol* 2002;13:399-405.
49. Tsang TS, Gersh BJ, Appleton CP, et al. Left ventricular diastolic dysfunction as a predictor of the first diagnosed nonvalvular atrial fibrillation in 840 elderly men and women. *J Am Coll Cardiol* 2002;40:1636-44.
50. Graham I, Atar D, Borch-Johnsen K, et al. European guidelines on cardiovascular disease prevention in clinical practice: executive summary: Fourth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (Constituted by representatives of nine societies and by invited experts). *Eur Heart J* 2007;28:2375-414.
51. Hamm CW, Bassand JP, Agewall S, et al. ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: The Task Force for the management of acute coronary syndromes (ACS) in patients presenting without persistent ST-segment elevation of the European Society of Cardiology (ESC). *Eur Heart J* 2011;32:2999-3054.
52. Van de Werf F, Bax J, Betriu A, et al. Management of acute myocardial infarction in patients presenting with persistent ST-segment elevation: the Task Force on the Management of ST-Segment Elevation Acute Myocardial Infarction of the European Society of Cardiology. *Eur Heart J* 2008;29:2909-45.
53. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA* 2001;285:2486-97.
54. Kannel WB. Some lessons in cardiovascular epidemiology from Framingham. *Am J Cardiol* 1976;37:269-82.
55. Anderson KM, Wilson PW, Odell PM, Kannel WB. An updated coronary risk profile. A statement for health professionals. *Circulation* 1991;83:356-62.
56. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97:1837-47.
57. D'Agostino RB, Wolf PA, Belanger AJ, Kannel WB. Stroke risk profile: adjustment for antihypertensive medication. The Framingham Study. *Stroke* 1994;25:40-3.
58. Schnabel RB, Sullivan LM, Levy D, et al. Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study. *Lancet* 2009;373:739-45.
59. Schnabel RB, Aspelund T, Li G, et al. Validation of an atrial fibrillation risk algorithm in whites and African Americans. *Arch Intern Med* 2010;170:1909-17.
60. Kannel WB, D'Agostino RB, Silbershatz H, Belanger AJ, Wilson PW, Levy D. Profile for estimating risk of heart failure. *Archives of Internal Medicine* 1999;159:1197-1204.

61. Conroy RM, Pyorala K, Fitzgerald AP, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003;24:987-1003.
62. Jong P, Yusuf S, Rousseau MF, Ahn SA, Bangdiwala SI. Effect of enalapril on 12-year survival and life expectancy in patients with left ventricular systolic dysfunction: a follow-up study. *Lancet* 2003;361:1843-8.
63. Zhang Y, Zhang P, Mu Y, et al. The role of renin-angiotensin system blockade therapy in the prevention of atrial fibrillation: a meta-analysis of randomized controlled trials. *Clin Pharmacol Ther* 2010;88:521-31.
64. Fauchier L, Pierre B, de Labriolle A, Grimard C, Zannad N, Babuty D. Antiarrhythmic effect of statin therapy and atrial fibrillation a meta-analysis of randomized controlled trials. *J Am Coll Cardiol* 2008;51:828-35.
65. Rahimi K, Emberson J, McGale P, et al. Effect of statins on atrial fibrillation: collaborative meta-analysis of published and unpublished evidence from randomised controlled trials. *BMJ* 2011;342:d1250.
66. Chamberlain AM, Agarwal SK, Folsom AR, et al. A clinical risk score for atrial fibrillation in a biracial prospective cohort (from the Atherosclerosis Risk in Communities [ARIC] study). *Am J Cardiol* 2011;107:85-91.
67. Levy S, Sutton G, Ng PC, et al. The diploid genome sequence of an individual human. *PLoS Biol* 2007;5:e254.
68. Wheeler DA, Srinivasan M, Egholm M, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 2008;452:872-6.
69. Smith JG, Newton-Cheh C. Candidate gene and genome-wide association studies. In: Roden DM, ed. *AHA Clinical Series: Cardiovascular Genetics and Genomics*. Chichester: Wiley-Blackwell, 2009.
70. Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet.* 2001;17:502-510.
71. Friedman MJ, Trager W. The biochemistry of resistance to malaria. *Sci Am* 1981;244:154-5, 158-64.
72. Goldstein JL, Brown MS. Familial hypercholesterolemia: identification of a defect in the regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity associated with overproduction of cholesterol. *Proc.Natl.Acad.Sci.U.S.A* 1973;70:2804-2808.
73. Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* 2008;40:161-9.
74. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat.Genet.* 2003;33:177-182.
75. Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype-phenotype associations. *Nature* 2007;447:655-60.
76. Smith JG, Newton-Cheh C. Genome-wide association study in humans. *Methods Mol Biol* 2009;573:231-58.
77. Wolff L. Familiar auricular fibrillation. *N Engl J Med* 1943;229:396-398.
78. Levy RL. Paroxysmal auricular fibrillation and flutter without signs of organic cardiac disease in two brothers. *J Mt Sinai Hosp* 1942;8:765-770.



79. Fox CS, Parise H, D'Agostino RB, Sr., et al. Parental atrial fibrillation as a risk factor for atrial fibrillation in offspring. *JAMA* 2004;291:2851-5.
80. Arnar DO, Thorvaldsson S, Manolio TA, et al. Familial aggregation of atrial fibrillation in Iceland. *Eur Heart J* 2006;27:708-12.
81. Ellinor PT, Yoerger DM, Ruskin JN, MacRae CA. Familial aggregation in lone atrial fibrillation. *Hum Genet* 2005;118:179-84.
82. Lubitz SA, Yin X, Fontes JD, et al. Association between familial atrial fibrillation and risk of new-onset atrial fibrillation. *JAMA* 2010;304:2263-9.
83. Christophersen IE, Ravn LS, Budtz-Joergensen E, et al. Familial aggregation of atrial fibrillation: a study in Danish twins. *Circ Arrhythm Electrophysiol* 2009;2:378-83.
84. Marcus GM, Smith LM, Vittinghoff E, et al. A first-degree family history in lone atrial fibrillation patients. *Heart Rhythm* 2008;5:826-30.
85. Ellinor PT, Moore RK, Patton KK, Ruskin JN, Pollak MR, Macrae CA. Mutations in the long QT gene, *KCNQ1*, are an uncommon cause of atrial fibrillation. *Heart* 2004;90:1487-8.
86. Ellinor PT, Petrov-Kondratov VI, Zakharova E, Nam EG, MacRae CA. Potassium channel gene mutations rarely cause atrial fibrillation. *BMC Med Genet* 2006;7:70.
87. Ackerman MJ, Priori SG, Willems S, et al. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies: this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). *Europace* 2011;13:1077-109.
88. Olson TM, Alekseev AE, Moreau C, et al. *KATP* channel mutation confers risk for vein of Marshall adrenergic atrial fibrillation. *Nat Clin Pract Cardiovasc Med* 2007;4:110-6.
89. Olson TM, Alekseev AE, Liu XK, et al. *Kv1.5* channelopathy due to *KCNA5* loss-of-function mutation causes human atrial fibrillation. *Hum Mol Genet* 2006;15:2185-91.
90. Chen YH, Xu SJ, Bendahhou S, et al. *KCNQ1* gain-of-function mutation in familial atrial fibrillation. *Science* 2003;299:251-254.
91. Otway R, Vandenberg JI, Guo G, et al. Stretch-sensitive *KCNQ1* mutation A link between genetic and environmental factors in the pathogenesis of atrial fibrillation? *J Am Coll Cardiol* 2007;49:578-86.
92. Das S, Makino S, Melman YF, et al. Mutation in the S3 segment of *KCNQ1* results in familial lone atrial fibrillation. *Heart Rhythm* 2009;6:1146-53.
93. Yang Y, Xia M, Jin Q, et al. Identification of a *KCNE2* gain-of-function mutation in patients with familial atrial fibrillation. *Am J Hum Genet* 2004;75:899-905.
94. Ravn LS, Aizawa Y, Pollevick GD, et al. Gain of function in *IKs* secondary to a mutation in *KCNE5* associated with atrial fibrillation. *Heart Rhythm* 2008;5:427-35.
95. Xia M, Jin Q, Bendahhou S, et al. A *Kir2.1* gain-of-function mutation underlies familial atrial fibrillation. *Biochem Biophys Res Commun* 2005;332:1012-9.
96. Hodgson-Zingman DM, Karst ML, Zingman LV, et al. Atrial natriuretic peptide frameshift mutation in familial atrial fibrillation. *N Engl J Med* 2008;359:158-65.
97. Oberti C, Wang L, Li L, et al. Genome-wide linkage scan identifies a novel genetic locus on chromosome 5p13 for neonatal atrial fibrillation associated with sudden death and variable cardiomyopathy. *Circulation* 2004;110:3753-9.

98. Darbar D, Kannankeril PJ, Donahue BS, et al. Cardiac sodium channel (SCN5A) variants associated with atrial fibrillation. *Circulation* 2008;117:1927-35.
99. Ellinor PT, Nam EG, Shea MA, Milan DJ, Ruskin JN, MacRae CA. Cardiac sodium channel mutation in atrial fibrillation. *Heart Rhythm* 2008;5:99-105.
100. Benito B, Brugada R, Perich RM, et al. A mutation in the sodium channel is responsible for the association of long QT syndrome and familial atrial fibrillation. *Heart Rhythm* 2008;5:1434-40.
101. Makiyama T, Akao M, Shizuta S, et al. A novel SCN5A gain-of-function mutation M1875T associated with familial atrial fibrillation. *J Am Coll Cardiol* 2008;52:1326-34.
102. McNair WP, Ku L, Taylor MR, et al. SCN5A mutation associated with dilated cardiomyopathy, conduction disorder, and arrhythmia. *Circulation* 2004;110:2163-7.
103. Olson TM, Michels VV, Ballew JD, et al. Sodium channel mutations and susceptibility to heart failure and atrial fibrillation. *JAMA* 2005;293:447-54.
104. Li Q, Huang H, Liu G, et al. Gain-of-function mutation of Nav1.5 in atrial fibrillation enhances cellular excitability and lowers the threshold for action potential firing. *Biochem Biophys Res Commun* 2009;380:132-7.
105. Watanabe H, Darbar D, Kaiser DW, et al. Mutations in sodium channel beta1- and beta2-subunits associated with atrial fibrillation. *Circ Arrhythm Electrophysiol* 2009;2:268-75.
106. Mahida S, Lubitz SA, Rienstra M, Milan DJ, Ellinor PT. Monogenic atrial fibrillation as pathophysiological paradigms. *Cardiovasc Res* 2011;89:692-700.
107. Ng PC, Murray SS, Levy S, Venter JC. An agenda for personalized medicine. *Nature* 2009;461:724-6.
108. Sinner MF, Lubitz SA, Pfeufer A, et al. Lack of replication in polymorphisms reported to be associated with atrial fibrillation. *Heart Rhythm* 2011;8:403-9.
109. Sinner MF, Pfeufer A, Akyol M, et al. The non-synonymous coding IKr-channel variant KCNH2-K897T is associated with atrial fibrillation: results from a systematic candidate gene-based analysis of KCNH2 (HERG). *Eur Heart J* 2008;29:907-14.
110. Schnabel RB, Kerr KF, Lubitz SA, et al. Large-Scale Candidate Gene Analysis in Whites and African-Americans Identifies IL6R Polymorphism in Relation to Atrial Fibrillation: The NHLBI CARE Project. *Circ Cardiovasc Genet* 2011.
111. Wirka RC, Gore S, Van Wagoner DR, et al. A common connexin-40 gene promoter variant affects connexin-40 expression in human atria and is associated with atrial fibrillation. *Circ Arrhythm Electrophysiol* 2011;4:87-93.
112. Gudbjartsson DF, Arnar DO, Helgadottir A, et al. Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature* 2007;448:353-7.
113. Gudbjartsson DF, Holm H, Gretarsdottir S, et al. A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat Genet* 2009;41:876-8.
114. Benjamin EJ, Rice KM, Arking DE, et al. Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. *Nat Genet* 2009;41:879-81.
115. Ellinor PT, Lunetta KL, Glazer NL, et al. Common variants in KCNN3 are associated with lone atrial fibrillation. *Nat Genet* 2010;42:240-4.

116. Gretarsdottir S, Thorleifsson G, Manolescu A, et al. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann Neurol* 2008;64:402-9.
117. Pfeufer A, van Noord C, Marciante KD, et al. Genome-wide association study of PR interval. *Nat Genet* 2010;42:153-9.
118. Berglund G, Elmstahl S, Janzon L, Larsson SA. The Malmo Diet and Cancer Study. Design and feasibility. *J Intern Med*. 1993;233:45-51.
119. Statistical yearbook of Sweden. Stockholm, Sweden: Statistics Sweden, 2012.
120. Manjer J, Carlsson S, Elmstahl S, et al. The Malmo Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. *Eur J Cancer Prev* 2001;10:489-99.
121. Melander O, Newton-Cheh C, Almgren P, et al. Novel and conventional biomarkers for prediction of incident cardiovascular events in the community. *JAMA* 2009;302:49-57.
122. Enhorning S, Wang TJ, Nilsson PM, et al. Plasma copeptin and the risk of diabetes mellitus. *Circulation* 2010;121:2102-8.
123. Calltorp J, Adami HO, Astrom H, et al. Country profile: Sweden. *Lancet* 1996;347:587-94.
124. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekblom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol* 2009;24:659-67.
125. Ludvigsson JF, Andersson E, Ekblom A, et al. External review and validation of the Swedish national inpatient register. *BMC Public Health* 2011;11:450.
126. Johansson LA, Westerling R. Comparing Swedish hospital discharge records with death certificates: implications for mortality statistics. *Int J Epidemiol* 2000;29:495-502.
127. Ingelsson E, Arnlov J, Sundstrom J, Lind L. The validity of a diagnosis of heart failure in a hospital discharge register. *Eur J Heart Fail* 2005;7:787-91.
128. Nilsson AC, Spetz CL, Carsjo K, Nightingale R, Smedby B. [Reliability of the hospital registry. The diagnostic data are better than their reputation]. *Lakartidningen* 1994;91:598, 603-5.
129. Linnarsjo A, Hammar N, Gustavsson A, Reuterwall C. Recent time trends in acute myocardial infarction in Stockholm, Sweden. *Int J Cardiol* 2000;76:17-21.
130. Elo SL, Karlberg IH. Validity and utilization of epidemiological data: a study of ischaemic heart disease and coronary risk factors in a local population. *Public Health* 2009;123:52-7.
131. 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;16:235-43.
132. Ahlbom A. Acute myocardial infarction in Stockholm--a medical information system as an epidemiological tool. *Int J Epidemiol* 1978;7:271-6.
133. Hammar N, Nerbrand C, Ahlmark G, et al. Identification of cases of myocardial infarction: hospital discharge data and mortality data compared to myocardial infarction community registers. *Int J Epidemiol* 1991;20:114-20.

134. Forberg JL, Green M, Bjork J, et al. In search of the best method to predict acute coronary syndrome using only the electrocardiogram from the emergency department. *J Electrocardiol* 2009;42:58-63.
135. Waldo AL, Feld GK. Inter-relationships of atrial fibrillation and atrial flutter mechanisms and clinical implications. *J Am Coll Cardiol* 2008;51:779-86.
136. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York: Wiley, 2000.
137. Cox DR. Regression models and life tables. *J R Stat Soc [B]* 1972;34:187-220.
138. Gronnesby JK, Borgan O. A method for checking regression models in survival analysis based on the risk score. *Lifetime Data Anal* 1996;2:315-28.
139. Hosmer DW, Lemeshow S. Goodness of Fit Tests for the Multiple Logistic Regression Model. *Comm Statist* 1980;A9:1043-1069.
140. Levin ML. The occurrence of lung cancer in man. *Acta Unio Int Contra Cancrum* 1953;9:531-41.
141. Walter SD. The estimation and interpretation of attributable risk in health research. *Biometrics* 1976;32:829-49.
142. Harrell FE, Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA* 1982;247:2543-6.
143. Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Ann Intern Med* 2009;150:795-802.
144. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157-72; discussion 207-12.
145. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88.
146. Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954;10:101-29.
147. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539-58.
148. Kaab S, Darbar D, van Noord C, et al. Large scale replication and meta-analysis of variants on chromosome 4q25 associated with atrial fibrillation. *Eur Heart J* 2009;30:813-9.
149. Viviani Anselmi C, Novelli V, Roncarati R, et al. Association of rs2200733 at 4q25 with atrial flutter/fibrillation diseases in an Italian population. *Heart* 2008;94:1394-6.
150. Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. *Nature* 2010;467:52-8.
151. Kirchhof P, Lip GY, Van Gelder IC, et al. Comprehensive risk reduction in patients with atrial fibrillation: emerging diagnostic and therapeutic options--a report from the 3rd Atrial Fibrillation Competence NETwork/European Heart Rhythm Association consensus conference. *Europace* 2012;14:8-27.

152. Benjamin EJ, Levy D, Vaziri SM, D'Agostino RB, Belanger AJ, Wolf PA. Independent risk factors for atrial fibrillation in a population-based cohort. The Framingham Heart Study. *JAMA* 1994;271:840-4.
153. Aviles RJ, Martin DO, Apperson-Hansen C, et al. Inflammation as a risk factor for atrial fibrillation. *Circulation* 2003;108:3006-10.
154. Wang TJ, Larson MG, Levy D, et al. Plasma natriuretic peptide levels and the risk of cardiovascular events and death. *N Engl J Med* 2004;350:655-63.
155. Frost L, Vestergaard P. Alcohol and risk of atrial fibrillation or flutter: a cohort study. *Arch Intern Med* 2004;164:1993-8.
156. Psaty BM, Manolio TA, Kuller LH, et al. Incidence of and risk factors for atrial fibrillation in older adults. *Circulation* 1997;96:2455-61.
157. Mozaffarian D, Psaty BM, Rimm EB, et al. Fish intake and risk of incident atrial fibrillation. *Circulation* 2004;110:368-73.
158. Stewart S, Hart CL, Hole DJ, McMurray JJ. Population prevalence, incidence, and predictors of atrial fibrillation in the Renfrew/Paisley study. *Heart* 2001;86:516-21.
159. Feinberg WM, Blackshear JL, Laupacis A, Kronmal R, Hart RG. Prevalence, age distribution, and gender of patients with atrial fibrillation. Analysis and implications. *Arch Intern Med* 1995;155:469-73.
160. Engstrom G, Hedblad B, Juul-Moller S, Tyden P, Janzon L. Cardiac arrhythmias and stroke: increased risk in men with high frequency of atrial ectopic beats. *Stroke* 2000;31:2925-9.
161. Schoonderwoerd BA, Smit MD, Pen L, Van Gelder IC. New risk factors for atrial fibrillation: causes of 'not-so-lone atrial fibrillation'. *Europace* 2008;10:668-73.
162. Huxley RR, Lopez FL, Folsom AR, et al. Absolute and attributable risks of atrial fibrillation in relation to optimal and borderline risk factors: the Atherosclerosis Risk in Communities (ARIC) study. *Circulation* 2011;123:1501-8.
163. Platonov PG, Mitrofanova LB, Orshanskaya V, Ho SY. Structural abnormalities in atrial walls are associated with presence and persistency of atrial fibrillation but not with age. *J Am Coll Cardiol* 2011;58:2225-32.
164. Yasue H, Yoshimura M, Sumida H, et al. Localization and mechanism of secretion of B-type natriuretic peptide in comparison with those of A-type natriuretic peptide in normal subjects and patients with heart failure. *Circulation* 1994;90:195-203.
165. Braunwald E. Biomarkers in heart failure. *N Engl J Med* 2008;358:2148-59.
166. Marott SC, Nordestgaard BG, Zacho J, et al. Does elevated C-reactive protein increase atrial fibrillation risk? A Mendelian randomization of 47,000 individuals from the general population. *J Am Coll Cardiol* 2010;56:789-95.
167. Schnabel RB, Larson MG, Yamamoto JF, et al. Relations of biomarkers of distinct pathophysiological pathways and atrial fibrillation incidence in the community. *Circulation* 2010;121:200-7.
168. Adamsson Eryd S, Smith JG, Melander O, Hedblad B, Engstrom G. Inflammation-sensitive proteins and risk of atrial fibrillation: a population-based cohort study. *Eur J Epidemiol* 2011;26:449-55.

169. Schnabel RB, Larson MG, Yamamoto JF, et al. Relation of multiple inflammatory biomarkers to incident atrial fibrillation. *Am J Cardiol* 2009;104:92-6.
170. Conen D, Ridker PM, Everett BM, et al. A multimarker approach to assess the influence of inflammation on the incidence of atrial fibrillation in women. *Eur Heart J* 2010;31:1730-6.
171. Lasky-Su J, Lyon HN, Emilsson V, et al. On the replication of genetic associations: timing can be everything! *Am J Hum Genet* 2008;82:849-58.
172. Lubitz SA, Sinner MF, Lunetta KL, et al. Independent susceptibility markers for atrial fibrillation on chromosome 4q25. *Circulation* 2010;122:976-84.
173. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol* 2004;159:882-90.
174. Shi L, Li C, Wang C, et al. Assessment of association of rs2200733 on chromosome 4q25 with atrial fibrillation and ischemic stroke in a Chinese Han population. *Hum Genet* 2009.
175. Lee KT, Yeh HY, Tung CP, et al. Association of RS2200733 but not RS10033464 on 4q25 with atrial fibrillation based on the recessive model in a taiwanese population. *Cardiology* 2010;116:151-6.
176. Li C, Wang F, Yang Y, et al. Significant association of SNP rs2106261 in the ZFHX3 gene with atrial fibrillation in a Chinese Han GeneID population. *Hum Genet* 2011;129:239-46.
177. Delaney JT, Jeff JM, Brown NJ, et al. Characterization of genome-wide association-identified variants for atrial fibrillation in african americans. *PLoS ONE* 2012;7:e32338.
178. Bhatia RS, Tu JV, Lee DS, et al. Outcome of heart failure with preserved ejection fraction in a population-based study. *N Engl J Med* 2006;355:260-9.
179. Cowie MR, Wood DA, Coats AJ, et al. Incidence and aetiology of heart failure; a population-based study. *Eur Heart J* 1999;20:421-8.
180. Yusuf S, Hawken S, Ounpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 2004;364:937-52.
181. Heeringa J, Hoogendoorn EH, van der Deure WM, et al. High-normal thyroid function and risk of atrial fibrillation: the Rotterdam study. *Arch Intern Med* 2008;168:2219-24.
182. Gammage MD, Parle JV, Holder RL, et al. Association between serum free thyroxine concentration and atrial fibrillation. *Arch Intern Med* 2007;167:928-34.
183. Cappola AR, Fried LP, Arnold AM, et al. Thyroid status, cardiovascular risk, and mortality in older adults. *JAMA* 2006;295:1033-41.
184. Sawin CT, Geller A, Wolf PA, et al. Low serum thyrotropin concentrations as a risk factor for atrial fibrillation in older persons. *N Engl J Med* 1994;331:1249-52.
185. Buch P, Friberg J, Scharling H, Lange P, Prescott E. Reduced lung function and risk of atrial fibrillation in the Copenhagen City Heart Study. *Eur Respir J* 2003;21:1012-6.
186. Gami AS, Hodge DO, Herges RM, et al. Obstructive sleep apnea, obesity, and the risk of incident atrial fibrillation. *J Am Coll Cardiol* 2007;49:565-71.

187. Alonso A, Lopez FL, Matsushita K, et al. Chronic kidney disease is associated with the incidence of atrial fibrillation: the Atherosclerosis Risk in Communities (ARIC) study. *Circulation* 2011;123:2946-53.
188. Fedorowski A, Hedblad B, Engstrom G, Gustav Smith J, Melander O. Orthostatic hypotension and long-term incidence of atrial fibrillation: the Malmo Preventive Project. *J Intern Med* 2010;268:383-9.
189. Emilsson L, Smith JG, West J, Melander O, Ludvigsson JF. Increased risk of atrial fibrillation in patients with coeliac disease: a nationwide cohort study. *Eur Heart J* 2011.
190. Ahlehoff O, Gislason GH, Jorgensen CH, et al. Psoriasis and risk of atrial fibrillation and ischaemic stroke: a Danish Nationwide Cohort Study. *Eur Heart J* 2011.
191. Walkey AJ, Wiener RS, Ghobrial JM, Curtis LH, Benjamin EJ. Incident stroke and mortality associated with new-onset atrial fibrillation in patients hospitalized with severe sepsis. *JAMA* 2011;306:2248-54.
192. Fernando HC, Jaklitsch MT, Walsh GL, et al. The Society of Thoracic Surgeons practice guideline on the prophylaxis and management of atrial fibrillation associated with general thoracic surgery: executive summary. *Ann Thorac Surg* 2011;92:1144-52.
193. Mathew JP, Parks R, Savino JS, et al. Atrial fibrillation following coronary artery bypass graft surgery: predictors, outcomes, and resource utilization. MultiCenter Study of Perioperative Ischemia Research Group. *JAMA* 1996;276:300-6.
194. Rostagno C, La Meir M, Gelsomino S, et al. Atrial fibrillation after cardiac surgery: incidence, risk factors, and economic burden. *J Cardiothorac Vasc Anesth* 2010;24:952-8.
195. Vaporciyan AA, Correa AM, Rice DC, et al. Risk factors associated with atrial fibrillation after noncardiac thoracic surgery: analysis of 2588 patients. *J Thorac Cardiovasc Surg* 2004;127:779-86.
196. Thornton JR. Atrial fibrillation in healthy non-alcoholic people after an alcoholic binge. *Lancet* 1984;2:1013-5.
197. Turagam MK, Velagapudi P, Kocheril AG. Atrial fibrillation in athletes. *Am J Cardiol* 2012;109:296-302.
198. van der Hooft CS, Heeringa J, van Herpen G, Kors JA, Kingma JH, Stricker BH. Drug-induced atrial fibrillation. *J Am Coll Cardiol* 2004;44:2117-24.
199. Ripatti S, Tikkanen E, Orho-Melander M, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet* 2010;376:1393-400.
200. Kathiresan S, Melander O, Anevski D, et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. *N Engl J Med* 2008;358:1240-9.
201. Palomaki GE, Melillo S, Bradley LA. Association between 9p21 genomic markers and heart disease: a meta-analysis. *JAMA* 2010;303:648-56.
202. Newton-Cheh C, Smith JG. What can human genetics teach us about the causes of cardiovascular disease? *J Am Coll Cardiol* 2010;55:2843-5.
203. Mommersteeg MT, Brown NA, Prall OW, et al. Pitx2c and Nkx2-5 are required for the formation and identity of the pulmonary myocardium. *Circ Res* 2007;101:902-9.

204. Haissaguerre M, Jais P, Shah DC, et al. Spontaneous initiation of atrial fibrillation by ectopic beats originating in the pulmonary veins. *N Engl J Med* 1998;339:659-66.
205. Wang J, Klysik E, Sood S, Johnson RL, Wehrens XH, Martin JF. Pitx2 prevents susceptibility to atrial arrhythmias by inhibiting left-sided pacemaker specification. *Proc Natl Acad Sci U S A* 2010;107:9753-8.
206. Kirchhof P, Kahr PC, Kaese S, et al. PITX2c is expressed in the adult left atrium, and reducing Pitx2c expression promotes atrial fibrillation inducibility and complex changes in gene expression. *Circ Cardiovasc Genet* 2011;4:123-33.
207. Nojiri S, Joh T, Miura Y, et al. ATBF1 enhances the suppression of STAT3 signaling by interaction with PIAS3. *Biochem Biophys Res Commun* 2004;314:97-103.
208. Hilfiker-Kleiner D, Hilfiker A, Fuchs M, et al. Signal transducer and activator of transcription 3 is required for myocardial capillary growth, control of interstitial matrix deposition, and heart protection from ischemic injury. *Circ Res* 2004;95:187-95.
209. Tsai CT, Lin JL, Lai LP, Lin CS, Huang SK. Membrane translocation of small GTPase Rac1 and activation of STAT1 and STAT3 in pacing-induced sustained atrial fibrillation. *Heart Rhythm* 2008;5:1285-93.
210. Burashnikov A. Are there atrial selective/predominant targets for "upstream" atrial fibrillation therapy? *Heart Rhythm* 2008;5:1294-5.
211. Lubitz SA, Benjamin EJ, Ruskin JN, Fuster V, Ellinor PT. Challenges in the classification of atrial fibrillation. *Nat Rev Cardiol* 2010;7:451-60.
212. Husser D, Adams V, Piorkowski C, Hindricks G, Bollmann A. Chromosome 4q25 variants and atrial fibrillation recurrence after catheter ablation. *J Am Coll Cardiol* 2010;55:747-53.