Helsinki University Biomedical Dissertations No. 188

GENETIC RISK PROFILES FOR CORONARY HEART DISEASE

Emmi Tikkanen

Hjelt Institute, Faculty of Medicine
University of Helsinki, Helsinki, Finland
and
Institute for Molecular Medicine Finland (FIMM)
University of Helsinki, Helsinki, Finland

ACADEMIC DISSERTATION

To be presented with the permission of the Faculty of Medicine, University of Helsinki, for public examination in Lecture Hall 1, Kytösuontie 9, on November 8th, 2013, at 12 noon.

Helsinki 2013

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

ISSN 1457-8433
ISBN 978-952-10-9288-6 (pbk.)
ISBN 978-952-10-9289-3 (PDF)
http://ethesis.helsinki.fi

Unigrafia Oy
Helsinki, Finland 2013

Supervisor	Professor Samuli Ripatti Hjelt Institute, Faculty of Medicine University of Helsinki Helsinki, Finland Institute for Molecular Medicine Finland (FIMM) University of Helsinki Helsinki, Finland Wellcome Trust Sanger Institute Cambridge, United Kingdom
Reviewers	Professor Esa Läärä Department of Mathematical Sciences, Faculty of Science University of Oulu Oulu, Finland Professor Hannes Lohi Department of Veterinary Biosciences, Faculty of Veterinary Medicine Research Programs Unit, Molecular Neurology, Faculty of Medicine University of Helsinki and Folkhälsan Research Center Helsinki, Finland
Opponent	Dr. Daniel MacArthur Analytic and Translational Genetics Unit Massachusetts General Hospital Boston, MA, USA Program in Medical and Population Genetics Broad Institute of Harvard and MIT Cambridge, MA, USA Department of Genetics Harvard Medical School Boston, MA, USA

To my family

ABSTRACT

Coronary heart disease (CHD) is a major burden for public health worldwide. Several factors are known to be associated with the disease risk, including high levels of low-density lipoprotein (LDL) cholesterol and blood pressure. The established risk factors do not, however, fully predict an individual's risk for the disease. In recent years, new candidate risk factors, including genetic markers, have been extensively studied. Genome-wide association studies (GWASs) have mapped over 40 genetic regions for CHD risk and hundreds of loci for CHD risk factors. The impact of these findings on public health remains obscure.

In this study, we utilized the findings from large-scale GWASs and constructed genetic risk scores (GRSs) based on panels of single-nucleotide polymorphisms (SNPs). The aim was to estimate the joint effects of common genetic markers on CHD and its risk factors, and to evaluate the incremental value of genetic information in CHD risk assessment.

In Projects I and II, we studied longitudinal effects of genetic loci associated with lipids and blood pressure, and evaluated the prediction of dyslipidemia and hypertension in young adults by using the genetic information in addition to clinical measurements. Our results show that the GRSs were significantly associated with longitudinal measurements of lipid traits and blood pressure throughout childhood, adolescence and adulthood. For some traits, the genetic effect was not consistent across age groups. For example, the GRS effect for high-density lipoprotein (HDL) cholesterol was considerably larger in children than in adults, and the proportion of variance explained by the SNPs in children was twice as much as in adults. The GRS for triglycerides improved the prediction of dyslipidemia in young adults when added to childhood lipid measurement. The blood pressure GRS increased the risk of hypertension, but did not improve risk discrimination over other risk factors.

In Projects III and IV, we found that the GRSs based on CHD SNPs predicted CHD events. The estimated relative risk for the GRS was similar in magnitude to the relative risk of other risk factors such as systolic blood pressure. The genetic effect was independent of family history of the disease, which has been used as a surrogate for genetic risk in many prediction algorithms. The GRS based on 28 SNPs improved the prediction of CHD events beyond traditional risk factors and family history when evaluated with reclassification or discrimination metrics. Genetic screening could be especially useful for individuals in the intermediate-risk group (10-year risk 10-20%), as current preventive strategies are focused mainly on the high-risk group (>20%).

In conclusion, these findings suggest that the genetic information obtained from GWASs could be used in early identification of individuals at increased risk for lipid disorders, hypertension and CHD.

Keywords: cardiovascular disease, genetic association, genetic epidemiology, risk factor

TIIVISTELMÄ

Sydänsairaudet muodostavat yhden suurimmista kansanterveydellisistä ongelmista maailmanlaajuisesti. Niiden syntyyn vaikuttavat monet riskitekijät, kuten korkea kolesteroli ja verenpaine, mutta huomattavaa osaa sairastumisriskistä ei voida selittää tunnetuilla riskitekijöillä. Tämä on motivoinut tutkijoita etsimään uusia, mukaan lukien geneettisiä, riskitekijöitä. Genominlaajuisissa assosiaatioanalyyseissa on löydetty yli 40 sepelvaltimotautiin yhteydessä olevaa geenialuetta. Useita geenialueita on myös yhdistetty sydän- ja verisuonitautien riskitekijöihin. Näiden löydösten merkitystä kansanterveydelle ei ole kuitenkaan vielä ymmärretty.

Tässä väitöstutkimuksessa selvitettiin tunnettujen geenimerkkien yhteisvaikutusta sepelvaltimotautiin ja sen riskitekijöihin geneettisten riskiprofiilien avulla. Tavoitteena oli myös arvioida, voiko genomitiedon avulla parantaa sepelvaltimotaudin ennustamista.

Tutkimuksen ensimmäisessä ja toisessa osatyössä selvitettiin geneettisten riskiprofiilien pitkittäisvaikutuksia veren kolesterolipitoisuuksiin, triglyserideihin ja verenpaineeseen, sekä ennustettiin rasva-aineenvaihdunnan häiriöitä ja hypertensiota nuorilla aikuisilla genomitiedon avulla. Tutkimus osoittaa, että geneettiset riskiprofiilit ovat yhteydessä näihin riskitekijöihin varhais-lapsuudesta keski-ikään. Joidenkin ominaisuuksien suhteen geeniprofiilien vaikutus ei ollut yhtenäinen läpi elämänkaaren, esimerkiksi estimoitu efekti HDL-kolesteroliin oli suurempi lapsilla kuin aikuisilla. Geneettiset profiilit nostivat riskiä sairastua hypertriglyseridemiaan ja hypertensioon. Genomitiedon lisääminen malleihin paransi erottelukykyä dyslipidemian, mutta ei hypertension ennustamisessa.

Tutkimuksen kolmannen ja neljännen osatyön mukaan tunnetut geenimerkit ennustavat sepelvaltimotautitapahtumia. Geneettisen riskiprofiilin suhteellisen riskin suuruus oli verrattavissa muiden riskitekijöiden, kuten korkean verenpaineen, aiheuttamaan riskiin. Tutkimus osoitti myös, että geeniprofiilin vaikutus oli riippumaton riskitekijä verrattuna potilaan perhehistoriaan sydän- ja verisuonitaudeista, jota on yleisesti käytetty geneettisen riskin indikaattorina. Tutkimuksemme mukaan genomitiedon lisääminen sepelvaltimotaudin ennustemalleihin perinteisten riskitekijöiden lisäksi paransi mallin erottelukykyä ja tarkensi riskiluokittelua. Geneettisestä testauksesta olisi erityisesti hyötyä niille henkilöille, jotka sijoittuvat sepelvaltimotaudin riskiluokittelussa keskivaiheille (10-vuoden riski 10-20%), sillä ennaltaehkäisevät toimenpiteet on kohdistettu vain korkean riskin ryhmään (yli 20%).

Tulokset viittaavat siihen, että genominlaajuisista assosiaatioanalyyseista saatu genomitieto voisi antaa aikaisessa vaiheessa elämää hyödyllistä lisätietoa henkilön riskistä sairastua rasva-aineenvaihdunnan häiriöihin, hypertensioon ja sepelvaltimotautiin.

Avainsanat: sydän- ja verisuonitaudit, geneettinen assosiaatio, geneettinen epidemiologia, riskitekijä

TABLE OF CONTENTS

ABBREVIATIONS

A	Adenine
ACS	Acute coronary syndrome
ATP-III	Adult Treatment Panel III
BMI	Body mass index
bp	Base pair
C	Cytosine
CEU	Utah residents with Northern and Western European ancestry (HapMap population)
CHD	Coronary heart disease
CI	Confidence interval
CRP	C-reactive protein
CVD	Cardiovascular disease
DBP	Diastolic blood pressure
DNA	Deoxyribonucleic acid
FR	Finrisk study
G	Guanine
GRS	Genetic risk score
GWAS	Genome-wide association study
HDL	High-density lipoprotein
HR	Hazard ratio
HWE	Hardy–Weinberg equilibrium
ICD	International Classification of Disease

IDI	Integrated discrimination index
IL-6	Interleukin-6
IQR	Interquantile range
LD	Linkage disequilibrium
LDL	Low-density lipoprotein
MAF	Minor allele frequency
MDC-CC	Malmö Diet and Cancer Study - Cardiovascular Cohort
MDCS	Malmö Diet and Cancer Study
MI	Myocardial infarction
MPP	Malmö Preventive Project
NICE	National Institute for Health and Clinical Excellence
NRI	Net reclassification improvement
OR	Odds ratio
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
SBP	Systolic blood pressure
SNP	Single-nucleotide polymorphism
T	Thymine
TC	Total cholesterol
TG	Triglycerides
TSI	Toscani in Italia (HapMap population)
YFS	The Cardiovascular Risk in Young Finns Study
YRI	Yoruba in Ibadan, Nigeria (HapMap population)

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals (I-IV). In addition, some unpublished data are presented.

I Tikkanen E*, Tuovinen T*, Widén E, Lehtimäki T, Viikari J, Kähönen M, Peltonen L, Raitakari OT, Ripatti S (2011): Association of Known Loci with Lipid Levels Among Children and Prediction of Dyslipidemia in Adults. *Circ Cardiovasc Genet*. **4**(6):673–680.

II Oikonen M*, Tikkanen E*, Juhola J*, Tuovinen T, Seppälä I, Juonala M, Taittonen L, Mikkilä V, Kähönen M, Ripatti S, Viikari J, Lehtimäki T, Havulinna AS, Kee F, Newton-Cheh C, Peltonen L, Schork NJ, Murray SS, Berenson GS, Chen W, Srinivasan SR, Salomaa V, Raitakari OT (2011): Genetic variants and blood pressure in a population-based cohort: the cardiovascular risk in young Finns study. *Hypertension*. **58**(6):1079-1085.

III Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, Guiducci C, Perola M, Jula A, Sinisalo J, Lokki M-L, Nieminen MS, Melander O, Salomaa V, Peltonen L, Kathiresan S (2010): A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet*. **376**(9750):1393–1400.

IV Tikkanen E, Havulinna AS, Palotie A, Salomaa V, Ripatti S (2013): Genetic Risk Prediction and a 2-Stage Risk Screening Strategy for Coronary Heart Disease. *Arterioscler Thromb Vasc Biol*. **33**(9):2261–2266.

* These authors contributed equally to this work.

1 INTRODUCTION

Diseases of the heart and vascular system are leading causes of mortality and morbidity worldwide. Coronary heart disease (CHD) is caused by an accumulation of lipids in the inner wall of the heart-supplying artery and often manifests as chest pain or discomfort. CHD is a complex disease, with both environmental and genetic risk factors modifying disease susceptibility.

The Framingham Heart Study, initiated in 1948, was established to identify risk factors for CHD, which were largely unknown at that time. During 1960–1970 the study uncovered several factors that modify the risk for CHD, nowadays referred to as traditional or conventional risk factors. These include high levels of low-density lipoprotein (LDL) cholesterol, low levels of high-density lipoprotein (HDL) cholesterol, high blood pressure, smoking, and diabetes. Based on these risk factors, the group has developed a prediction algorithm that allows clinicians to estimate the CHD risk for subjects without established cardiovascular disease (Wilson et al. 1998; Framingham Heart Study 2013). Several other similar risk estimation tools exist (Assmann et al. 2002; Conroy et al. 2003; Hippisley-Cox et al. 2007; Ridker et al. 2007; D'Agostino et al. 2008; Ridker et al. 2008).

However, a high number of events occur in individuals who do not score high based on the predictions with current risk estimation algorithms. Even though genetic factors are known to contribute to the risk of disease, information on genetic factors is neglected in most risk estimation tools. Some algorithms (Assmann et al. 2002; Ridker et al. 2007; Ridker et al. 2008) use information on family history as a surrogate for genetic effects. Although relatively easy to measure, family history might not accurately capture the genetic effects since families usually share environmental conditions as well. Instead, measured genetic variants provide more precise information on an individual’s genetic susceptibility to CHD.

Genome-wide association studies (GWASs) have identified several common genetic susceptibility loci for CHD and its risk factors (Levy et al. 2009; Newton-Cheh et al. 2009; Teslovich et al. 2010; Coronary Artery Disease Genetics Consortium 2011; Ehret et al. 2011; Schunkert et al. 2011; Deloukas et al. 2012). Individually, these variants typically have modest effect sizes (OR 1.06–1.30 for CHD) and are not useful in risk estimation studies, but jointly they might have sufficient predictive power. A simple way to evaluate the joint effects of multiple independent genetic variants is to generate genetic risk scores (GRSs) by summing the number of risk alleles at each individual locus. These scores provide a tool for studying the effects of common genetic variation on trait variability and disease risk.

Although the GRS based on the current GWAS findings only capture approximately 10% of the CHD heritability (Deloukas et al. 2012), there could be clear advantages of using

genetic variants in estimation of CHD risk. Unlike traditional risk factors, the genetic information is determined at conception and could be available from the first health examination. Since CHD evolves over a long period and atherosclerotic processes are initiated already in childhood (Enos et al. 1953; Holman et al. 1958; McNamara et al. 1971), detecting high-risk individuals early in life is important. Nevertheless, since the GWAS findings are based on cross-sectional adult samples, it is not clear whether these genetic effects are consistent throughout the lifespan.

In this study, we utilized the results from large-scale GWASs for CHD, lipids, and blood pressure, and generated GRSs for each trait based on the reported single-nucleotide polymorphisms. We estimated longitudinal genetic effects for lipids and blood pressure and evaluated the genetic risk for dyslipidemia, hypertension, CHD, and other end-points in over 50,000 Finnish and Swedish study subjects.

2 REVIEW OF THE LITERATURE

2.1 Coronary heart disease

2.1.1 Diagnosis and symptoms

Cardiovascular disease (CVD) encompasses a range of diseases of the heart and blood vessels, including CHD, cerebrovascular disease, and other diseases of the cardiovascular system. CHD covers several acute and chronic medical conditions (**Table 1**), and acute coronary syndrome (ACS) includes only acute events such as myocardial infarction (MI), unstable angina or sudden cardiac death.

Table 1. Selected ICD-10 codes for coronary heart disease.

I20	**Angina pectoris** Unstable angina; angina pectoris with documented spasm; other forms of angina pectoris (e.g. angina of effort); unspecified angina pectoris
I21	**Acute myocardial infraction** Acute transmural myocardial infarction of anterior wall / inferior wall / other sites / unspecified site; acute subendocardial myocardial infarction; unspecified myocardial infarction
I22	**Subsequent myocardial infraction** Subsequent myocardial infarction of anterior wall / inferior wall / other sites / unspecified site
I23	**Certain current complications following acute myocardial infarction** Hemopericardium; atrial septal defect; ventricular septal defect; rupture of cardiac wall without hemopericardium; rupture of chordae tendineae; rupture of papillary muscle; thrombosis of atrium, auricular appendage, and ventricle; other current complications following acute myocardial infarction
I24	**Other acute ischemic heart diseases** Coronary thrombosis not resulting in myocardial infarction; Dressler syndrome; other forms of acute ischemic heart disease; unspecified acute ischemic heart disease
I25	**Chronic ischemic heart disease** Atherosclerotic heart disease; old myocardial infarction; aneurysm of heart; coronary artery aneurysm; ischemic cardiomyopathy; silent myocardial ischemia; other forms of chronic ischemic heart disease; unspecified chronic ischemic heart disease
I46	**Cardiac arrest** Cardiac arrest with successful resuscitation; sudden cardiac death; unspecified cardiac arrest

Abbreviations: ICD-10, International Classification of Diseases (version 10). Data adapted from World Health Organization (2012b).

A typical symptom in CHD is chest pain or discomfort radiating to the arm, neck, jaw, or epigastrium. Other symptoms include dizziness, fatigue, nausea and diaphoresis. Sudden cardiac death might also be the first manifestation of the disease. Diagnosis of CHD is based on the patient's cardiac symptoms, changes in electrocardiogram, biomarker measurements, and in fatal cases, autopsy findings (Luepker et al. 2003).

2.1.2 Burden of the disease

CVD, especially CHD, is the most common cause of death worldwide. In Europe, 22% of women and 20% of men die from CHD. Overall, CVD accounts for 52% of deaths in women and 42% of deaths in men (Nichols et al. 2012). In Finland, the prevalence for CHD (ICD-10 codes I20-I25) in 2010 was 800 per 100,000 in males and 254 per 100,000 in females (age-standardized for 25–74-year-olds to standard European population) (Laatikainen et al. 2009).

Regional variation in age-standardized CHD mortality rates exists between and within countries in Europe. The highest mortality rates are in Central and Eastern Europe, with rates decreasing gradually when moving towards southwestern Europe. A clear within-country variation exists in Germany, the UK, Poland, and Finland. For example, people living in the northeastern part of Finland have a higher mortality from MI than individuals in the southwestern part of the country (annual incidence rate during 1991–2003 per 100,000: 855.6 vs. 334.7 in men and 351.4 vs. 210.6 in women in northeast and southwest Finland, respectively). The regional differences in mortality rates can partly be explained by prevalence of the classic risk factors (e.g. smoking and hypertension) (Havulinna et al. 2008; Muller-Nordhorn et al. 2008).

Mortality rates for CHD have declined in many countries during the past decades. The decreasing trend has been especially strong in Finland, which had the highest CHD mortality in the world in the 1970s. Today, the mortality rate is closer to that of other Nordic countries, but remains high (World Health Organization 2012a). Reductions in risk factors (mainly LDL cholesterol levels) and improved treatment are the main causes for reduced CHD mortality rates in Finland (Vartiainen et al. 1994; Laatikainen et al. 2005).

Despite progress in many developed countries, the global burden of the disease persists. CHD rates have increased and are higher than ever in many former Soviet Union states (Mirzaei et al. 2009). The CHD epidemic is dynamic, reflecting changes in CHD risk factor levels within a population. Poor dietary habits, heavy smoking and a sedentary lifestyle due to urbanization have triggered the epidemic in developing countries, where the majority of all cardiovascular events occur (Okrainec et al. 2004; Perk et al. 2012). The emerging epidemic of obesity and diabetes poses a new threat to cardiovascular health worldwide. Since 2000, the prevalence of diabetes has increased in most European countries (Nichols et al. 2012).

The economic burden of CHD consists of direct costs in health care (33% of total cost), productivity loss (29%), and informal care (38%). The annual total cost of CHD is estimated to be over 60 billion Euros in the European Union (Nichols et al. 2012).

2.1.3 Disease pathology

Atherosclerosis is an underlying cause of CHD, but the etiology of the disease is complex. Both environmental and genetic factors act at several stages of atherosclerosis. The development of atherosclerosis (**Figure 1**) begins early in life as an accumulation of lipids beneath the endothelial layer in the artery. These fatty streaks are common in youth and do not necessarily cause atherosclerosis. The critical stage in the progression of the disease is when fatty streaks evolve into atherosclerotic plaque (atheroma). Several CHD risk factors play a role in this process and may trigger the disease pathogenesis. Large atheromas congest arteries and diminish blood flow to the heart. MI occurs as a result of thrombus caused by rupture or erosion of the fibrous plaque (Lusis 2000; Moore and Tabas 2011).

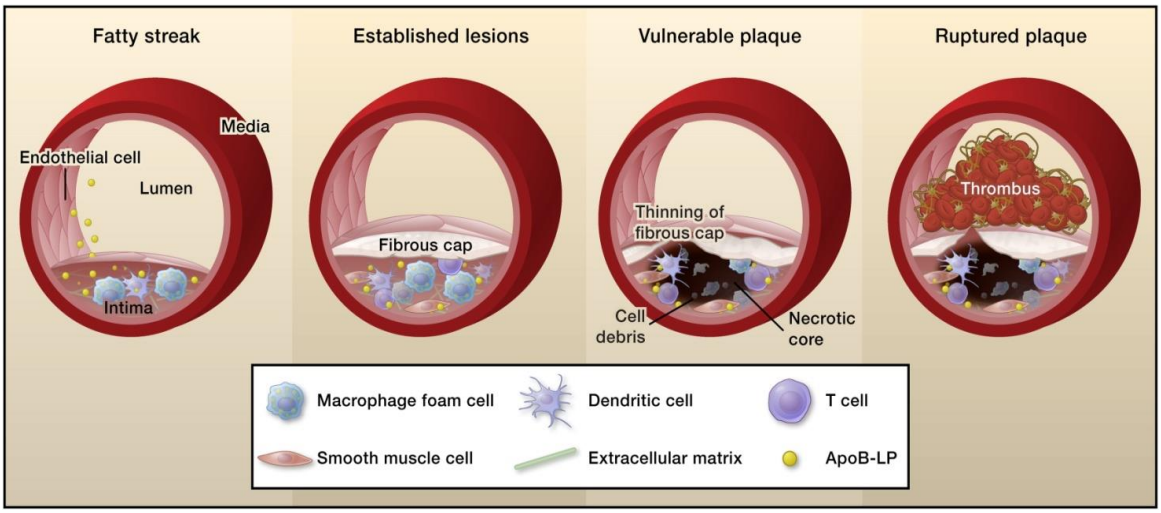


Figure 1. Development of atherosclerosis and plaque rupture in the artery. Reprinted from Moore and Tabas (2011) with permission from Elsevier.

The risk for a thrombus depends on the vulnerability of the plaque. Atheromas with thin fibrous caps are more likely to develop thrombus than plaques with a thick endothelial layer. Thus, the development of thrombus with severe clinical complications depends more on plaque vulnerability than severity of stenosis. Plaque vulnerability is largely affected by inflammation process, which is initiated by the accumulation of lipids and hemodynamic strain in the artery (Lusis 2000; Hansson 2005).

2.1.4 Risk factors

Advanced age and male gender are two major risk factors for CHD. The prevalence of CHD increases with age, affecting less than 1% of the population aged under 40 years, 6% of 40- to 59-year-olds, and over 10% of 60+ females and over 20% of 60+ males (Roger et al. 2012). The lifetime risk at the age of 40 for developing CHD is one in two for men and one in three for women (Lloyd-Jones et al. 1999). However, the difference in risk between genders decreases with age; in fact, heart disease is a more common cause of death in elderly women than in men (Nichols et al. 2012).

Other established risk factors for CHD, identified originally mainly by the Framingham Heart Study (Kannel et al. 1961; Kannel et al. 1967; Abbott et al. 1988), include high levels of LDL and low levels of HDL cholesterol, high blood pressure, obesity, diabetes mellitus, smoking and family history of the disease. In recent years, the role of inflammation has been under intensive research and novel biomarkers reflecting inflammatory status have been tested for their predictive properties. Other factors that have been found to increase or decrease the risk for CHD include physical fitness and activity (Williams 2001) and psychosocial factors such as stress and anxiety (Roest et al. 2010; Kivimäki et al. 2012), diet (Dauchet et al. 2006), and alcohol consumption (Corrao et al. 2004), among others.

In Finland, the prevalence of standard risk factors has substantially changed over the last 40 years. Especially men living in the North Karelia had exceptionally high risk factor values in the 1970s: 54% were smokers and their mean levels of total cholesterol and blood pressure were 6.96 mmol/l and 147/90 mmHg, respectively. The North Karelia project was established in 1972 to reduce the risk factor burden in Eastern Finland, and after 1977 the interventions were applied to the rest the of the country (Puska et al. 1976; Puska et al. 1998). Interventions have resulted in changes in health behavior, especially reduced dietary salt and saturated fat intake, and consequently have decreased the risk factor levels. For example, the pooled mean total cholesterol level in 2007 in males was 5.39 mmol/l (5.45 mmol/l in the North Karelia region). Since 1972, CHD mortality has declined by 80% in the middle-aged population and changes in standard risk factor levels predict 60% of the reduced mortality (Vartiainen et al. 2010).

2.1.4.1 Blood lipids

Some blood lipids have a central role in the pathogenesis of atherosclerosis due to their subendothelial accumulation. High levels of LDL cholesterol in the blood increase the accumulation of LDL in the arteries and accelerate the development of atherosclerotic lesions. Therefore, current strategies for preventing CHD focus on lowering blood LDL cholesterol levels. Treatment with hydroxymethol glutryl coenzyme (statins) has been successful in both primary and secondary prevention of CHD (Scandinavian Simvastatin Survival Study Group 1994; Shepherd et al. 1995; Sacks et al. 1996; Heart Protection

Study Collaborative Group 2002; Baigent et al. 2005). Randomized trials have reported 20-30% reductions of cardiovascular events in subjects treated with statins compared with placebo group subjects (Shepherd et al. 1995; Baigent et al. 2005). Furthermore, studies conducted in both CHD patients and healthy subjects with elevated lipid levels have reported similar 5-year number needed to treat (NNT) estimates; approximately 40 to 60 subjects need to be treated with pravastatin or lovastatin to prevent one coronary event (Superko and King 2008; Ridker et al. 2009). However, a considerable residual risk for CHD remains in individuals treated with statins. Despite achieving ideal LDL levels with treatment, plaque progression continues in approximately 20% of patients with established CHD (Bayturan et al. 2010).

HDL cholesterol has an inverse association with CHD. It removes excess cholesterol from arterial walls (a process called reverse cholesterol transport) and has anti-inflammatory and antioxidant properties (Barter et al. 2004). Because of these anti-atherogenic mechanisms, high levels of blood HDL may decrease the risk for CHD. Despite efforts, pharmacological agents that raise HDL levels have not been demonstrated to be an effective treatment for CHD (Libby et al. 2011). One clinical trial was even terminated prematurely due to an increased risk for cardiovascular events and death in those treated with torcetrapib, a drug that inhibits cholesterol ester transfer protein (CETP) and raises HDL (Barter et al. 2007). Recently, the causal role of HDL in development of CHD has been questioned since the genetic variants that lower HDL levels have not been shown to increase CHD risk (Voight et al. 2012). The results indicate that even though HDL cholesterol is a marker of the risk it might not be causally related to the disease. As HDL cholesterol is highly correlated with other metabolic traits, such as triglycerides, obesity, and insulin resistance, it is possible that some other trait is the causal factor underlying the association between HDL and CHD. However, more studies are needed to unravel this mystery.

Increased triglyceride concentrations are associated with higher CHD risk in univariate analysis, but when adjusted for other risk factors, the effect is attenuated (Sarwar et al. 2007; Di Angelantonio et al. 2009). This is because blood triglyceride levels are highly correlated with other CHD risk factors and are therefore not considered to be an independent risk factor. In metabolic syndrome, elevated triglycerides usually coexist with low HDL cholesterol levels, abdominal obesity, high blood pressure, and elevated fasting plasma glucose levels (Alberti et al. 2005).

European guidelines for lipid levels in subjects at low or moderate cardiovascular risk (Perk et al. 2012) recommend that HDL cholesterol should be at least 1.2 mmol/l or more in females and 1.0 mmol/l or more in males. Other conventional blood lipids have the same guidelines for both genders: LDL cholesterol less than 3.0 mmol/l and triglycerides (fasting) less than 1.7 mmol/l. Total cholesterol level is calculated as a function of other lipid measurements and should not exceed 5.0 mmol/l.

2.1.4.2 Blood pressure

High blood pressure is one of the major risk factors for CHD, sudden death, heart failure and stroke. High blood pressure raises the mechanical stress of blood vessels and increases the workload of the heart. Increased hemodynamic burden decreases the elasticity of blood vessels and promotes the development of atheromas. High blood pressure might also induce structural changes in the heart. In response to mechanical stress, the left ventricular wall of the heart is thickened (i.e. left ventricular hypertrophy). This condition often precedes systolic or diastolic dysfunction and heart failure (Lorell and Carabello 2000; Gradman and Alfayoumi 2006).

Hypertension is measured as systolic (maximum) and diastolic (minimum) blood pressure. Systolic indicates the pressure in the arteries when the heart muscle contracts, and diastolic is the pressure between heartbeats, when the heart is filled with blood. Blood pressure 120/80 mmHg is considered ideal. Classification of blood pressure levels and the definition of hypertension are given in **Table 2**.

Environmental risk factors such as diet, physical exercise, and alcohol and salt intake, have an influence on blood pressure levels (Frisoli et al. 2011). Thus, managing these risk factors is an important preventive strategy for CHD. In addition, antihypertensive drugs (such as beta blockers and calcium channel blockers) are prescribed especially for subjects with severe hypertension (systolic blood pressure ≥ 180 and/or diastolic blood pressure ≥ 110 mmHg) or with a high total risk for cardiovascular events (Perk et al. 2012).

Table 2. Definitions and classification of blood pressure levels.*

Category	SBP (mmHg)		DBP (mmHg)
Optimal	< 120	and	< 80
Normal	120–129	and/or	80–84
High normal	130–139	and/or	85–89
Hypertension	≥ 140	and/or	≥ 90

* Blood pressure levels in untreated individuals. Abbreviations: SBP, systolic blood pressure; DBP, diastolic blood pressure. Data adapted from Perk et al. (2012).

2.1.4.3 Obesity

The prevalence of worldwide obesity has almost doubled since 1980. In 2008, 10% of men and 14% of women had a body mass index (BMI) of 30 kg/m^2 or more (World Health Organization 2012a). In adults, BMI of 25 kg/m^2 or more is defined as overweight and BMI of 30 kg/m^2 or more as obesity. In the general population, overweight and obesity are associated with increased overall mortality, which is mainly attributable to

cardiovascular and diabetic deaths (Whitlock et al. 2009). However, for patients with established CHD, the relationship between obesity and cardiovascular mortality is more controversial. Overweight and mild obesity appear to be protective and improve survival of CHD patients (Romero-Corral et al. 2006). This "obesity paradox" could be partly due to the inaccuracy of BMI as an indicator of harmful obesity. Although useful as a standard measure, BMI is a rough estimate of overall obesity, as it provides no information on fat distribution. Two individuals with the same BMI could have very different body composition of fat and lean mass.

Abdominal obesity, measured by waist circumference or waist-to-hip ratio, has been found to increase the risk of diabetes and CHD over and above BMI. Abdominal measurements may also be useful in more accurate risk classification of BMI-defined obese and overweight people (Canoy et al. 2007; The InterAct Consortium 2012).

2.1.4.4 Diabetes mellitus

Diabetes mellitus includes a range of medical conditions. Two main subtypes are type 1 diabetes (juvenile-onset diabetes) and type 2 diabetes (adult-onset diabetes), but other subtypes also exist. Of the two main subtypes, type 2 is the most prevalent form of diabetes and thus, a more relevant risk factor, but both subtypes have been shown to increase CHD risk (Grundy et al. 1999; Orchard et al. 2006). Diabetes is manifested by high levels of glucose in the blood caused by insulin resistance. Insulin is a hormone produced by the pancreas that lowers blood glucose levels. In type 1 diabetes, beta cells that produce insulin have been destroyed by autoimmune reactions, and the patients are dependent on insulin injections to maintain normal glucose levels. The onset of type 2 diabetes is usually later in life (> 40 years of age) and is often preceded by metabolic syndrome (Wilson et al. 2005). Weight control is an essential part of the treatment of patients with type 2 diabetes.

Diagnostic criteria for diabetes are fulfilled if the measured fasting plasma glucose is 7 mmol/l or more or 11.1 mmol/l or more two hours after the oral dose in the glucose tolerance test. As dyslipidemia and hypertension are especially harmful in diabetics, more stringent thresholds are applied for blood lipid levels and blood pressure. For example, the target value for blood pressure is below 130/80 mmHg. Statins might be considered for patients with unfavorable lipid values, as they have been shown to decrease the risk for vascular events in people with diabetes (Kearney et al. 2008; Working group set up by the Finnish Medical Society Duodecim and the Finnish Respiratory Society 2011).

2.1.4.5 Smoking

Smoking increases the risk of several chronic diseases. Smokers have a two-fold risk for developing heart disease compared with nonsmokers, and the relative risk is 25% higher in women than men (Huxley and Woodward 2011). Furthermore, people who have never

smoked but who live with a smoker have a 30% excess risk for heart disease, which is comparable to the risk of subjects who smoke one cigarette per day (Law et al. 1997).

Exposure to cigarette smoke has many adverse effects on health. It promotes inflammation, increases oxidative stress and has unfavorable effects on lipid profile (Ambrose and Barua 2004). Smoking thus acts in concert with other risk factors, and its coexistence with such risk factors as high blood pressure or high cholesterol multiplies the risk for cardiovascular events. For example, smoking reduces the amount of oxygen in the blood, which increases the workload of the heart and raises blood pressure. Consequently, when smoking and hypertension are combined they increase the risk for arterial stiffness more than either factor alone (Scallan et al. 2010).

2.1.4.6 Family history of cardiovascular disease

Positive parental family history of heart disease refers to a situation where at least one parent of an individual has suffered from the disease. This simple measure, usually evaluated with a questionnaire, has been thought to represent a combination of both environmental (shared risk factors in the family) and genetic effects. Many studies have shown that family history is a risk factor for CHD, and the effect is independent of traditional risk factors (Jousilahti et al. 1996; Hawe et al. 2003; Lloyd-Jones et al. 2004). Also, adjusting for behavioral risk factors (tobacco use, alcohol use, physical activity, and fruit and vegetable intake), psychosocial risk factors (depression, permanent stress, financial stress, stressful events, and perceived locus of control) and a panel of common genetic variants explains only modestly the association between MI and family history (Chow et al. 2011). Thus, the composition of family history measurement remains largely unknown.

2.1.4.7 Inflammatory biomarkers

The development of atherosclerosis is a complex interplay between metabolic and inflammatory processes. Focal endothelial activation triggers an immune response, which attracts immune cells to the site. The cells infiltrate the lesion and produce inflammatory cytokines (interferon-γ, interleukin-1, and tumor necrosis factor), inducing the production of interleukin-6 (IL-6) (Hansson 2005). IL-6 stimulates the production of C-reactive protein (CRP), a commonly used indicator of acute infections, but the levels are also elevated in chronic diseases. Fibrinogen is another protein involved in the inflammation process, playing a role in blood clotting and platelet aggregation. These inflammatory biomarkers have also been associated with CVD and CHD in the general population and their use in cardiovascular risk evaluation has been studied (Ridker et al. 2000; Danesh et al. 2004; Melander et al. 2009; Kaptoge et al. 2012).

Nevertheless, there are some drawbacks to using these inflammatory biomarkers in risk assessment. First, they are associated with other classic risk factors, and thus, they might

not provide much additional information beyond these risk factors. For example, inflammatory processes have been shown to have a role in a low HDL cholesterol state (Laurila et al. 2013). Second, inflammation markers are not specific to CVD, as they are elevated in other nonvascular diseases as well. Third, the causality or dose-effect relationship of these markers and the risk of CVD has not been established. Fourth, evidence on therapeutic agents that target circulating levels of these markers and affect CVD incidence is lacking. Fifth, the assays have a higher cost than classic risk factors such as blood lipids (Perk et al. 2012).

2.1.5 Early atherosclerotic changes and risk factor fluctuation in young people

CHD usually manifests in adulthood, but atherosclerotic changes begin early in life. Early atherosclerotic changes in children were documented for the first time in autopsy studies in the 1950s (Holman et al. 1958). Fatty streaks were found in aortas of children as young as 9 months, and every child over 7 years of age had at least some fatty streaks. Advanced atherosclerotic lesions in young persons were found in autopsies of young male casualties from the Korean and Vietnam wars, where atherosclerotic lesions were observed in 50–75% of soldiers (Enos et al. 1953; McNamara et al. 1971). In the Bogalusa Heart Study, the extent of atherosclerosis was evaluated in autopsies of young people (aged 2- to 39 years) and correlated with several cardiovascular risk factors (BMI, systolic and diastolic blood pressure, blood lipids, and smoking). The study not only found strong associations between the individual risk factors and the extent of preclinical atherosclerosis, but it also showed that the amount of lesions in the aorta and coronary arteries increased as the number of risk factors rose (Berenson et al. 1998).

Cardiovascular risk factors in children, adolescents, and young adults have been investigated in many longitudinal studies, where the subjects have been followed from childhood to early adulthood and clinical measurements of risk factors taken several times during their lifespan. The studies include the Bogalusa Heart Study (initiated in 1973), the Muscatine Study (1971), and the Cardiovascular Risk in Young Finns Study (1980), among others (Lauer et al. 1975; Srinivasan et al. 1976; Raitakari et al. 2008). One important finding from these studies is that the risk factors have a tendency to track from childhood to adulthood, and thus, measurements on childhood risk factors are predictive of risk factor levels in adulthood (Webber et al. 1991; Porkka et al. 1994). Therefore, risk factor measurements taken at a young age might provide an important tool for the early assessment of cardiovascular risk. However, due to insufficient follow-up time, linking childhood risk factors to adulthood CVD in these cohorts has not yet been possible.

The development of noninvasive imaging methods for preclinical atherosclerosis has provided a useful collection of intermediate cardiovascular end-points that measure atherosclerotic changes in a healthy population, long before the clinical manifestation of the disease. These methods include measures of coronary calcium, carotid intima-media

thickness, and increased left ventricular mass, among others. Many of the longitudinal studies have incorporated measurements of preclinical atherosclerosis into their follow-up investigations and found that these measures in young adulthood are associated with abnormalities in childhood risk factors (Mahoney et al. 1996; Davis et al. 2001; Li et al. 2003; Raitakari et al. 2003).

2.1.6 Cardiovascular risk assessment

Cardiovascular risk refers to the likelihood of having a cardiovascular event over a defined time period. It can be estimated by using the information on traditional risk factors with existing prediction tools (Wilson et al. 1998; Assmann et al. 2002; Conroy et al. 2003; Hippisley-Cox et al. 2007; Ridker et al. 2007; D'Agostino et al. 2008; Ridker et al. 2008). For example, the most widely used Framingham risk score (Wilson et al. 1998) estimates the absolute 10-year risk for all CHD events, whereas SCORE (Conroy et al. 2003) is primarily aimed at predicting fatal CVD events. The risk estimation is highly dependent on the choice of prediction tool, with different risk calculators providing different risk estimates and categorization for the same individuals (Allan et al. 2013).

Risk assessment is an important tool in primary prevention of cardiovascular events. Based on the risk estimates, subjects can be classified into predefined risk categories (e.g. 10-year risk 0–5%, 5–10%, 10–20%, >20%). These risk categories can then be used by clinicians to guide treatment decisions. For example, current guidelines from ATP-III (2001) and the National Institute for Health and Care Excellence (2008) recommend that statin treatment be allocated to subjects with a 10-year risk 20% or more. As subjects with established CVD have a high risk regardless of their risk score, these prediction algorithms are applicable only for disease-free subjects. In terms of effectiveness, directing preventive strategies to high-risk individuals only is not ideal, as most of the cardiovascular events occur within a population with a risk lower than 20%. In the era of low-cost statins, it has been proposed that the target group for preventative actions be widened (Lloyd-Jones 2010; Holmes et al. 2011). This could be done by lowering the risk threshold that defines those eligible for statins or by improving the risk categorization with a more accurate risk factor panel.

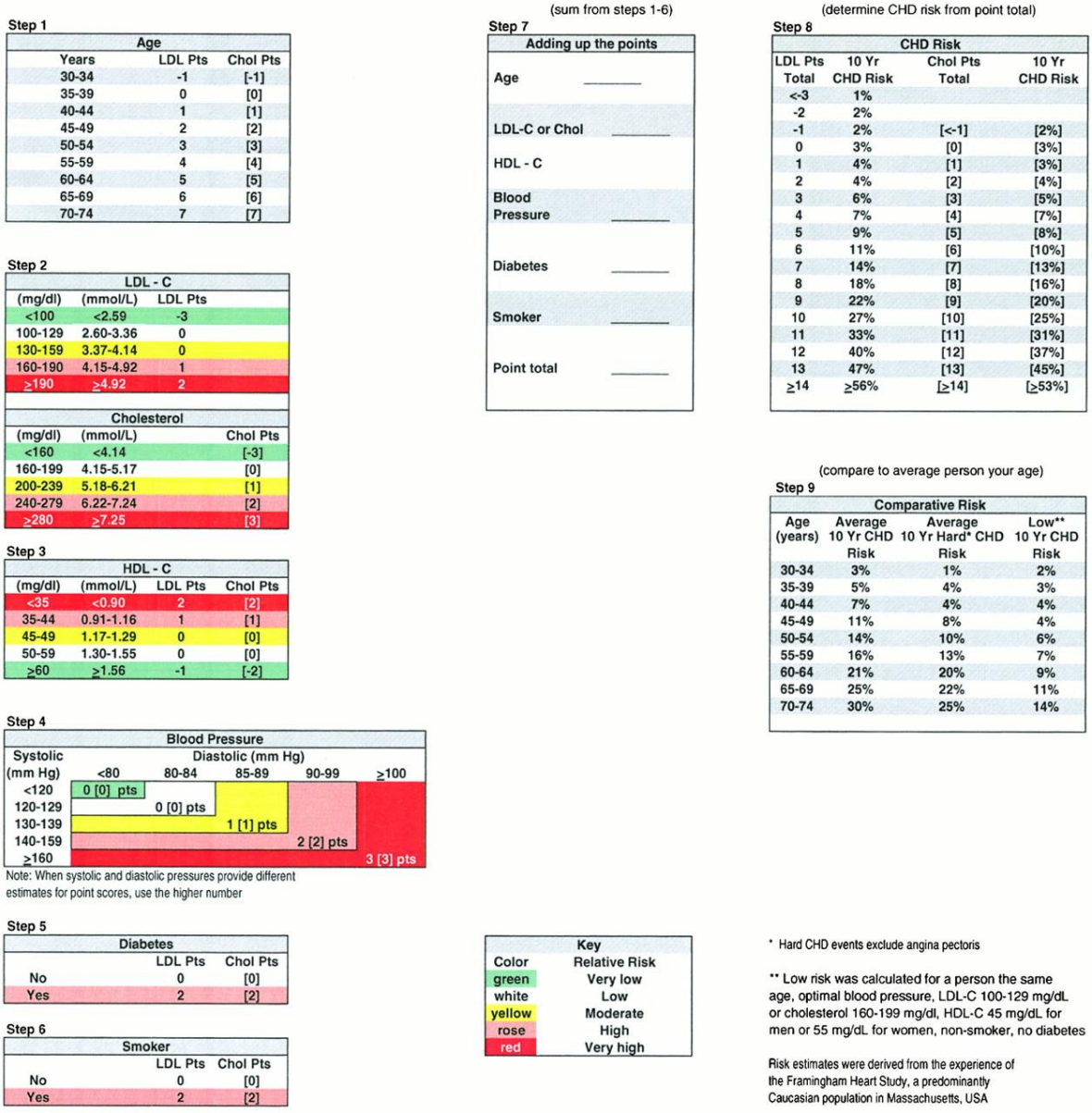
The high-risk strategy might have clear benefits for an individual, but it has only a small impact on the disease burden at a population level. Another approach for disease prevention is a population strategy that aims to shift the whole risk factor distribution via health campaigns or other safe interventions. The population strategy has a strong effect on disease incidence, but it yields little benefit at an individual level. Fortunately, the two approaches are not mutually exclusive (Rose 1981; Rose 2001).

2.1.6.1 Risk evaluation based on traditional risk factors

The Framingham risk score (Wilson et al. 1998) estimates the 10-year risk for CHD by using information on sex, age, total or LDL cholesterol, HDL cholesterol, blood pressure, diabetes, and smoking (**Figures 2 and 3**). Each risk factor is first scored separately, and then the total risk is calculated as the sum of the risk points. Total sum of the risk points is then converted to absolute risk, which can be compared with the risk of an average person of the same age.

The relative risk might be more informative especially for young people. Since advanced age is the most important risk factor for CHD, the estimated 10-year risk for young people remains low, even if they have several risk factors. For example, a 36-year-old male who smokes and, has total cholesterol of 6 mmol/l and a blood pressure of 140/90 mmHg has a 10-year risk of 8%. By contrast, the risk for a 56-year-old male with similar values is 20%, which means that he is at high risk for developing CHD over a period of 10 years and should be considered for statin treatment. However, the average 10-year risk for 35- to 39-year-old males is 5%. Thus, the relative risk shows that 8% is a clearly elevated risk for a 36-year-old male. The protective effect of young age is even stronger in females. Even though CHD is the most common cause of death for women, the onset of the disease is later in life than for men. With the same risk factor values as above, the 10-year CHD risk for a 36-year-old female is only 2% and for a 56-year-old 15%. Nevertheless, the risks are still higher than for an average person of the same age.

Several similar prediction algorithms exist, but only PROCAM (Assmann et al. 2002) and the Reynolds Risk Score (Ridker et al. 2007) uses information on family history, which is considered to reflect genetic risk. Current prediction algorithms are not perfect; their accuracy for discriminating cases from noncases is around 80% (Wilson et al. 1998; Assmann et al. 2002; Conroy et al. 2003). This has motivated the search for new potential risk factors that could improve the risk estimation over and above traditional risk factors.



Step 1

Age		
Years	LDL Pts	Chol Pts
30-34	-1	[-1]
35-39	0	[0]
40-44	1	[1]
45-49	2	[2]
50-54	3	[3]
55-59	4	[4]
60-64	5	[5]
65-69	6	[6]
70-74	7	[7]

Step 2

LDL - C		
(mg/dl)	(mmol/L)	LDL Pts
<100	<2.59	-3
100-129	2.60-3.36	0
130-159	3.37-4.14	0
160-190	4.15-4.92	1
≥190	≥4.92	2

Cholesterol		
(mg/dl)	(mmol/L)	Chol Pts
<160	<4.14	[-3]
160-199	4.15-5.17	[0]
200-239	5.18-6.21	[1]
240-279	6.22-7.24	[2]
≥280	≥7.25	[3]

Step 3

HDL - C			
(mg/dl)	(mmol/L)	LDL Pts	Chol Pts
<35	<0.90	2	[2]
35-44	0.91-1.16	1	[1]
45-49	1.17-1.29	0	[0]
50-59	1.30-1.55	0	[0]
≥60	≥1.56	-1	[-2]

Step 4

Blood Pressure					
Systolic (mm Hg)	Diastolic (mm Hg) <80	80-84	85-89	90-99	≥100
<120	0 [0] pts				
120-129		0 [0] pts			
130-139			1 [1] pts		
140-159				2 [2] pts	
≥160					3 [3] pts

Note: When systolic and diastolic pressures provide different estimates for point scores, use the higher number

Step 5

Diabetes		
	LDL Pts	Chol Pts
No	0	[0]
Yes	2	[2]

Step 6

Smoker		
	LDL Pts	Chol Pts
No	0	[0]
Yes	2	[2]

(sum from steps 1-6)

Step 7

Adding up the points	
Age	_____
LDL-C or Chol	_____
HDL - C	
Blood Pressure	_____
Diabetes	_____
Smoker	_____
Point total	_____

Key	
Color	Relative Risk
green	Very low
white	Low
yellow	Moderate
rose	High
red	Very high

(determine CHD risk from point total)

Step 8

CHD Risk			
LDL Pts Total	10 Yr CHD Risk	Chol Pts Total	10 Yr CHD Risk
<-3	1%		
-2	2%		
-1	2%	[<-1]	[2%]
0	3%	[0]	[3%]
1	4%	[1]	[3%]
2	4%	[2]	[4%]
3	6%	[3]	[5%]
4	7%	[4]	[7%]
5	9%	[5]	[8%]
6	11%	[6]	[10%]
7	14%	[7]	[13%]
8	18%	[8]	[16%]
9	22%	[9]	[20%]
10	27%	[10]	[25%]
11	33%	[11]	[31%]
12	40%	[12]	[37%]
13	47%	[13]	[45%]
≥14	≥56%	[≥14]	[≥53%]

(compare to average person your age)

Step 9

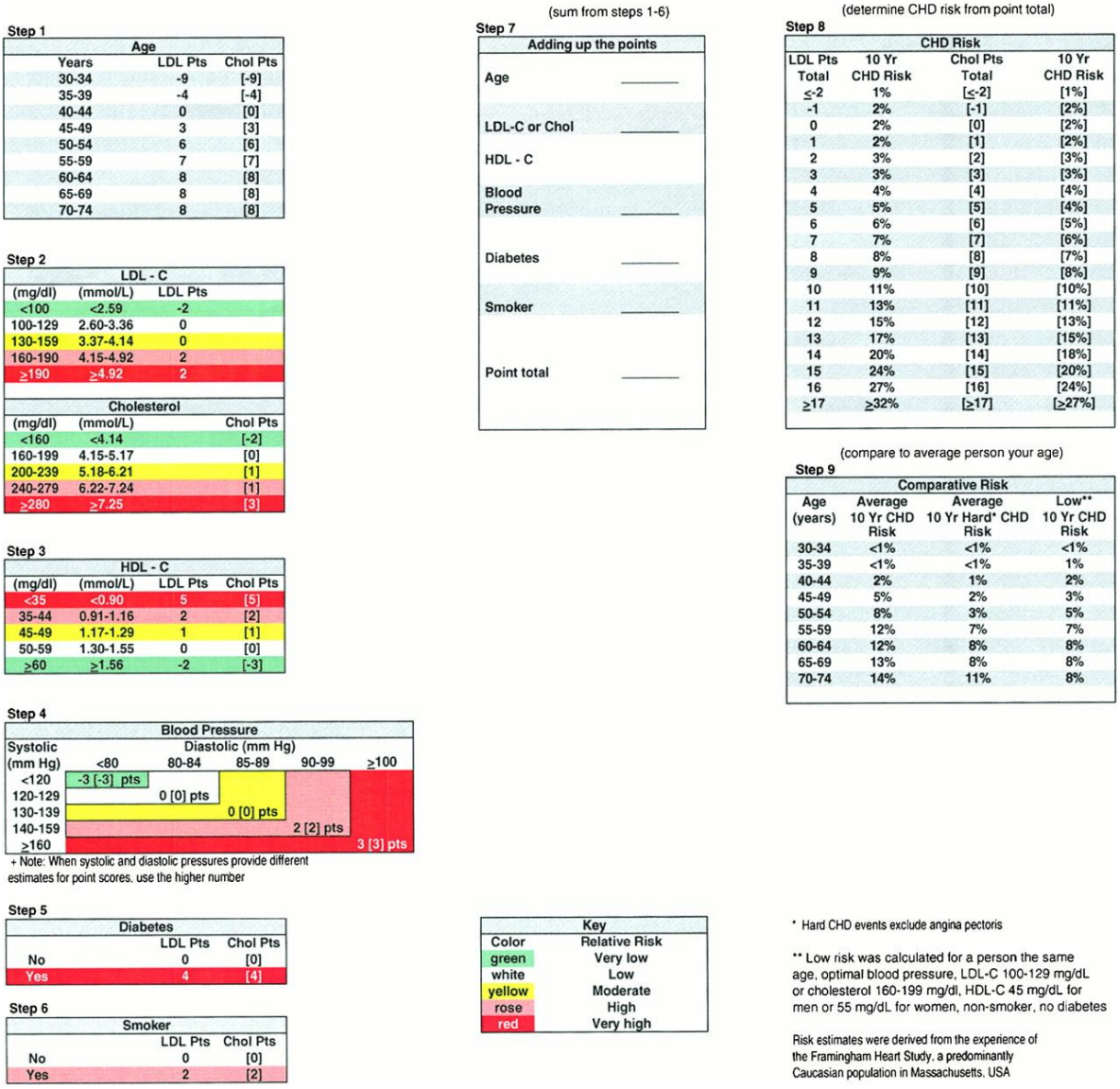
Comparative Risk			
Age (years)	Average 10 Yr CHD Risk	Average 10 Yr Hard* CHD Risk	Low** 10 Yr CHD Risk
30-34	3%	1%	2%
35-39	5%	4%	3%
40-44	7%	4%	4%
45-49	11%	8%	4%
50-54	14%	10%	6%
55-59	16%	13%	7%
60-64	21%	20%	9%
65-69	25%	22%	11%
70-74	30%	25%	14%

* Hard CHD events exclude angina pectoris

** Low risk was calculated for a person the same age, optimal blood pressure, LDL-C 100-129 mg/dL or cholesterol 160-199 mg/dl, HDL-C 45 mg/dL for men or 55 mg/dL for women, non-smoker, no diabetes

Risk estimates were derived from the experience of the Framingham Heart Study, a predominantly Caucasian population in Massachusetts, USA

Figure 2. Coronary heart disease (CHD) risk score sheet for men. Estimates risk for CHD over a period of 10 years based on the Framingham experience in men 30-74 years old at baseline. Abbreviations: Pts, points. Reprinted from Wilson et al. (1998) with permission from Wolters Kluwer Health.



Step 1

Age		
Years	LDL Pts	Chol Pts
30-34	-9	[-9]
35-39	-4	[-4]
40-44	0	[0]
45-49	3	[3]
50-54	6	[6]
55-59	7	[7]
60-64	8	[8]
65-69	8	[8]
70-74	8	[8]

Step 2

LDL - C		
(mg/dl)	(mmol/L)	LDL Pts
<100	<2.59	-2
100-129	2.60-3.36	0
130-159	3.37-4.14	0
160-190	4.15-4.92	2
≥190	≥4.92	2

Cholesterol		
(mg/dl)	(mmol/L)	Chol Pts
<160	<4.14	[-2]
160-199	4.15-5.17	[0]
200-239	5.18-6.21	[1]
240-279	6.22-7.24	[1]
≥280	≥7.25	[3]

Step 3

HDL - C			
(mg/dl)	(mmol/L)	LDL Pts	Chol Pts
<35	<0.90	5	[5]
35-44	0.91-1.16	2	[2]
45-49	1.17-1.29	1	[1]
50-59	1.30-1.55	0	[0]
≥60	≥1.56	-2	[-3]

Step 4

Blood Pressure				
Systolic (mm Hg)	Diastolic (mm Hg) <80	80-84	85-89	90-99	≥100
<120	-3 [-3] pts				
120-129		0 [0] pts			
130-139			0 [0] pts		
140-159				2 [2] pts	
≥160					3 [3] pts

+ Note: When systolic and diastolic pressures provide different estimates for point scores, use the higher number

Step 5

Diabetes		
	LDL Pts	Chol Pts
No	0	[0]
Yes	4	[4]

Step 6

Smoker		
	LDL Pts	Chol Pts
No	0	[0]
Yes	2	[2]

(sum from steps 1-6)

Step 7

Adding up the points	
Age	_____
LDL-C or Chol	_____
HDL - C	
Blood Pressure	_____
Diabetes	_____
Smoker	_____
Point total	_____

Key	
Color	Relative Risk
green	Very low
white	Low
yellow	Moderate
rose	High
red	Very high

(determine CHD risk from point total)

Step 8

CHD Risk			
LDL Pts Total	10 Yr CHD Risk	Chol Pts Total	10 Yr CHD Risk
≤-2	1%	[≤-2]	[1%]
-1	2%	[-1]	[2%]
0	2%	[0]	[2%]
1	2%	[1]	[2%]
2	3%	[2]	[3%]
3	3%	[3]	[3%]
4	4%	[4]	[4%]
5	5%	[5]	[4%]
6	6%	[6]	[5%]
7	7%	[7]	[6%]
8	8%	[8]	[7%]
9	9%	[9]	[8%]
10	11%	[10]	[10%]
11	13%	[11]	[11%]
12	15%	[12]	[13%]
13	17%	[13]	[15%]
14	20%	[14]	[18%]
15	24%	[15]	[20%]
16	27%	[16]	[24%]
≥17	≥32%	[≥17]	[≥27%]

(compare to average person your age)

Step 9

Comparative Risk			
Age (years)	Average 10 Yr CHD Risk	Average 10 Yr Hard* CHD Risk	Low** 10 Yr CHD Risk
30-34	<1%	<1%	<1%
35-39	<1%	<1%	1%
40-44	2%	1%	2%
45-49	5%	2%	3%
50-54	8%	3%	5%
55-59	12%	7%	7%
60-64	12%	8%	8%
65-69	13%	8%	8%
70-74	14%	11%	8%

* Hard CHD events exclude angina pectoris

** Low risk was calculated for a person the same age, optimal blood pressure, LDL-C 100-129 mg/dL or cholesterol 160-199 mg/dl, HDL-C 45 mg/dL for men or 55 mg/dL for women, non-smoker, no diabetes

Risk estimates were derived from the experience of the Framingham Heart Study, a predominantly Caucasian population in Massachusetts, USA

Figure 3. Coronary heart disease (CHD) risk score sheet for women. Estimates risk for CHD over a period of 10 years based on the Framingham experience in women 30-74 years old at baseline. Abbreviations: Pts, points. Reprinted from Wilson et al. (1998) with permission from Wolters Kluwer Health.

2.1.6.2 Statistical methods for testing new potential risk factors

Epidemiological studies have evaluated several novel risk factors for CHD. These include additional lipid-related markers and inflammatory biomarkers such as CRP and fibrinogen. These have been found to be associated with CHD in prospective studies (Ridker et al. 2000; Danesh et al. 2005; Melander et al. 2009; Di Angelantonio et al. 2012; Kaptoge et al. 2012). However, statistical significance does not necessarily indicate clinical usefulness. While a statistical association is a precondition, novel potential risk factors should also fulfill other criteria: First, the biomarker should be accurately measurable at a reasonable cost; second, it should provide incremental information beyond existing risk factors; and third, it should have an influence on clinical decision-making (Morrow and de Lemos 2007). Therefore, additional methods have been developed to address practical and clinical utility of novel risk factors. These methods include discrimination, reclassification, and calibration.

Discrimination is assessed with a Receiver Operating Characteristic (ROC) curve and C-index (area under the ROC). These are functions based on commonly used performance rates for medical tests. The rates can be defined as follows: 1) true positive rate is the probability of a positive test result for a case; 2) false positive rate is the probability of a positive test result for a noncase; 3) true negative rate is the probability of a negative test result for a noncase; and 4) false negative rate is the probability of a negative test result for a case. As a good medical test should be able to discriminate cases from noncases as accurate as possible, the test should maximize the true positive and negative rates and minimize the false positive and negative rates. ROC curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity). In cardiovascular risk assessment, the evaluation is based on the predicted risk rather than a binary test result. Thus, the C-index gives the probability that the predicted risk is higher for a case than for a noncase. C-indices vary from 0.5 (poor discrimination) to 1.0 (perfect discrimination). To evaluate whether a new risk factor improves prediction of future CHD, one can compare C-indices of two models (with and without the new risk factor). A limitation of C-index is that as a rank-based measure it does not quantify the absolute difference in risk. According to Cook (2007), "differences between 2 individuals who are at very low risk (e.g. 1.0% versus 1.1%) have the same impact on the c-statistic as 2 individuals who are at moderate versus high risk (e.g. 5% versus 20%) if their differences in rank are the same". However, the risk difference between the latter two individuals is more relevant clinically.

To overcome the limitations for C-index, Pencina et al. (2008) proposed two additional measures to evaluate the predictive performance of the model. Net Reclassification Improvement (NRI) compares the movements in the scale of predicted probabilities when the new risk factor is added to the model. Usually, the predicted risk is categorized into clinically meaningful risk categories (e.g. 0–5%, 5–10%, 10–20%, >20%), and the movements between the categories are recorded (**Table 3**). For cases, upward movements

imply improved prediction and downward movements worse prediction. The interpretation is the opposite for noncases. Thus, NRI indicates how much more correct versus incorrect reclassification occurs. Reclassification of the subjects in the intermediate-risk category (e.g. 10–20%), denoted as "clinical NRI", is particularly interesting, as in this group the treatment decisions are less clear than in high-risk individuals (Cook 2008; Pencina et al. 2008; Pencina et al. 2011).

Table 3. Risk reclassification with a biomarker score.

Predicted 10-year risk (without biomarker score)	**Predicted 10-year risk (with biomarker score)**			
CVD cases	0–5%	5–10%	10–20%	>20%
0–5%	0	0	0	0
5–10%	2	66	20	2
10–20%	0	20	99	21
>20%	0	0	4	37
Noncases	0–5%	5–10%	10–20%	>20%
0–5%	152	33	1	0
5–10%	80	869	101	6
10–20%	0	197	702	49
>20%	0	0	24	66

The basic model included area, age, sex, high-density lipoprotein (HDL) cholesterol, nonHDL cholesterol, systolic blood pressure, body mass index, smoking, diabetes, and cardiovascular drugs. Biomarker score included troponin I, N-terminal pro-brain natriuretic peptide, and C-reactive protein. Abbreviations: CVD, cardiovascular disease. Data adapted from Blankenberg et al. (2010) with permission from Wolters Kluwer Health.

Another discrimination measure, Integrated Discrimination Index (IDI), compares the mean differences in predicted risk between cases and noncases for the models with and without the new risk factor (Pencina et al. 2008). It indicates how far, on average, individuals move along the risk scale with the addition of new risk marker. If IDI is small and NRI is high, then most of the reclassification occurs adjacent to the risk thresholds (Lloyd-Jones 2010). Under some assumptions, IDI has also been shown to be equivalent with the change in mean absolute residuals and with the change in proportion of the explained variation (a generalization of R2 from linear to binary regression) when the new risk marker is added to the model (Pepe et al. 2008).

In addition to the methods described above, the model calibration needs to be on an acceptable level. Calibration is a measure of overall goodness-of-fit of the model, and it is usually evaluated visually by dividing data into deciles and tested with Hosmer-Lemeshow test, $P<0.05$ generally indicating poor calibration (Hosmer and Lemeshow

1999). It is a measure of how well predicted risk approximates with the actual observed risk.

As an example, some studies have evaluated whether substituting traditional risk factors, LDL and HDL cholesterol, with apoliprotein B (apoB) and apolipoprotein A1 (apoA1) would provide more accurate predictions of cardiovascular risk. These lipid markers have some benefits over LDL and HDL cholesterol. For example, they are less prone to measurement errors in the laboratory. Also, the apo(B)/apo(A1) ratio has been shown to be a strong risk marker for MI (Yusuf et al. 2004). However, these markers are not measured in all laboratories, possibly because they are more costly to measure than conventional lipid markers. In a large study of Emerging Risk Factors Collaboration (Di Angelantonio et al. 2012), the model containing conventional risk factors had a C-index of 0.7244. Substituting conventional lipid markers (total cholesterol and HDL) with apoB and apoA1 worsened CVD prediction, and addition of apoB and apoAI, lipoprotein(a), or lipoprotein-associated phospholipase A2 mass to conventional lipid markers led to modest improvements in risk discrimination (change in C-index 0.0006−0.002) or reclassification (NRI<1% for all lipid biomarkers). The modest improvements in prediction are most likely explained by the high correlation between tested biomarkers and traditional biomarkers. **Table 4** shows discrimination and reclassification metrics for some other candidate risk markers.

Table 4. Discrimination and reclassification of novel biomarkers.

Biomarker	End-point	C-index*	C-index change**	NRI	Reference
C-reactive protein	CVD	0.714	0.004	1.5%	Kaptoge et al. (2012)
Fibrinogen	CVD	0.717	0.003	0.8%	Kaptoge et al. (2012)
N-BNP	CHD	0.760	0.006	1.2%	Melander et al. (2009)
MR-proADM	CHD	0.760	0.004	2.4%	Melander et al. (2009)
Cystatin C	CHD	0.760	0.004	0.9%	Melander et al. (2009)
Carotid intima-media thickness	CVD	0.757	NR	0.8%	Den Ruijter et al. (2012)
Coronary artery calcium score	CHD	NR	0.05	19.3%	Kavousi et al. (2012)

* Basic model included age, sex, total and high-density lipoprotein cholesterol, systolic blood pressure, treatment of hypertension, smoking, and diabetes. ** Novel marker was added to the basic model. Abbreviations: NRI, net reclassification improvement; CVD, cardiovascular disease; CHD, coronary heart disease; N-BNP, N-terminal pro-B-type natriuretic peptide; MR-proADM, midregional proadrenomedullin; NR, not reported.

2.2 Complex disease genetics

CHD, Parkinson's disease, multiple sclerosis, and asthma are examples of complex diseases. A common feature for these diseases is that they are heritable in a multifactorial manner, which means that several genes together with lifestyle and environmental factors contribute to disease onset. For most diseases, many of the risk factors are unknown and under intensive research.

According to the common disease – common variant hypothesis, the genetic risk for complex diseases is composed of several common genetic variants, each of them having a small additive effect on the total risk. Accumulation of genetic risk variants increases the genetic susceptibility, but does not directly cause the disease. Thus, there is an important distinction between complex diseases and monogenic diseases that can be caused by a mutation in a single gene. While genetic risk is determined at conception, maintaining a healthy lifestyle can substantially decrease an individual's risk for a complex disease. The degree to which genetic variation contributes to disease onset depends on heritability of the disease.

2.2.1 Heritability

Phenotypic variation of complex traits is composed of both environmental and genetic sources. In a population, the total variance (V_T) of a trait is the sum of genetic (V_G) and environmental (V_e) variance. The broad-sense heritability (H^2) is defined as the fraction of phenotypic variability that is due to genetic variation:

$$H^2 = \frac{V_G}{V_T}$$

Broad-sense heritability takes into account all potential sources of genetic variation (additive, dominant, epistatic), whereas narrow-sense heritability (h^2) is the proportion of total variance that is due to the additive genetic variance (V_A):

$$h^2 = \frac{V_a}{V_T}$$

For binary disease traits, heritability can be defined similarly on an underlying, normally distributed liability scale (Visscher et al. 2008).

A common method for estimating heritability is to compare phenotypic differences of dizygotic and monozygotic twin pairs. The narrow-sense heritability can also be estimated from genome-wide SNP data in unrelated individuals, which has some advantages over traditional family-based methods (Speed et al. 2012). For example, heritability estimates from twin studies might be biased upwards due to shared environment (Zaitlen et al. 2013).

Variations in environmental factors, allele frequencies, and genetic effects have an impact on heritability estimates. Thus, heritability estimates in one population might not be generalizable to another population. Moreover, heritability of a trait is not constant over time and might fluctuate due to changes in genetic and environmental effects (Wray and Visscher 2008).

2.2.2 Structure of the human genome

Deoxyribonucleic acid (DNA) is a double-helical molecule packed into 22 autosomal chromosome pairs and sex chromosomes X and Y in the cell nucleus. Half of the chromosomes are inherited from the mother and the other half from the father. Females carry two X chromosomes and males carry one X and one Y chromosome. In addition to nuclear DNA, a small amount of DNA is located in the cell mitochondria. Mitochondrial DNA is inherited solely from the mother.

DNA consists of two strands of base pairs (bp), which are formed so that adenine (A) bonds with thymine (T) and cytosine (C) bonds with guanine (G) (**Figure 4**). Together these chemical bases (A,T,C,G) form a sequence of approximately 3 billion characters. Genes comprise on average a 21,000-bp-long stretch of DNA and are distributed unevenly along the genome (Roberts and McNally 2011). They are composed of segments that encode functional products (RNA transcripts and proteins), noncoding introns and promoter regions, which control the expression of the gene. The minority of the genome (~1.5%) encodes protein sequences, and according to the latest estimates, there are only approximately 21,000 protein-coding genes in the human genome (Clamp et al. 2007). The genetic code in the protein-coding genes determines how DNA is translated into proteins. The sequences of three consecutive nucleotides, three-letter terms known as codons, correspond to certain amino acids, building blocks for proteins. However, genes can produce multiple functionally distinct proteins due to a process called alternative splicing.

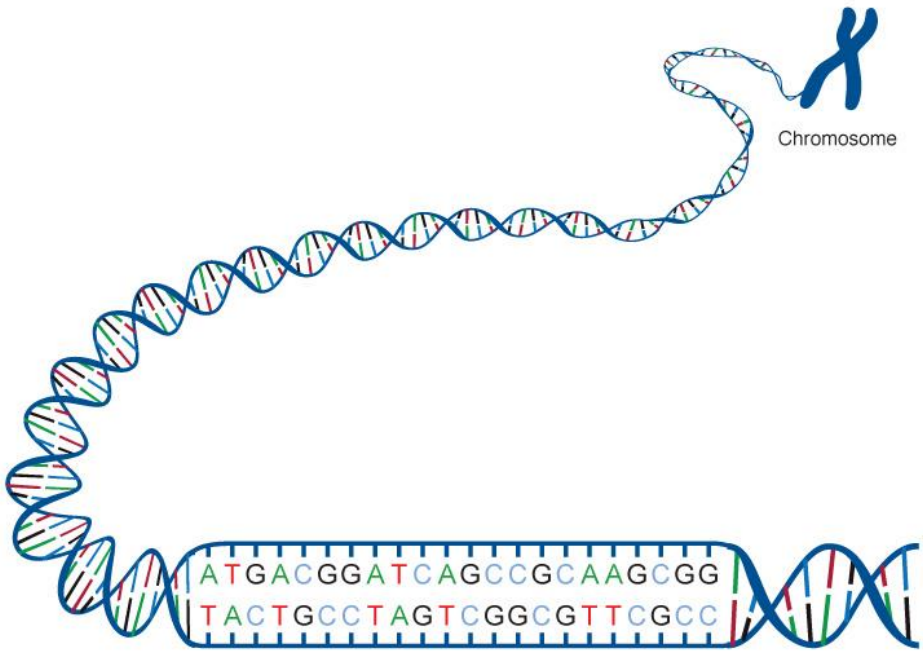


Figure 4. The DNA molecule consists of two strands held together by bonds between the bases. Adenine (A) pairs with thymine (T), and cytosine (C) with guanine (G). Reprinted from the National Institutes of Health. National Human Genome Research Institute (2012).

2.2.3 Genetic variation

The first draft sequence of the human genome was published in 2001 by the Human Genome Project and Celera Genomics (Lander et al. 2001; Venter et al. 2001). The genetic sequence is 99.9% similar for all humans. Genetic variation between individuals comes in many different forms and different sizes, from a single base to large segments of DNA. The types of variation include deletions, insertions, substitutions, repeated sequences, inversions, and other rearrangements. The most common type of variation is a single-nucleotide polymorphism (SNP), which is a variation in a single base. For example, the SNP might have two alleles, A and G, in a population with frequencies 40% and 60%, respectively. In this case, A is the minor allele of the locus, with a minor allele frequency (MAF) of 40%. On average, SNPs that have MAF of at least 1% occur in every 300 bases of the human genome. SNPs are widely utilized in gene mapping studies for common and complex traits. Depending on their location, SNPs may have different

effects on the phenotype. Most of the SNPs have small or neutral effects, but especially SNPs located in the coding regions that change the amino acid sequence might change the protein encoded by the gene.

SNPs are selected for gene mapping studies on the basis of a phenomenon called linkage disequilibrium (LD), the nonrandom association of alleles at different loci. When two chromosomes separate during meiosis, they cross over at various points. If two consecutive markers derive from different chromosomes, the event is called a recombination. Regions with a low number of recombinations, 'recombination coldspots', have large LD patterns, whereas regions with abundant recombinations, 'recombination hotspots', have smaller LD. Regions with high LD are usually described as haplotype blocks. Haplotype blocks can be used for SNP selection and are useful in regions where the LD structure is relatively discrete. Since LD between the markers within a haplotype block is high, sometimes only one SNP is sufficient to capture the genetic variation in that region. In some regions, however, the LD patterns are more complicated and it is not possible to accurately define the haplotype blocks (Wall and Pritchard 2003).

The international HapMap project was launched in 2002 to determine the common variation in the human DNA sequence. Using samples from different ethnic populations, the HapMap project created a haplotype map of the human genome (International HapMap Consortium 2003). Haplotype structures vary according to the population and reflect the immigration history of humans. The genetic variation is greatest in Africa, and haplotype diversity decreases with distance from Africa (Conrad et al. 2006). The genetic diversity between the populations needs to be taken into account when designing gene mapping studies. Tag SNPs used in one population might not be representative of another population. For example, two SNPs at chromosome 9 (rs4977574 and rs1333049), which were among the first association signals for CHD (Wellcome Trust Case Control Consortium 2007; Kathiresan et al. 2009b), tag the functional SNP rs1333047 in the HapMap2 CEU population. Due to the absence of LD, however, the association between rs1333049 or rs4977574 and CHD has not been replicated in the HapMap2 YRI population (**Figure 5**) (Schaub et al. 2012). The LD structure also affects genotype imputation (see Section 2.2.4.1).

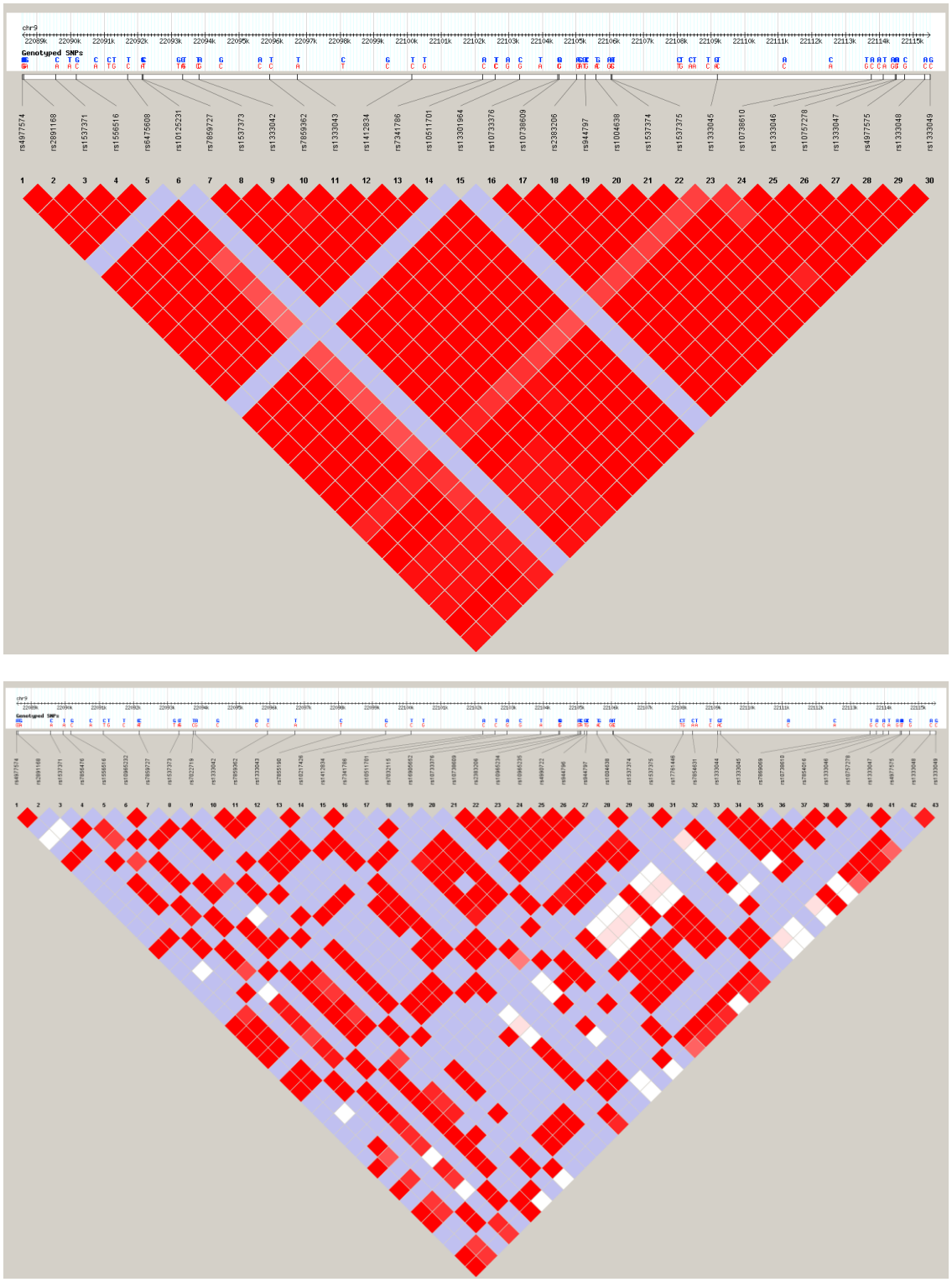


Figure 5. Patterns of linkage disquilibrium (LD) at chromosome 9p21 in the HapMap2 CEU (upper) and YRI (lower) populations. Abbreviations: CEU, Utah residents of Northern and Western European ancestry; YRI, Yoruba in Ibadan, Nigeria.

2.2.4 Genome-wide association studies

GWASs have uncovered common genetic variants for complex traits. The international HapMap project (2003) together with advances in large-scale genotyping technology and computational methods triggered the wave of GWASs in 2005, and since then, the number of GWAS publications has grown each year (Hindorff et al. 2013). The aim in GWAS is to identify genetic loci for common and complex traits by testing single variants one-by-one through the genome. The first GWASs utilized the SNP arrays of 100,000–250,000 variants (Klein et al. 2005; Maraganore et al. 2005), but the number of tested SNPs is growing as denser arrays are becoming available. Today, GWAS chips from Illumina Inc. include up to 5 million SNPs.

In GWAS, genetic effects are studied based on allele frequencies. The method can be applied for many different study settings. The simplest way is to compare allele frequencies between disease cases and controls. If the allele is more common in cases, it is said to be a risk allele for the disease. The magnitude of the effect can be estimated using a simple Chi-square test or logistic regression, which allows adjustment for several confounding factors. The other common type of association analysis is to study genetic effects for quantitative traits (e.g. BMI, LDL cholesterol), which are of interest because they are usually intermediates between the genetic factors and the disease. In this approach, the phenotypic distributions for each genotype are compared and modeled with linear regression analysis.

Testing genetic associations in thousands of loci creates the problem of multiple testing, which can lead to false-positive findings. Thus, a stringent threshold for statistical significance needs to be applied when interpreting the results. Bonferroni correction (significance level/number of tests) is a commonly used, albeit conservative, method. The significance level of 5×10^{-8} has been considered a standard in most GWASs.

2.2.4.1 Meta-analysis and imputation

Meta-analysis can be used to increase the power to detect SNP associations, as it pools the results from several studies. Effect sizes from different data sets can be combined with fixed or random effects meta-analysis. A fixed effects model assumes that there is a common genetic effect in all data sets and that observed differences are due to chance alone. A random effects model assumes that the effects are different in all data sets (Ioannidis et al. 2009).

Meta-analysis requires careful harmonization of data analysis, phenotype definitions, and SNP data. Since SNP arrays are usually not comparable in different studies, genotype imputation is applied to harmonize the coverage. This means that the missing genotypes are predicted based on the observed genotypes and the haplotype structure of the reference data, such as the HapMap2 or HapMap3 reference panel (Marchini and Howie 2010). The most extensive reference data today, produced by the 1000 Genomes Project,

comprise of sequences of 1092 individuals from 14 populations by combining low-coverage whole-genome and exome sequencing. It captures up to 98% of SNPs at a frequency of 1% (Abecasis et al. 2012). Using the 1000 Genomes reference population to impute an array of ~550,000 SNPs extends the data to up to ~14,000,000 polymorphic variants. Thus, the coverage of genetic data can be substantially increased by imputation. To perform analysis with this kind of data requires bioinformatics tools and high computational efficiency.

2.2.4.2 Success and limitations

Since 2005, GWASs have yielded over 10,000 SNPs associated with common, complex traits (**Figure 6**) (Hindorff et al. 2013). The majority of these SNPs are located within noncoding functional elements of the genome (Dunham et al. 2012).

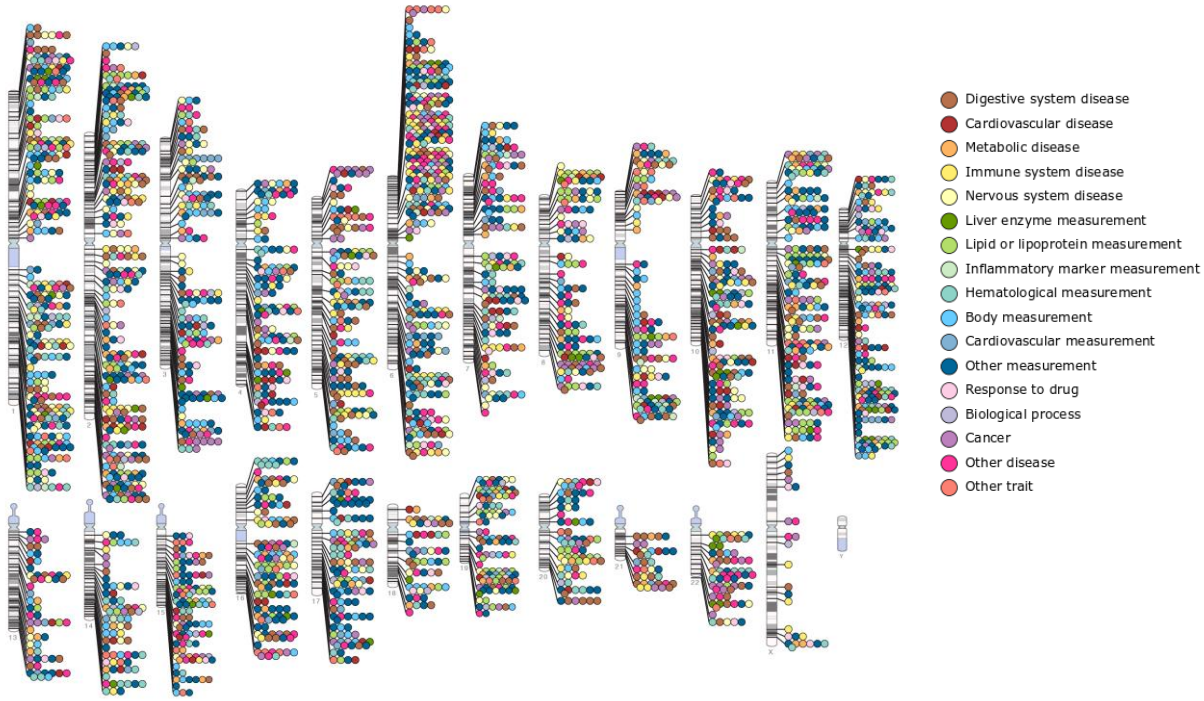


Figure 6. Published genome-wide associations ($P \leq 5 \times 10^{-8}$) through 12/2012 for 17 trait categories. Reprinted from Hindorff et al. (2013).

GWASs have shown that the genetic effect sizes for complex traits are modest. Exceptions include the first GWAS signal found for CHD at 9p21 (see Section 2.3.3), the variant in *FTO* gene associated with BMI (Frayling et al. 2007), and the SNP associated

with age-related macular degeneration that was identified by using only 96 cases and 50 controls (Klein et al. 2005). However, to detect variants with smaller effect sizes considerably larger study samples than those used in the initial GWASs are needed. Usually, it is not possible to gain sufficient power in a single data set, and thus, meta-analyses soon became the method of choice. As mentioned earlier, data harmonization is a challenge in meta-analysis. As a consequence, GWASs have been more successful for traits (e.g. blood lipids) that have similar, established measurement assays worldwide. Data harmonization might be expected to be more challenging for traits measured with a questionnaire (e.g. alcohol consumption).

The modest-effect common variants identified to date explain only a small proportion of the total genetic variability of complex traits. The unexplained part of the variability is probably accounted for by a large number of common SNPs with even smaller effects, but also by other types of genetic variation, like rare variants with large effects (Manolio et al. 2009). However, since GWASs rely on allele frequencies, rare variants are poorly detected in GWASs.

2.3 Genetics of coronary heart disease and its risk factors

2.3.1 Heritability of coronary heart disease and its risk factors

The contribution of genetic factors to the risk for CHD is about equal to the contribution of environmental risk factors. In the studies of Swedish and Danish twins, the heritability for CHD death varied from 52% to 57% in males and from 38% to 58% in females (Wienke et al. 2001; Zdravkovic et al. 2002). The heritability estimate based on the genome-wide SNP data of approximately 2000 cases and 3000 controls was 41% (Speed et al. 2012). A substantial part of the phenotypic variability of cardiovascular risk factors is due to genetic variation. Heritability varies from 61% to 71% for blood lipids (Kettunen et al. 2012). Long-term measurements of systolic and diastolic blood pressure have heritability estimates of 57% and 56%, respectively, whereas the corresponding estimates for temporal blood pressure measurements are 42% and 39% (Levy et al. 2000).

2.3.2 Mendelian forms of coronary heart disease

Monogenic forms of CHD that may result from a mutation in a single gene are called Mendelian lipid disorders. One example is familial hypercholesterolemia (FH), which is characterized by high LDL cholesterol levels already at an early age. Studies have revealed several genetic mutations that cause severe FH. The identification of Mendelian CHD genes has led to better understanding of the pathology of these diseases and supported the development of novel diagnostics and treatment. FH is usually caused by a mutation in the LDL receptor (*LDLR*) gene (Brown and Goldstein 1986). *LDLR* is responsible for removing LDL cholesterol from the blood stream, and the mutation in the gene leads to accumulation of LDL in the blood. As a consequence, FH, in particular the homozygous form of the disease, is characterized by extremely high LDL cholesterol and early-life CHD. Brown and Goldstein were awarded the Nobel Prize in 1985 for their identification of *LDLR* in FH. Indeed, the discovery of *LDLR* has been a success story in the field of cardiovascular genetics, as the knowledge gained from this finding regarding *LDLR* function, structure, and regulation eventually explained the LDL cholesterol-lowering effects of statins (Lagor and Millar 2010), originally recognized as inhibitors of cholesterol biosynthesis (Endo et al. 1976).

Mutations in *PCSK9*, *APOB*, *ABCG5*, *ABCG8* and *ARH* have also been found to cause FH (Soria et al. 1989; Garcia et al. 2001; Abifadel et al. 2003; Kathiresan and Srivastava 2012). Later, common variants in these regions were observed to be associated with general CHD or blood lipids in GWASs (Kathiresan et al. 2008; Willer et al. 2008; Aulchenko et al. 2009; Kathiresan et al. 2009a; Kathiresan et al. 2009b; Deloukas et al. 2012). This genetic overlap suggests that at least partially similar biological mechanisms could underlie Mendelian monogenic and common forms of CHD.

2.3.3 Genome-wide association studies of coronary heart disease

First GWAS finding for CHD was locus 9p21, identified simultaneously in three studies (Helgadottir et al. 2007; McPherson et al. 2007; Wellcome Trust Case Control Consortium 2007). The risk allele is common (frequency ~50% in European populations) and confers ~1.2-fold risk for CHD (Deloukas et al. 2012), which is a substantially higher risk than for other identified common variants (**Figure 7**). However, the genetic variants in this region are located in the gene desert, and the functional basis for the observed associations was long unknown. Recent functional studies have, however, reported that the SNPs in the locus play a role in inflammatory signaling response (Harismendy et al. 2011; Schaub et al. 2012), supporting the role of inflammation in CHD. Interestingly, there is also an association signal for type 2 diabetes in 9p21 (Saxena et al. 2007; Scott et al. 2007; Zeggini et al. 2007), but the lead SNP is completely uncorrelated with the lead SNP in the CHD locus.

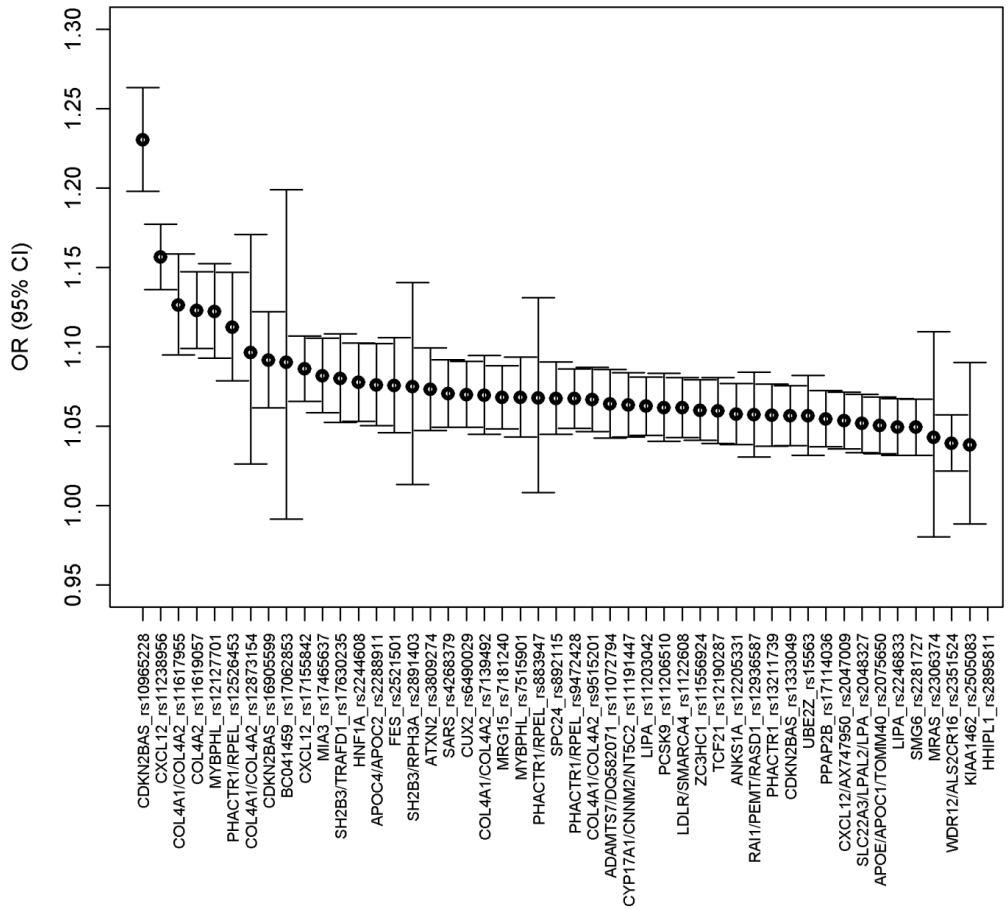


Figure 7. Odds ratios (ORs) and 95 % confidence intervals (CIs) for SNPs associated with coronary heart disease. Data adapted from Deloukas et al. (2012).

The largest CHD GWAS to date (Deloukas et al. 2012), including 63,746 CHD cases and 130,681 controls, identified 15 novel loci for CHD. Combined with the previously known loci, the number of independent genome-wide significant SNPs for CHD is currently over 40. Of these loci, 12 SNPs are associated also with lipids (mainly LDL cholesterol) and 5 with blood pressure. The authors also report 104 independent genetic variants ($r^2 < 0.2$) that were associated with CHD at a 5% false discovery rate. Together, these variants explain approximately 10.6% of CHD heritability. For most loci, the functional mechanism is unknown, but the authors report that the main pathways for the CHD loci are lipid metabolism and inflammation.

Most of the CHD GWASs have been performed in European populations, but recently, the GWAS in Han Chinese identified four loci (*TTC32-WDR35*, *GUCY1A3*, *C6orf10-BTNL2* and *ATP2B1*) not previously linked with CHD in European populations (Lu et al. 2012). Nevertheless, the SNPs in *GUCY1A3* and *ATP2B1* have been mapped to blood pressure levels also in European populations (Levy et al. 2009; Ehret et al. 2011; Wain et al. 2011). Another example of population-specific association signals is found in the study of Takeuchi et al. (2012), where two novel susceptibility loci (*ALDH2* and *HLA/DRB-DQB*) were identified in a Japanese study sample.

2.3.4 Genome-wide association studies of lipids and blood pressure

Triglycerides, LDL, HDL, and total cholesterol have been studied extensively in GWASs. A large study comprising > 100,000 individuals of European ancestry reported 95 independent loci associated with blood lipids, 59 of which were reported for the first time (Teslovich et al. 2010). Only a handful of these variants were also associated with CHD in the same study. These included the following variants in or near previously reported CHD genes: *SORT1*, *LPA*, *HNF1A*, and *LDLR* (Erdmann et al. 2009; Kathiresan et al. 2009b; Tregouet et al. 2009). Some loci were later identified in CHD GWASs: *LPL*, *TRIB1*, *ABO*, and *APOA1* (Schunkert et al. 2011; Deloukas et al. 2012). All of these variants had pleiotropic effects, that is, they were associated with more than one lipid phenotype. For most loci, however, the lead trait was LDL or total cholesterol, supporting the causal role of LDL cholesterol in CHD. Contradictory, four loci (*IRS1*, *KLF14*, *C6orf106*, and *NAT2*) had specific associations only with HDL cholesterol and triglycerides, whose pathological roles in CHD are less clear. Nevertheless, these genes have also been linked with type 2 diabetes (Voight et al. 2010) and other metabolic traits (Kilpeläinen et al. 2011). Due to pleiotropic effects with other cardiovascular risk factors, it remains unclear which of these risk factors are responsible for the detected CHD association.

Many of the lipid GWAS hits overlap with of loci causing Mendelian dyslipidemic syndromes (**Figure 8**). This indicates that the genetic architecture of lipid traits is a mixture of rare variants with large effects and common variants with small effects (Kathiresan and Srivastava 2012). Rare large-effect variants might especially affect the

higher end of the phenotypic distribution. For example, in addition to common risk variants, cases with hypertriglyceridemia have been reported to carry a significantly higher burden of rare risk variants (MAF < 1% in controls) in four lipid genes (*APOA5*, *GCKR*, *LPL*, *APOB*) (Johansen et al. 2010).

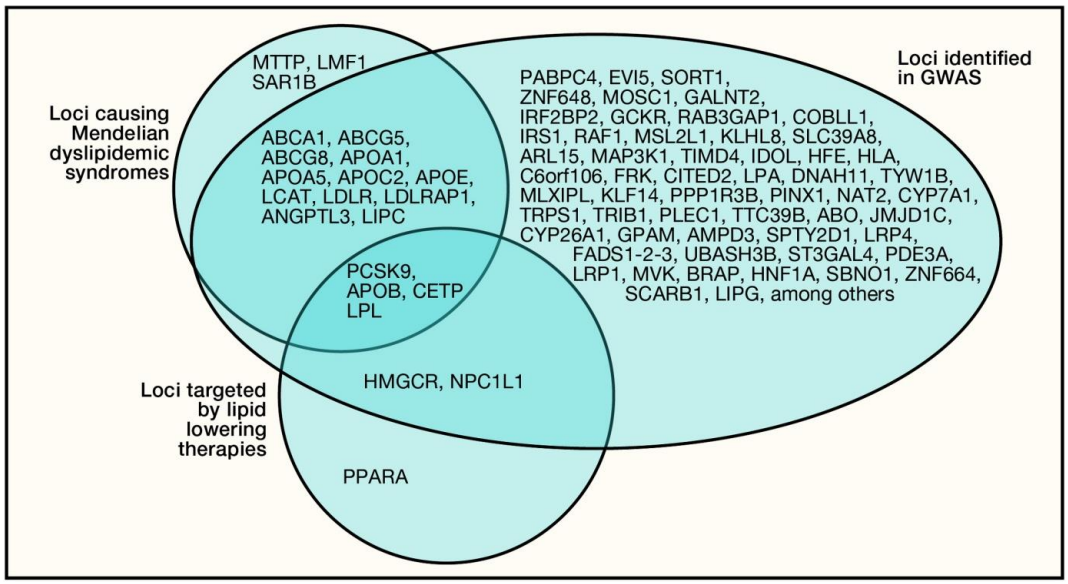


Figure 8. Overlap of genetic loci causing Mendelian lipid syndromes, those targeted by lipid-lowering therapies, and those identified in GWASs. Reprinted from Kathiresan and Srivastava (2012) with permission from Elsevier.

The first GWASs for blood pressure identified SNPs in or near 13 genes: *ZNF652*, *CACNB2*, *FGF5*, *c10orf107*, *SH2B3*, *CYP1A2*, *ULK4*, *TBX3/TBX5*, *ATP2B1*, *CYP17A1*, *MTHFR*, *PLCD3*, and *PLEKHA7* (Levy et al. 2009; Newton-Cheh et al. 2009). Of these, the same variant in or near *SH2B3* (rs3184504) has also been associated with CHD (Schunkert et al. 2011) and platelet counts (Gieger et al. 2011). The other variant near *CYP17A1* (rs12413409), which is in perfect LD with the blood pressure SNP (rs11191548), is also associated with CHD (Schunkert et al. 2011). In addition, *CYP17A1* is one of the 12 genes identified to cause monogenic hypertension. All 12 genes are involved in two groups of pathways: renal sodium handling and steroid hormone metabolism. The loss-of-function mutation in *CYP17A1* causes rare congenital adrenal hyperplasia, which is characterized by hypokalemia, gonadal deficiency, and hypertension (Ehret and Caulfield 2013).

In 2011, a large GWAS consortium (N=200,000) identified 16 novel blood pressure loci (Ehret et al. 2011). The majority of the SNPs identified so far are located in regions that do not contain previously found candidate genes for blood pressure. Altogether, 29 variants explain only 1-2% of blood pressure variance.

3 AIMS OF THE STUDY

In this study, we used the genetic information obtained from several large-scale genome-wide association studies of coronary heart disease and two of its important risk factors, lipids and blood pressure. The aim was to estimate the genetic effects in independent data sets and to evaluate the clinical and practical utility of these findings for public health.

The study has four specific aims:

1) To test whether previously identified genetic markers are associated with lipids and blood pressure in a longitudinal study design.

2) To estimate longitudinal and age-specific effects of genetic markers for lipids and blood pressure and to evaluate the ability of genetic risk scores to predict dyslipidemia and hypertension in young adults.

3) To estimate the genetic risk for coronary heart disease in independent prospective and case-control data sets.

4) To evaluate predictive ability of genetic risk scores in cardiovascular risk assessment.

4 MATERIALS AND METHODS

4.1 Study populations

4.1.1 FINRISK 1992, 1997, and 2002

FINRISK studies have been conducted every 5 years to monitor trends in CVD risk factor levels in Finland. The survey was initiated in 1972 in eastern Finland and North Karelia. The FINRISK 1992−2002 surveys have been sampled from up to five geographical areas. Each FINRISK survey is an independent, stratified random sample drawn from the national population register. Health information has been collected by mailed questionnaire, physical measurements, and clinical examination (Vartiainen et al. 2010).

Study protocols have been approved by the Institutional Review Board of Helsinki University Hospital, Helsinki, Finland. All participants provided written informed consent. FINRISK cohorts were used in Projects III and IV.

4.1.2 Health 2000

The Health 2000 survey was conducted in 2000 and 2001 with a stratified two-stage cluster sampling design. In the first stage, 80 health centers of 249 in total were selected to represent clusters. Then, individual persons were randomly sampled from these clusters. The data were collected with interviews, questionnaires, measurements, and clinical examinations (National Public Health Institute 2008).

The study protocol was approved by the Ethics Committees of the National Public Health Institute and Research in Epidemiology and Public Health at the Hospital District of Helsinki and Uusimaa. All participants provided written informed consent. The cohort was included in Projects III and IV.

4.1.3 Young Finns Cohort

The Cardiovascular Risk in Young Finns Study started in 1980, when 3596 children and adolescents aged 3, 6, 9, 12, 15, and 18 years participated in the first cross-sectional survey. Participants were randomly selected from five cities (Helsinki, Kuopio, Oulu, Tampere, and Turku) and their rural surroundings. The follow-up surveys were conducted in 1983, 1986, 2001, and 2007, when the participants were aged between 30 and 45. Various measurements (questionnaires, physical measurements, blood tests) of cardiovascular risk factors were taken in each of the follow-ups. DNA was extracted in 2001.

The study protocol was approved by the local ethics committees (University Hospitals of Helsinki, Kuopio, Oulu, Tampere, and Turku) and all subjects gave written informed consent (Raitakari et al. 2008). The cohort was included in Projects I and II.

4.1.4 Corogene

The Corogene cohort was collected from Finnish patients assigned to undergo coronary angiogram in the Helsinki University Central Hospital between June 2006 and March 2008. Patients were assigned to four groups: 1) patients without CHD, 2) patients with stable CHD, 3) patients with ACS, and 4) patients with other ischemic events. Data collection included a questionnaire, hospital records, physical examinations, and laboratory sampling (Vaara et al. 2012).

Altogether 2500 patients with ACS or previous MI were selected for genome-wide genotyping. Controls for these samples were obtained from FINRISK participants living in the Helsinki-Vantaa region. For each case, two ACS-free controls were matched by sex and birth year. In total, 2101 cases and 3914 controls (1453 unique) were included in this study.

The study protocol was approved by the Ethics Committee of Helsinki University Hospital, Internal Medicine. All participants provided written informed consent. The Corogene study cohort was included in Project III.

4.1.5 Malmö Diet and Cancer Study – Cardiovascular Cohort

The baseline examination of the Malmö Diet and Cancer Study (MDCS) was conducted between 1991 and 1996 for 53,000 middle-aged subjects. The Cardiovascular Cohort (MDC-CC) consists of 6103 randomly selected participants from MDCS. The data were collected with a questionnaire, medical history assessment, physical examination, and laboratory measurements.

The MDCS was approved by the Ethics Committee of Lund University, Sweden. All participants provided written informed consent (Berglund et al. 1993; Persson et al. 2007). The cohort was included in Project III.

4.1.6 Malmö Preventive Project

The Malmö Preventive Project (MPP) was set up in 1974. A total of 33,346 individuals participated in health screening during 1974–1992. Baseline data were collected with a questionnaire, physical examination, and biochemical analyses. Altogether 17,284 individuals were re-screened during 2002–2006. Of these individuals, 2400 were excluded from the present study because of a lack of DNA or crucial clinical information

or if they were already included in the MDC-CC cohort. Thus, 14,884 individuals comprise this study population.

The study protocol was approved by the Ethics Committee of Lund University, Sweden. All participants provided written informed consent. The cohort was included in Project III.

4.2 Genetic markers

4.2.1 SNP selection

SNP selection was based on GWASs of CHD or MI (Erdmann et al. 2009; Kathiresan et al. 2009b; Coronary Artery Disease Genetics Consortium 2011; Schunkert et al. 2011), lipids (Teslovich et al. 2010), and blood pressure (Levy et al. 2009; Newton-Cheh et al. 2009). In total, 28 CHD/MI SNPs, 131 lipid SNPs (from 95 independent loci), and 13 blood pressure SNPs were selected for Projects I–IV (**Table 5**).

Table 5. Selected single-nucleotide polymorphisms (SNPs) for coronary heart disease, blood pressure, and lipids.

SNP	LOCUS	TRAIT(S)	SNP	LOCUS	TRAIT(S)	SNP	LOCUS	TRAIT(S)
rs646776	*CELSR2, PSRC1, SORT1*	CHD	rs4759375	*SBNO1*	HDL	rs11065987	*BRAP*	LDL, TC
rs11206510	*PCSK9*	CHD	rs4660293	*PABPC4*	HDL	rs2479409	*PCSK9*	LDL, TC
rs17465637	*MIA3*	CHD	rs605066	*CITED2*	HDL	rs2332328	*NYNRIN*	LDL
rs6725887	*WDR12*	CHD	rs6450176	*ARL15*	HDL	rs1129555	*GPAM*	LDL
rs2306374	*MRAS*	CHD	rs1800961	*HNF4A*	HDL, TC	rs2807834	*MOSC1*	LDL, TC
rs12526453	*PHACTR1*	CHD	rs4148008	*ABCA8*	HDL	rs11220462	*ST3GAL4*	LDL
rs3798220	*LPA*	CHD	rs2925979	*CMIP*	HDL	rs1030431	*CYP7A1*	LDL, TC
rs4977574	*CDKN2A/B, ANRIL*	CHD	rs643531	*TTC39B*	HDL	rs651007	*LDLR*	TC
rs1746048	*CXCL12*	CHD	rs3741414	*LRP1*	HDL	rs2072183	*NPC1L1*	TC
rs3184504	*SH2B3*	CHD, DBP	rs7134594	*MVK*	HDL	rs2902940	*MAFB*	TC
rs1122608	*LDLR*	CHD	rs1515100	*IRS1*	HDL	rs6759321	*RAB3GAP1*	TC
rs9982601	*MRPS6*	CHD	rs174601	*FADS1/2/3*	HDL	rs7515577	*EVI5*	TC
rs2259816	*HNFA1*	CHD	rs4731702	*KLF14*	HDL	rs4297946	*TOP1*	TC
rs17114036	*PPAP2B*	CHD	rs2652834	*LACTB*	HDL	rs7239867	*LIPG*	TC
rs17609940	*ANKS1A*	CHD	rs1084651	*LPA*	HDL	rs1961456	*NAT2*	TC
rs11556924	*ZC3HC1*	CHD	rs4420638	*APOE/C1/C2*	HDL, LDL, TC	rs9488822	*FRK*	TC
rs579459	*ABO*	CHD	rs181362	*UBE2L3*	HDL	rs2737229	*TRPS1*	TC
rs12413409	*CYP17A1, CNNM2, NT5C2*	CHD	rs2293889	*TRPS1*	HDL	rs1260326	*GCKR*	TC, TG
rs964184	*ZNF259, APOA5/A4/C3/A1*	CHD, HDL, LDL, TC, TG	rs2814944	*C6orf106*	HDL	rs7206971	*OSBPL7*	TC
rs4773144	*COL4A1, COL4A2*	CHD	rs881844	*STARD3*	HDL	rs10832963	*SPTY2D1*	TC
rs2895811	*HHIPL1*	CHD	rs838880	*SCARB1*	HDL	rs174550	*FADS1/2/3*	TC
rs3825807	*ADAMTS7*	CHD	rs7134375	*PDE3A*	HDL	rs2814982	*C6orf106*	TC
rs12936587	*RASD1, SMCR3, PEMT*	CHD	rs10808546	*TRIB1*	HDL	rs7941030	*UBASH3B*	TC
rs216172	*SMG6, SRR*	CHD	rs13107325	*SLC39A8*	HDL	rs492602	*FLJ36070*	TC
rs46522	*UBE2Z, GIP, ATP5G1, SNF8*	CHD	rs4765127	*ZNF664*	HDL	rs2255141	*GPAM*	TC
rs1412444	*LIPA*	CHD	rs737337	*LOC55908*	HDL	rs2285942	*DNAH11*	TC
rs4380028	*ADAMTS7-MORF4L1*	CHD	rs12328675	*COBLL1*	HDL	rs2290159	*RAF1*	TC

rs10953541	*NR*	CHD	rs12967135	*MC4R*	HDL	rs11220463	*ST3GAL4*	TC
rs16948048	*ZNF652*	DBP	rs4082919	*PGS1*	HDL	rs581080	*TTC39B*	TC
rs11014166	*CACNB2*	DBP	rs6511720	*LDLR*	LDL, TC	rs2277862	*ERGIC3*	TC
rs16998073	*FGF5*	DBP	rs649129	*ABO*	LDL	rs439401	*APOE/C1/C2*	TG
rs1530440	*c10orf107*	DBP	rs10401969	*CILP2*	LDL, TC, TG	rs261342	*LIPC*	TG
rs1378942	*CSK-CYP1A2*	DBP	rs12916	*HMGCR*	LDL, TC	rs1495743	*NAT2*	TG
rs9815354	*ULK4*	DBP	rs4299376	*ABCG5/8*	LDL, TC	rs7811265	*MLXIPL*	TG
rs2384550	*TBX3/X5*	DBP	rs1367117	*APOB*	LDL, TC	rs11613352	*LRP1*	TG
rs2681492	*ATP2B1*	SBP	rs629301	*SORT1*	LDL, TC	rs4810479	*PLTP*	TG
rs11191548	*CYP17A1*	SBP	rs2000999	*HPR*	LDL, TC	rs442177	*KLHL8*	TG
rs17367504	*MTHFR-NPPB*	SBP	rs2902941	*MAFB*	LDL	rs1321257	*GALNT2*	TG
rs12946454	*PLCD3*	SBP	rs217386	*NPC1L1*	LDL	rs2131925	*ANGPTL3*	TG
rs381815	*PLEKHA7*	SBP	rs11153594	*FRK*	LDL	rs11776767	*PINX1*	TG
rs3764261	*CETP*	HDL, TC	rs909802	*TOP1*	LDL	rs174546	*FADS1/2/3*	TG
rs1532085	*LIPC*	HDL, TC	rs247616	*CETP*	LDL	rs5756931	*PLA2G6*	TG
rs386000	*LILRA3*	HDL	rs3850634	*ANGPTL3*	LDL, TC	rs12310367	*ZNF664*	TG
rs7241918	*LIPG*	HDL	rs12027135	*LDLRAP1*	LDL, TC	rs2247056	*HLA*	TG
rs1689800	*ZNF648*	HDL	rs3757354	*MYLIP*	LDL, TC	rs7205804	*CETP*	TG
rs6065906	*PLTP*	HDL	rs6882076	*TIMD4*	LDL, TC	rs2412710	*CAPN3*	TG
rs3136441	*LRP4*	HDL	rs1800562	*HFE*	LDL, TC	rs10761731	*JMJD1C*	TG
rs1883025	*ABCA1*	HDL, TC	rs1169288	*HNF1A*	LDL, TC	rs2068888	*CYP26A1*	TG
rs9987289	*PPP1R3B*	HDL	rs3177928	*HLA*	LDL, TC	rs10195252	*COBLL1*	TG
rs12678919	*LPL*	HDL, TG	rs7225700	*OSBPL7*	LDL	rs2943645	*IRS1*	TG
rs17145738	*MLXIPL*	HDL	rs514230	*IRF2BP2*	LDL, TC	rs645040	*MSL2L1*	TG
rs16942887	*LCAT*	HDL	rs11136341	*PLEC1*	LDL, TC	rs11649653	*CTF1*	TG
rs2923084	*AMPD3*	HDL	rs12670798	*DNAH11*	LDL	rs13238203	*TYW1B*	TG
rs1042034	*APOB*	HDL, TG	rs2954022	*TRIB1*	LDL, TC	rs2954029	*TRIB1*	TG
rs4846914	*GALNT2*	HDL	rs1564348	*LPA*	LDL, TC	rs2929282	*FRMD5*	TG
rs7115089	*UBASH3B*	HDL	rs2126259	*PPP1R3B*	LDL, TC	rs9686661	*MAP3K1*	TG
rs7255436	*ANGPTL4*	HDL	rs174583	*FADS1/2/3*	LDL	rs1553318	*TIMD4*	TG

Abbreviations: CHD, coronary heart disease; SBP, systolic blood pressure; DBP, diastolic blood pressure; TC, total cholesterol; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; TG, triglycerides.

4.2.2 Genotyping and quality control

Genotyping of selected CHD SNPs was done at the Institute for Molecular Medicine Finland FIMM, Wellcome Trust Sanger Institute (UK), Broad Institute (USA) and Lund University (Sweden) with the Sequenom platform (iPLEX MassARRAY, San Diego, CA, USA) or Tagman (Applied Biosystems, Foster City, CA, USA). SNPs were in Hardy-Weinberg equilibrium (HWE) and uncorrelated ($r^2 < 0.4$), and had a genotype call rate > 98% and a sample call rate > 95%.

Genome-wide SNP data for Corogene and Young Finns cohorts were obtained using Illumina genotyping arrays (Illumina HumanHap 610-Quad SNP array, Illumina Human670K BeadChip). Genotyped SNPs went through the quality control with the following inclusion thresholds; call rate ≥ 95%, MAF ≥ 1%, and a HWE P-value ≥ 10^{-6}. Genome-wide SNP data were imputed with IMPUTE 1.0 using two different reference sets: 1) HapMap3 CEU and TSI populations extended with the Finnish reference data consisting of 81 individuals, and 2) HapMap2 CEU population. We extracted 131 genotyped or imputed lipid SNPs and 13 blood pressure SNPs from the Young Finns data. Directly genotyped SNPs were coded as 0/1/2 and imputed SNPs as predicted allele dosage ranging from 0.0 to 2.0. We primarily chose SNPs imputed with an extended HapMap3 reference panel, as it yields better imputation quality (Surakka et al. 2010). If the SNP was not found in the data imputed with the Hapmap3 reference, we used HapMap2 imputed data. For Project III, 13 directly genotyped CHD SNPs were extracted from the genome-wide Corogene data.

4.2.3 Genetic risk scores

Common genetic variants that have been identified in GWASs typically have modest effects on the phenotype of interest. Thus, it has become a common practice to generate GRSs that reflect the joint genetic effect of individual SNPs. Based on the assumption of additional genetic effects, GRSs can be constructed simply as the sum of the risk alleles. In a weighted approach, individual SNPs are weighted with the effect sizes, obtained usually from the original GWASs. In this approach, rather than giving equal weight for each SNP, SNPs with large effects contribute to the score most.

We selected the effect sizes (Projects I and II: beta coefficients or Projects III and IV: odds ratios) for weights for each SNP from discovery GWASs or later GWASs with a larger study sample (Clarke et al. 2009; Erdmann et al. 2009; Gudbjartsson et al. 2009; Kathiresan et al. 2009b; Levy et al. 2009; Newton-Cheh et al. 2009; Teslovich et al. 2010; Coronary Artery Disease Genetics Consortium 2011; Schunkert et al. 2011). GRSs were calculated for each trait as the weighted sum of the risk alleles. The average number of the risk alleles was used to impute missing genotypes.

We calculated two scores for CHD; one based on 13 SNPs (Project III) and the other based on 28 SNPs (Project IV). Four GRSs were generated for blood lipids (Project I); total cholesterol (comprising 52 SNPs), LDL (37 SNPs), HDL (47 SNPs), and triglycerides (32 SNPs). Blood pressure score (Project II) was constructed by using 13 SNPs, but we also calculated GRSs separately for diastolic (8 SNPs) and systolic (5 SNPs) blood pressure. For each GRS, SNPs were independent, thus representing independent genetic loci. However, GRSs for CHD, lipids, and blood pressure are partly overlapping (**Figure 9**).

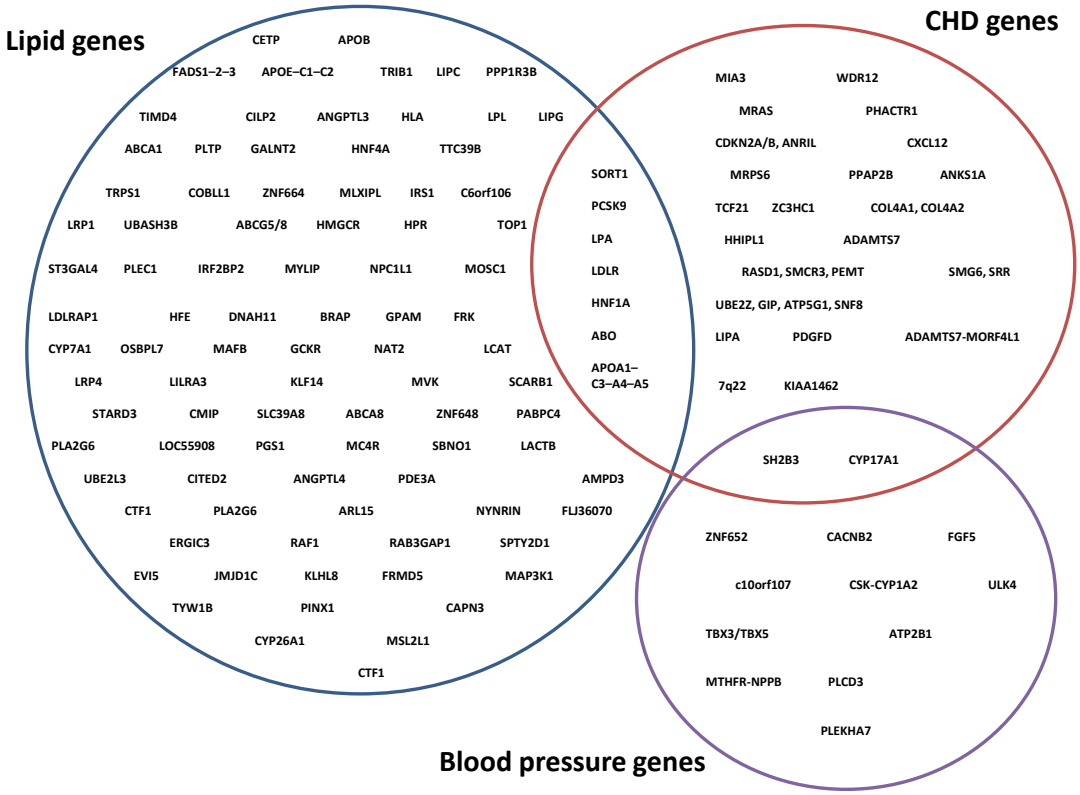


Figure 9. Reported genes for lipids, blood pressure and coronary heart disease (CHD) in genome-wide association studies.

4.3 Phenotypes

4.3.1 Cardiovascular events

In prospective Finnish data sets (FINRISK 1992–2002, Health 2000) follow-up morbidity and mortality data were obtained from the Finnish National Hospital Discharge Register and the Finnish National Cause-of-Death Register (Laatikainen et al. 2009). Diagnoses of cardiovascular events were recorded using ICD-9 during 1987–1995 and ICD-10 from 1996 onwards. The validity of diagnoses in the Finnish Hospital Discharge Register varies from satisfactory to very good, with a positive predictive value between 73 and 97 for vascular diagnoses (Sund 2012). For Project III, the follow-up ended on Dec 31, 2007. For Project IV, the follow-up was extended to Dec 31, 2010 for the FINRISK studies and to Dec 31, 2008 for Health 2000.

In Projects III and IV, the main end-point of interest was CHD, which was defined as MI, unstable angina pectoris, coronary revascularization (coronary artery bypass graft or percutaneous transluminal coronary angioplasty), or death due to CHD. CVD included CHD and ischemic stroke events. ACS was defined as MI, unstable angina, or death due to CHD. In the Corogene study, CHD was defined as coronary artery obstruction > 50% in at least one coronary artery.

Cardiovascular diagnoses for MDC-CC and MPP were obtained from the Swedish Hospital Discharge Register, the Swedish Cause-of-Death Register, and the Stroke Register of Malmö (Jerntorp and Berglund 1992). Diagnoses were recorded using ICD-9 and ICD-10 codes. Cardiovascular events included MI, ischemic stroke, and death due to CHD. Ischemic stroke events in the Swedish Hospital Discharge Register have been validated by using the Stroke Register of Malmö. In this study, the follow-up ended on Dec 31, 2006. ICD-9 and ICD-10 codes for event diagnoses in Finnish and Swedish cohorts are reported in **Table 6**.

Table 6. Cardiovascular event definitions in Finnish (FIN) and Swedish (SWE) cohorts.

	ICD-9 codes		ICD-10 codes	
	FIN*	**SWE****	**FIN***	**SWE****
Myocardial infarction	410	410	I21–I22	I21
Acute coronary syndrome	410, 798, except 7980A	NA	I20–I22, I46, R96, R98	NA
Coronary heart disease†	410–414, 798, except 7980A	NA	I20–I25, I46, R96, R98	NA
Stroke	430–438	434, 436	I60–I69, G45	I63–I64
Cardiovascular disease	MI, CHD or stroke	412, 414	MI, CHD or stroke	I22–I23, I25

ICD-9
410: Acute myocardial infarction
411: Other acute and subacute forms of ischemic heart disease
412: Old myocardial infarction
413: Angina pectoris
414: Other forms of chronic ischemic heart disease
430: Subarachnoid hemorrhage
431–432: Hemorrhagic strokes
433–438: Ischemic strokes or other cerebrovascular diseases
798: Sudden death, cause unknown; 7980A: Sudden infant death syndrome

ICD-10
I20: Unstable angina
I21: Acute myocardial infarction
I22: Subsequent myocardial infarction
I23: Certain current complications following acute myocardial infarction
I24: Other acute ischemic heart diseases
I25: Chronic ischemic heart disease
I60–I62: Hemorrhagic strokes
I63–I64, G45: Ischemic strokes
I65–I69: Other cerebrovascular diseases
I46: Cardiac arrest
R96: Other sudden death, cause unknown
R98: Unattended death

* FINRISK 1992, 1997, 2002, and Health 2000; ** MDC-CC; † The definition of CHD includes all myocardial infarction and acute coronary syndrome cases. Abbreviations: ICD, International Classification of Disease. Data adapted from Laatikainen et al. (2009) and the World Health Organization (2012b).

4.3.2 Lipids and blood pressure

In the Young Finns cohort, clinical measurements were taken at baseline in 1980 (participants aged 3-18 years) and during four follow-up examinations between 1983 and 2007. Serum lipid concentrations (LDL, HDL, total cholesterol, and triglycerides) were

assayed using standard enzymatic laboratory methods in the Research and Development Unit of the Social Insurance Institution, Turku. Variation in lipid levels in different study years due to change in measurement device was corrected by using correction factor equations (Juonala et al. 2004; Raiko et al. 2010).

Systolic and diastolic blood pressure was measured with a standard mercury gravity sphygmomanometer in 1980 and 1983, and with a random-zero sphygmomanometer in 1986, 2001, and 2007. For 3-year-olds, only systolic blood pressure was measured using an ultrasound device. Final blood pressure value was determined as an average value of three measurements taken between 8 and 10 a.m. from the right arm. The first measurement was taken from fasting subjects after 5 min of rest in the sitting position, and subsequent measurements after 2-3 min of rest.

4.3.2.1 Definition of dyslipidemia and hypertension

Dyslipidemia status was determined for 1204 subjects who had lipid measurements taken at the ages of 9 and 30–33 years. We defined lipid-specific dyslipidemias by using thresholds for cases as HDL < 1 mmol/l, LDL > 3 mmol/l, and triglycerides > 2 mmol/l at 30 or 33 years old. Additionally, we determined combined dyslipidemia status, in which an individual was defined as a case if any of the thresholds were exceeded.

Hypertension status was assigned for individuals over 18 years of age in 2007. It was defined based on the following criteria: 1) systolic blood pressure ≥ 140 mmHg, 2) diastolic blood pressure ≥ 90 mmHg, or 3) use of antihypertensive medication.

4.4 Statistical methods

4.4.1 Quality checks and exclusions

In Project I, we excluded nonfasted individuals (1.3%) and subjects with lipid medication (2.1%). Normality of the phenotype distributions was visually inspected and outliers removed. Triglyceride measurements were transformed using natural logarithm. In Project II, we used a substitution method for those subjects with antihypertensive medication in 2001 (N=58) and 2007 (N=143); 15 mmHg was added to the systolic and 10 mmHg to the diastolic blood pressure value. This method has been shown to be more effective than adjustment for a binary treatment covariate or excluding individuals with antihypertensive therapy (Tobin et al. 2005). In Projects III and IV, we excluded individuals with prevalent CVD and those ≥ 75 years of age at baseline. During the follow-up individuals turning 80 were censored on their 80th birthday.

4.4.2 Longitudinal and age-specific genetic effects for lipids and blood pressure

In the Young Finns study, each subject has been measured repeatedly at several time-points during their life course. The study design enables estimation of longitudinal effects, although it requires analysis methods that take into account the correlation between the repeated measurements for the same individual. We estimated longitudinal genetic effects on lipids and blood pressure with linear mixed effects models. Mixed effects model consists of two parts: fixed effects (β) and subject-specific random effects (u):

$$y = X\beta + Zu + \varepsilon$$

Here y is the response vector, ε is the vector of random errors, X is the model matrix for fixed effects, and Z is the model matrix for random effects. The random effects define the within-individual covariance structure of the observations. The choice of the covariance structure depends on the time intervals between measurements; the correlation is assumed to be higher for measurements taken closer together in time than those taken far apart in time (Littell et al. 2000).

In Project I, we performed the analyses separately for males and females. We tested each SNP individually and as in weighted multi-locus GRSs, which were standardized (mean=0, SD=1). We estimated genetic effects on lipid levels in seven age groups: 3–6, 9, 12, 15, 18, 21–30, and 33–45 years. Due to low number of 3-year-old participants, we combined them with 6-year-olds. Since the main focus of this project was to study the genetic effects in children and adolescents, we combined adult age groups so that they would provide a reasonable reference with which to compare the results of the younger age groups. For each lipid trait, the change in the GRS effects over age was tested with age × GRS interaction analysis with the restricted maximum likelihood (REML) method with fixed covariates of age group and birth year and a random normally distributed intercept.

The proportions of variance explained by genetic variants were calculated by linear models for each lipid trait as the difference in R^2 of the models with and without the genetic data. Models were constructed for each age and sex group separately.

In Project II, the longitudinal effects of SNPs and standardized GRSs for systolic and diastolic blood pressure were estimated with mixed models adjusted for age, sex, BMI, and study year. The models for systolic blood pressure were also adjusted for age squared. The analyses were replicated in the Bogalusa Heart Study (N=1194), which has been described elsewhere (Smith et al. 2010). As a secondary analysis, we performed a cross-sectional analysis for the baseline and 2007 data. Finally, we studied the modifying effects of sex and dietary salt intake on genetic effects. Salt intake was measured with a 13-item food questionnaire.

In Projects I and II, the models were optimized (e.g. the choice of the covariance structure) based on the Akaike information criterion (Akaike 1974). The model assumptions were checked by graphical inspection of residuals.

4.4.3 Prediction of dyslipidemia and hypertension

In Projects I and II, we studied how well the GRSs predict dyslipidemia and hypertension in early adulthood (see definition of dyslipidemia and hypertension in Section 4.3.2.1). In Project I, we restricted the data to birth cohorts of 1971, 1974, and 1977 that had lipid measurements performed at the ages of 9 and 30–33 years. Two models were fitted with logistic regression: 1) age, birth year, sex, and a lipid measurement at 9 years old as predictors; and 2) in addition, the lipid related GRS as a predictor. In Project 2, we also compared two models with and without the 13-SNP GRS in the cross-sectional 2007 cohort. Models were adjusted for age, sex, BMI, and family history of hypertension. To further quantify the genetic effects, the blood pressure GRS was divided into quintiles and the effect estimates for the highest and lowest groups were compared. In both projects, we estimated model discrimination with the ROC curves and calculated the change in the area under the ROC curve (C-index) between the models with and without the GRS.

4.4.4 Associations between genetic variants and cardiovascular events

In Projects III and IV, we studied associations between the selected CHD SNPs and incident cardiovascular events with Cox proportional hazards models:

$$h(t \mid X) = h_0(t)exp(\beta^T X)$$

The hazard at time t depends on the covariates X and the baseline hazard $h_0(t)$, which is an unknown and unspecified nonnegative function. The effects of covariates (β) are assumed to be constant over time. Models were adjusted for traditional risk factors, and age was used as the time scale. We studied each SNP individually in separate models and jointly as a GRS in one model.

In Project III, we categorized the GRS into five equal-sized groups based on quintiles. We estimated the effect sizes with 95% confidence intervals for each group by using the lowest group as a reference, and tested the null hypothesis of no linear effect over the quintiles. In Project IV, we estimated the effect for standardized (mean=0, SD=1) continuous GRS. To further quantify genetic effects for the subjects with high GRS, we divided the GRS into ten groups according to deciles of the GRS distribution. The effect estimates for the highest group were compared with the middle 20% of the GRS.

Model diagnostics were inspected in each cohort separately with scaled Schoenfeld residuals, deviance residuals, and martingale residuals. Residuals were plotted against time and examined visually. In addition, a statistical test was performed for the

proportional hazards assumption. Scaled Schoenfeld residuals are used to test the assumption of proportional hazards, which is that the effect of a covariate should not change with time. Deviance residuals are used to check the case influence, which means that the data for some subjects may contradict the model prediction. Martingale residuals can be plotted against covariates to detect nonlinearity.

Associations between the genetic variants and prevalent cardiovascular events were studied with logistic regression analysis adjusted for age and sex. The Corogene data were analyzed with conditional logistic regression.

Information on family history was available in FINRISK studies (N=19,001). We studied the relationship between the GRS and family history of CVD by further adjusting the models for binary family history indicator (0 = no family history of CVD; 1 = family history of CVD). We also studied the effects of family history with and without the adjustment for the GRS.

Each cohort was analyzed separately and fixed effect meta-analysis was used to combine the results from incident and prevalent data analyses.

4.4.5 Improvement in prediction

We evaluated improvement in prediction when genetic data were added to the model with the other risk factors. Discrimination, reclassification, and calibration indices were calculated by comparing the predicted probabilities (p) from the models with and without the GRS. Risk discrimination was estimated with C-index and the statistical significance of the change in C-index was tested with the correlated C-index approach (Antolini et al. 2004). We estimated IDI, which is the difference in mean predicted probabilities ($\bar{p}$) between cases and noncases of two models (Pencina et al. 2008):

$$IDI = (\bar{p}_{cases} - \bar{p}_{noncases})_{new\ model} - (\bar{p}_{cases} - \bar{p}_{noncases})_{old\ model}$$

Standard errors for IDI were calculated as paired differences in predicted probabilities across all cases and noncases. Risk reclassification was evaluated with NRI (Pencina et al. 2008), which is defined as

$$NRI = \left(p_{up,cases} - p_{down,cases}\right) - \left(p_{up,noncases} - p_{down,noncases}\right)$$

We tested risk reclassification with NRI and Clinical NRI using FINRISK 1992 and 1997 cohorts. We modeled risk reclassification jointly in these two datasets with restricted follow-up of FINRISK 1992 and adjusted the analysis with the cohort indicator. We calculated NRI, which takes censored observations into account, by using the Kaplan-Meier approach with bootstrap-based confidence interval estimation (Steyerberg and Pencina 2010; Pencina et al. 2011). NRI was estimated using four risk categories (0–5%,

5–10%, 10–20%, and > 20%). Clinical NRI was calculated for the subjects classified to the intermediate-risk group (10–20%) by the model without genetic data.

In Project IV, we performed external validation for discrimination and reclassification results. The effects were estimated in an independent training data set (FINRISK 2002) and applied as weights in the test data set (joint FINRISK 1992 and 1997 data).

To evaluate the effect of improved risk reclassification at a population level, we generalized our NRI results for 100,000 subjects with age and gender structure similar to the standard European population. The data was divided into eight groups based on gender and age (age was categorized into four groups: 40–50, 50–60, 60–70, ≥ 70), and reclassification tables were calculated for each of these groups separately. Then, reclassification tables were weighted with the estimated incidence rates multiplied by the group-specific counts based on the standard European population. Assuming that age- and sex-specific incidences of CHD in the European standard population are comparable with the current study, we estimated incidence rates from the FINRISK 1992 and 1997.

In a two-stage screening strategy, we assumed that all participants were first classified into cardiovascular risk categories based on traditional risk factors, and then additional GRS screening was targeted to those in the intermediate-risk category (10–20%). Subjects at intermediate risk were considered a clinically relevant subgroup based on the following assumptions: 1) statin medication is allocated to subjects in the high-risk category (≥ 20%) and subjects with diabetes, and 2) statins reduce cardiovascular risk by 20% in subjects without prevalent CVD (Baigent et al. 2005).

Calibration of the models was tested with the Hosmer-Lemeshow goodness-of-fit test (Hosmer and Lemeshow 1999).

5 RESULTS

5.1 Background characteristics

Background characteristics of the study cohorts are presented in **Table 7**. From the Young Finns Study, 2443 subjects were included in Project I and 2357 subjects in Project II. The mean age of Young Finns Study participants at the latest follow-up examination in 2007 was 37.6 (SD=5.0) years. In total, 55,123 subjects from seven cohorts (FINRISK 1992–2002, Health 2000, Corogene, MDC-CC, and MPP) had information on prevalent events and risk factors. After exclusions, 3829 CHD cases (7%) and 48,897 controls were included in cross-sectional analysis in Project III. Incident data were available in the FINRISK 1992–2002, Health 2000, and MDC-CC cohorts (total N=34,224). Altogether 30,725 participants were included in prospective cohort analysis in Project III. During the median follow-up of 10.7 years (IQR 6.7–13.6), 1264 incident CHD cases (4%) occurred. Only Finnish prospective cohorts (N=29,120) were included in Project IV. After exclusions, 24,124 subjects were included to the study. The median follow-up time was 12 years (IQR 8.8–15.3 years), and 1093 incident CHD events (5%) were observed.

Table 7. Background characteristics of study cohorts.

	FR-92	FR-97	FR-02	Health 2000	Coro-gene	YFS*	MDC-CC	MPP
N	6024	8388	8616	6092	6015	2443	5104	14884
N Males	2844	4225	3996	2738	4192	1123	2141	9773
N Females	3180	4163	4620	3354	1823	1320	2963	5111
Age	44.8 (11.4)	48.6 (13.5)	47.9 (13.2)	54.5 (15.4)	66.5 (10.9)	37.6 (5.0)	57.4 (5.9)	45.3 (7.0)
TC mmol/l	5.6 (1.1)	5.5 (1.1)	5.6 (1.1)	5.9 (1.1)	NA	5.0 (0.9)	6.2 (1.1)	5.6 (1.0)
LDL mmol/l	3.6 (1.0)	3.5 (0.9)	3.4 (0.9)	3.8 (1.2)	NA	3.1 (0.8)	4.2 (1.0)	NA
HDL mmol/l	1.4 (0.3)	1.4 (0.4)	1.5 (0.4)	1.3 (0.4)	NA	1.3 (0.3)	1.4 (0.4)	NA
Log(TG)	0.3 (0.6)	0.3 (0.5)	0.2 (0.5)	0.3 (0.5)	NA	0.2 (0.5)	NA	NA
SBP mmHg	135.9 (19.5)	136.2 (20.1)	134.9 (20.0)	135.5 (21.5)	NA	125.2 (14.2)	141.1 (18.9)	127.7 (14.4)
DBP mmHg	81.5 (11.9)	82.4 (11.3)	78.9 (11.4)	81.7 (11.4)	NA	79.1 (11.6)	86.9 (9.5)	84.2 (8.8)
Current smoker N (%)	1664 (27.6%)	1936 (23.1%)	2230 (25.9%)	1688 (27.7%)	NA	455 (10.0%)	1307 (26%)	5639 (38%)
Prevalent CVD N (%)	138 (2.3%)	315 (3.8%)	301 (3.5%)	327 (5.4%)	2101 (35%)	NA	107 (2%)	1749 (12%)
Incident CVD N (%)	642 (10.7%)	822 (9.8%)	343 (4.0%)	458 (7.5%)	NA	NA	468 (9%)	NA
Incident CHD N (%)	472 (7.8%)	592 (7.1%)	252 (2.9%)	352 (5.8%)	NA	NA	NA	NA
Incident ACS N (%)	365 (6.1%)	481 (5.7%)	203 (2.4%)	289 (4.7%)	NA	NA	NA	NA
Incident MI N (%)	259 (4.3%)	327 (3.9%)	146 (1.7%)	208 (3.4%)	NA	NA	266 (2%)	NA

* Values are from the 2007 follow-up of YFS. Data are mean (SD) or number (%). Abbreviations: FR, FINRISK; YFS, Young Finns Study; MDC-CC, Malmö Diet and Cancer Study - Cardiovascular Cohort; MPP, Malmö Preventive Project; N, number; TC, total cholesterol; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; TG, triglycerides; SBP, systolic blood pressure; DBP, diastolic blood pressure; CVD, cardiovascular disease; CHD, coronary heart disease; ACS, acute coronary syndrome; MI, myocardial infarction; NA, information not available

5.2 Longitudinal trends in lipid and blood pressure levels

Longitudinal mean levels for lipids and blood pressure are shown in **Figure 10**. We observed an increasing trend with age in total cholesterol, LDL, triglycerides, and blood pressure. The increase in risk factor levels starts at the age of 10–15 years and is stronger in males than in females. Especially triglyceride levels show a constant increase in males, but remain relatively stable in females after 30 years of age. The decreasing trend in total cholesterol and LDL levels in the early 1980s has been previously reported (Viikari et al. 2006) and might reflect the change in mean cholesterol levels in the Finnish population. Cholesterol levels have decreased substantially in Finland since the 1970s, when Finnish children had mean cholesterol levels ≥ 6 mmol/l (Räsänen et al. 1978). HDL levels decreased in males after 10 years of age, but in females HDL was less variable with age. Overall, males have more unfavorable trends in lipid profiles and blood pressure after puberty onset than females, which could partly explain the difference in adulthood CHD incidence between genders. The sex difference is at least partly attributable to changes in sex hormone levels after puberty (Garces et al. 2010).

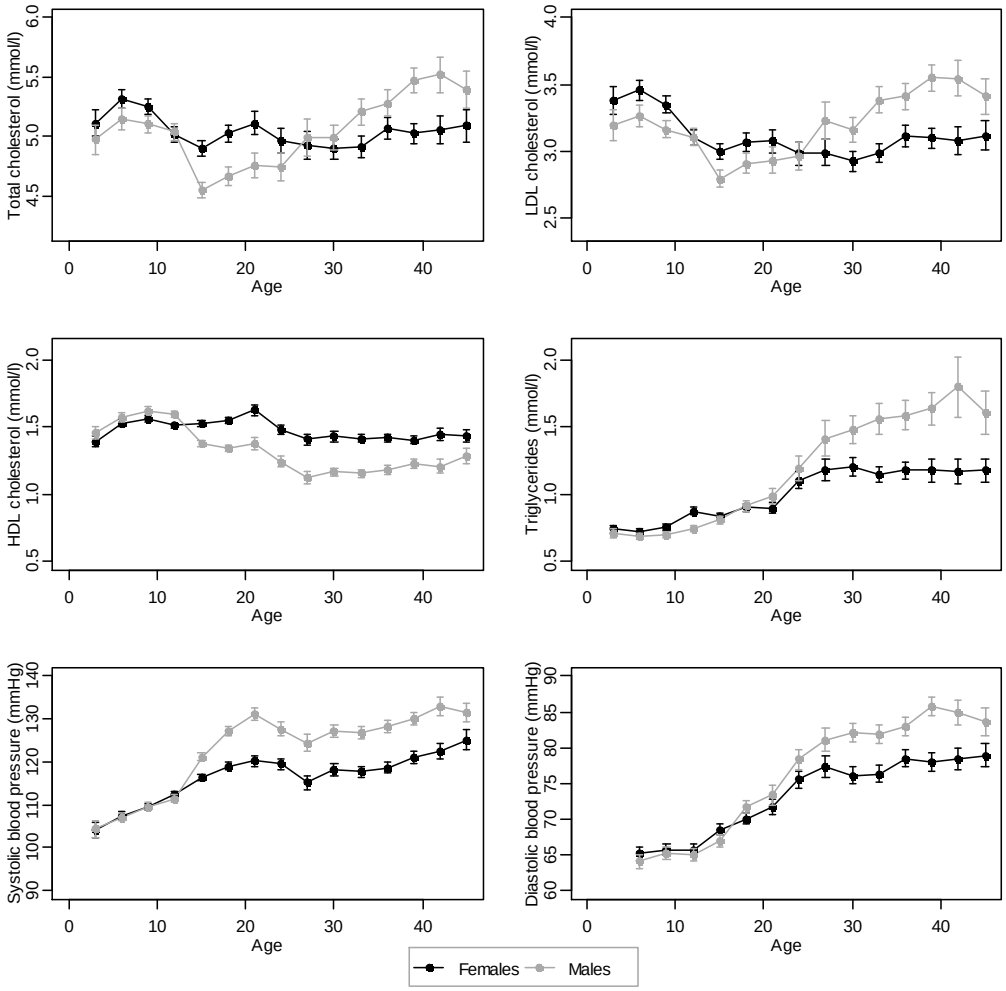


Figure 10. Mean levels and 95 % confidence intervals for lipids and blood pressure by age and gender.

5.3 Genetic effects on lipids and blood pressure

5.3.1 Associations of individual SNPs with lipids and blood pressure

Longitudinal associations for the ten strongest SNPs reported by Teslovich et al. (2010) are shown for each lipid trait in **Figure 11**. The top SNPs were longitudinally associated with lipids in our data. Overall, 30 of 47 HDL, 30 of 37 LDL , 33 of 52 total cholesterol and 21 of 32 triglycerides loci were associated with the corresponding lipid trait between the ages of 3 and 15 years at a significance level of $P<0.05$. Thus, the results from GWASs performed using adult samples seem to generalize to younger age groups.

Of 13 blood pressure SNPs, three were longitudinally associated with systolic and/or diastolic blood pressure ($P<0.05$). Of these, SNPs rs16948048 (in or near *ZNF652*) and rs11191548 (in or near *CYP17A1*) were associated only with diastolic blood pressure, whereas rs11014166 (in or near *CACNB2*) was associated with both blood pressure traits. We also found a sex-specific effect for rs11191548; the sex × SNP interaction for systolic blood pressure was significant ($P=0.005$), and in sex-stratified analyses, the SNP effect was significant in females ($P=0.003$), but not in males ($P=0.13$). A similar result has been reported in Chinese children (Wu et al. 2012). We did not observe significant interactions between the SNPs and dietary salt intake.

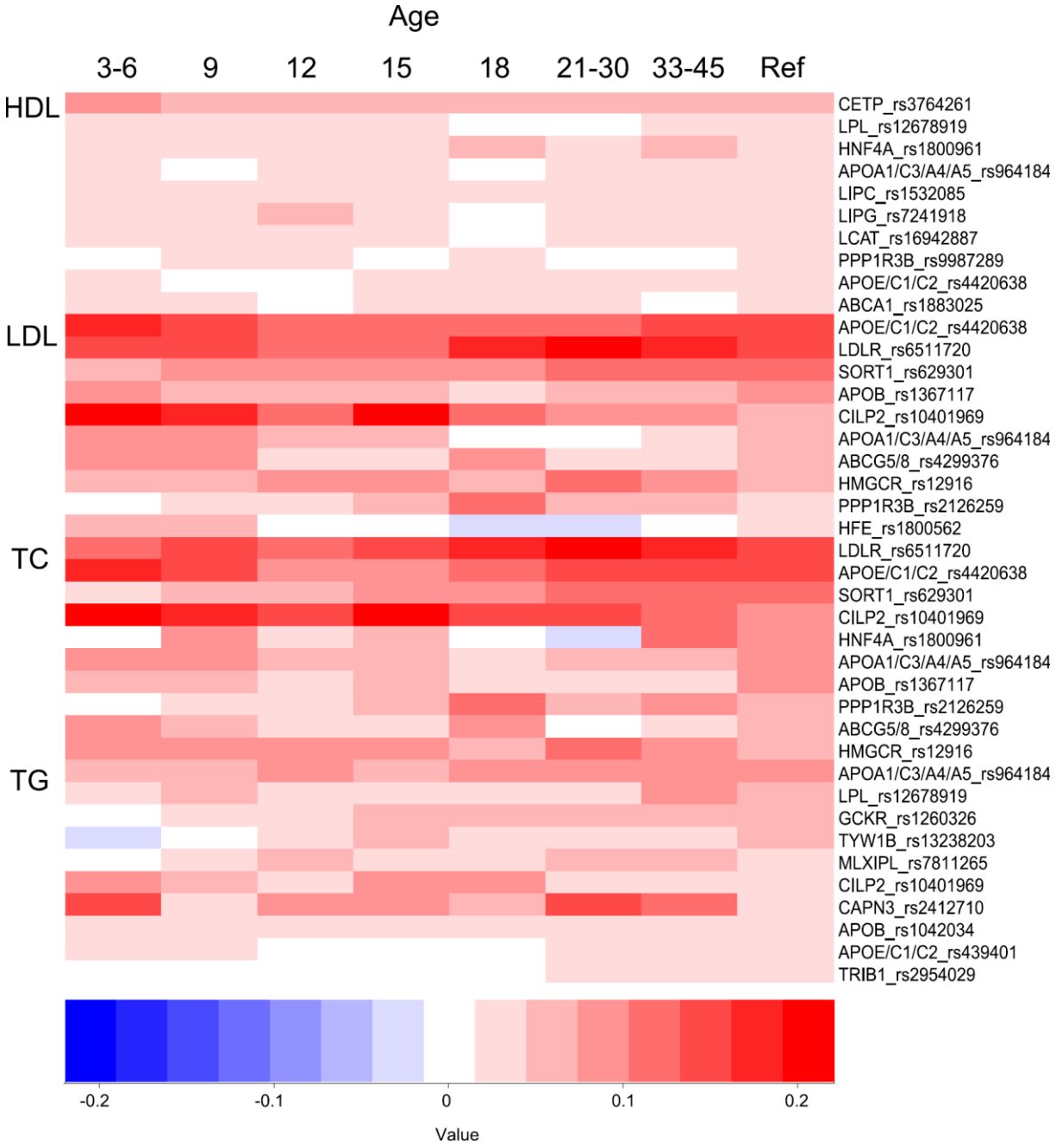


Figure 11. Longitudinal effects for the top 10 SNPs in the reference study (Teslovich et al. 2010). Effect sizes have been flipped to illustrate the increasing effect in the reference study. Colors correspond to the effect sizes in mmol/l or natural log-transformed mmol/l (TG). The gene name is a plausible biological candidate gene in the locus or the nearest annotated gene to the SNP. Abbreviations: TC, total cholesterol; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol TG, triglycerides; Ref, reference study.

5.3.2 Association of GRSs with lipids and blood pressure

The GRSs were statistically associated with lipids in all age and sex groups ($P<8.4\times10^{-5}$). Overall, the genetic effects for LDL and total cholesterol were higher than for HDL and triglycerides, and they were stable across all age groups. The childhood effect size (3- to 6-year-olds) for LDL (β=0.22, SE=0.03 in both genders) was comparable with the effect size in adulthood (β=0.25, SE=0.03 for males; β=0.22, SE=0.02 for females). The results for total cholesterol were similar.

Genetic effects for HDL cholesterol and triglycerides vary by age group (**Figure 12**). In males, the genetic effects for HDL were larger in 3- to 6-year-old children than in adults (β=0.11, SE=0.01 vs. β=0.08, SE=0.01), and the interaction between HDL-GRS and age group was significant (P=0.02). Females had a similar decreasing trend, but the GRS × age interaction was not statistically significant. By contrast, there was an increasing trend with age in GRS effects on triglycerides (P=0.0001 in males, P=0.009 in females). The effect of triglycerides-GRS was more than twofold in adult males compared with 3- 6-year-olds (β=0.05, SE=0.01 vs. β=0.14, SE=0.02). In females, the genetic effect was highest in the age group 21–30 years (β=0.12, SE=0.01) and lowest in 3- to 6-year-olds (β=0.06, SE=0.01).

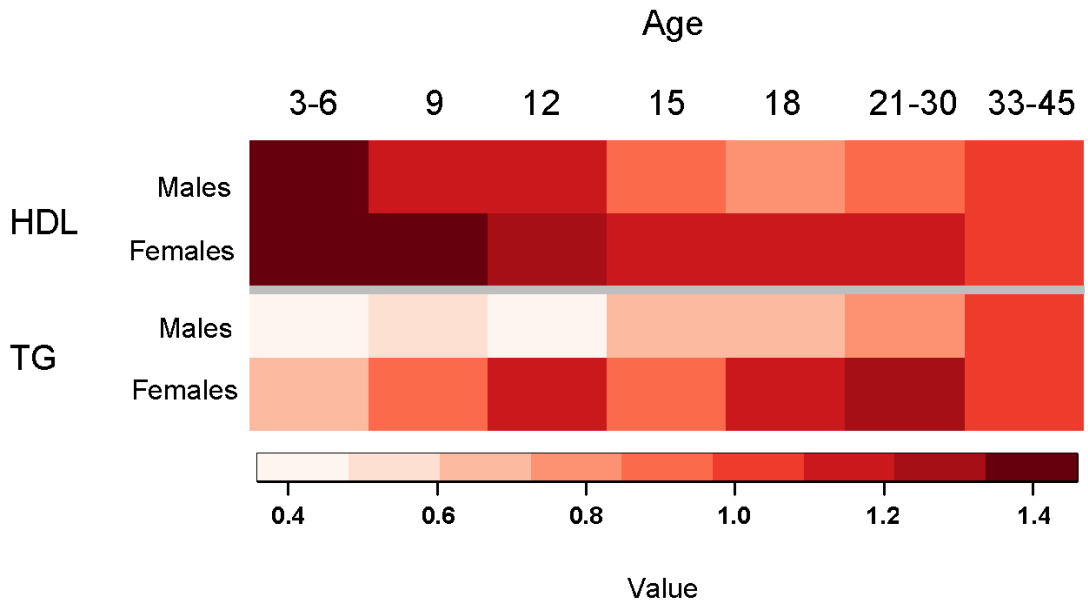


Figure 12. Different patterns of association by age for HDL cholesterol and triglycerides (TG). The genetic effect sizes decrease for HDL and increase for TG with age. Colors in different age groups correspond to the effect sizes proportional to adulthood effect size. P-values in each cell < 8.4×10^{-5}.

Jointly, the lipid loci explained 11.8–26.7% of the total variance in lipids among 3- to 6-year-olds. The proportion diminished with age for HDL, LDL, and total cholesterol. Proportion of variance explained by the SNPs in 21- to 45-year-olds was around 10%, which corresponds to the estimates reported for adults earlier (Teslovich et al. 2010). Estimates for triglycerides fluctuated between 7.4% and 14.8% and did not show a similar decreasing trend.

Blood pressure GRS was associated with longitudinal systolic (β=0.47, SE=0.18, P=0.008) and diastolic (β=0.53, SE=0.14, P=0.0003) blood pressure. These results were replicated in the African-American population of the Bogalusa Heart Study (both P≤0.005), but not in the European population of the Bogalusa Heart Study. The effects were smaller in cross-sectional analysis at baseline in 1980 (systolic blood pressure: β=0.11, SE=0.21, P=0.61; diastolic blood pressure: β=0.35, SE=0.21, P=0.10) but weaker than in 2007 (systolic blood pressure: β=1.0, SE=0.31, P=0.0006; diastolic blood pressure: β=0.76, SE=0.25, P=0.002) (**Table 8**). Sex or dietary salt intake did not modify the effect of the GRS (P-values for sex × GRS and salt intake × GRS nonsignificant).

Table 8. Effects of the 13-SNP GRS on blood pressure traits.

Longitudinal effect* from 1980 to 2007		Cross-sectional effect* at baseline in 1980		Cross-sectional effect* at 2007		Odds ratio** for hypertension in adulthood
SBP	**DBP**	**SBP**	**DBP**	**SBP**	**DBP**	
β=0.47, SE=0.18, P=0.008	β=0.53, SE=0.14, P=0.0003	β=0.11, SE=0.21 P=0.61	β=0.35, SE=0.21, P=0.10	β=1.0, SE=0.31, P=0.0006	β=0.76, SE=0.25, P= 0.002	OR=1.82, CI=1.53–2.17, P<0.0001

* per SD of GRS. ** Between the extreme quintiles of the GRS. Longitudinal models adjusted for measurement year, age (and age squared for systolic blood pressure), sex, and BMI; cross-sectional and logistic models adjusted for age, sex, BMI, and family history. Abbreviations: GRS, genetic risk score; SNP, single nucleotide polymorphism; SBP, systolic blood pressure; DBP, diastolic blood pressure; OR, odds ratio; SE, standard error; CI, confidence interval

5.3.3 Prediction of adulthood dyslipidemia and hypertension

More than half (54.9%) of the 30- to 33-year-old subjects fulfilled the criteria for dyslipidemia. For lipid-specific dyslipidemias, most of the cases were due to high LDL cholesterol (45.4%). HDL dyslipidemia was defined for 17.3% and hypertriglyceridemia for 12.7% of participants. The basic model including age, birth year, sex, and a lipid measurement at 9 years of age had a C-index of 0.81, 0.77, and 0.67 for HDL, LDL and triglycerides, respectively. GRS comprising SNPs associated with triglycerides improved risk discrimination of hypertriglyceridemia in young adults when added to childhood measurement of triglycerides (change in C-index: 0.04, P=0.01). The change in estimates

was nonsignificant for HDL (change in C-index: 0.01, P=0.14) and LDL (change in C-index: 0.005, P=0.54) dyslipidemias. IDI was significant for all comparisons (HDL: IDI=0.018, 95% CI 0.006–0.030, P=0.003; LDL: IDI=0.008, 95% CI 0.002–0.014, P=0.01; triglycerides: 0.041, 95% CI 0.024–0.059, $P=2.6\times10^{-6}$). Only LDL-GRS was associated with combined dyslipidemia status (P=0.02), but did not improve C-index over clinical lipid measurements. The calibration was good for models with ($0.26<P<0.58$) and without ($0.13<P<0.80$) GRSs.

The proportion of hypertensive subjects in childhood did not differ according to highest and lowest 20% of the 13-SNP blood pressure GRS. In adulthood (age 45 years), 20% of individuals in the lowest and 25% of individuals in the highest GRS were hypertensive. Blood pressure GRS increased the risk for hypertension in adulthood, when adjusted for age, sex, BMI, and family history of premature hypertension. Individuals in the highest 20% of GRS had a 1.8-fold (95% CI 1.53–2.17; P<0.0001) risk for hypertension, compared with individuals in the lowest 20%. However, addition of the GRS to the model did not improve risk discrimination (C-index: 0.72 vs. 0.71). In comparison, a recent study (Lieb et al. 2013) reported 59–70% higher relative risks for CHD in individuals in the top 20% of blood pressure GRS.

5.4 Traditional risk factors for cardiovascular events

The effects of traditional risk factors on incident CHD are shown in **Table 9**. The strongest predictors were gender, smoking and total cholesterol. Family history of premature CVD was significantly associated with CHD (HR=1.46, 95% CI 1.27–1.67), CVD (HR=1.37, 95% CI 1.22–1.53), ACS (HR=1.45, 95% CI 1.24–1.71), and MI (HR=1.31, 95% CI 1.09–1.58).

Table 9. Effects of traditional risk factors on incident CHD in the FINRISK 1992, 1997, and 2002 studies.*

	HR (95% CI)
Sex (1=male, 2=female)	0.42 (0.36,0.49)
Systolic blood pressure	1.18 (1.11,1.26)
Blood pressure treatment	1.40 (1.19,1.63)
Total cholesterol	1.29 (1.22,1.37)
HDL cholesterol	0.72 (0.66,0.79)
Type 2 diabetes	2.05 (1.69,2.50)
Current smoking	2.08 (1.80,2.40)
Body mass index	1.11 (1.03,1.19)
Family history of CVD	1.46 (1.27,1.67)

* Estimated with Cox proportional hazards model by using age as the time scale. Continuous variables were standardized (mean=0, SD=1). Abbreviations: HR, hazard ratio; CI, confidence interval; CHD, coronary heart disease; CVD, cardiovascular disease; HDL, high-density lipoprotein.

5.5 Genetic effects on cardiovascular events

5.5.1 Association of individual SNPs with cardiovascular events

In total, 13 of 28 CHD SNPs in or near *WDR12*, *PHACTR1*, *LPA*, *CDKN2A*/*B*, *CXCL12*, *MRPS6*, *ABO*, *COL4A1/2*, *HHIPL1*, *ADAMTS7*, *RASD1*/*SMCR3*/*PEMT*, *SMG6*/*SRR*, and *ADAMTS7*/*MORF4L1* were associated with at least one of the cardiovascular end-points. SNPs rs6725887 (*WDR12*), rs12526453 (*PHACTR1*), rs4977574 (*CDKN2A/B, ANRIL*), rs1746048 (*CXCL12*), and rs3825807 (*ADAMTS7*) were associated with all three end-points. The strongest effect size was detected for rare variant (MAF 1% in the study population) rs3798220 in the *LPA* locus (CHD: HR 1.43, 95% CI 1.02–2.00, P=0.04). The SNP rs579459 (*ABO*) was significantly associated only with ACS (HR 1.13, 95% CI 1.00–1.27, P=0.05), suggesting that its functional role may be specific to acute events. The variant has previously been shown to be associated with adverse cardiac outcome after ACS (Wauters et al. 2012).

5.5.2 Association of GRSs with cardiovascular events

GRS comprising 13 CHD SNPs increased the risk of incident CHD linearly over and above traditional risk factors (Project III). Subjects in the highest 20% of the GRS distribution were estimated to have a 1.66-fold increased risk of CHD compared with those in the lowest 20% (95% CI 1.35–2.04, P-value for linear trend across

quintiles=7.3×10^{-10}, **Figure 13**, left). The magnitude of the estimated effect for the 13-SNP GRS was comparable with other established risk factors such as systolic blood pressure (HR=1.66, 95% CI 1.19–2.30, for top vs. bottom quintile of systolic blood pressure in FINRISK studies). The 13-SNP GRS also increased the risk for other end-points, although the effect sizes for the highest versus lowest groups were slightly lower; the hazard ratio for CVD was 1.50 (95% CI 1.29–1.75, $P=1.9\times10^{-10}$) and for MI 1.46 (95% CI 1.15–1.86, $P=2.8\times10^{-5}$).

The genetic effects were comparable in the case-control analysis; the odds ratio for the highest versus lowest group was 1.63 (95% CI 1.24–2.15, P-value for linear trend across quintiles=4.8×10^{-5}, **Figure 13**, right) for prevalent CHD, 1.30 (95% CI 1.15–1.47, $P=2.6\times10^{-8}$) for CVD, and 1.56 (95% CI 1.38–1.76, $P=1.2\times10^{-15}$) for MI.

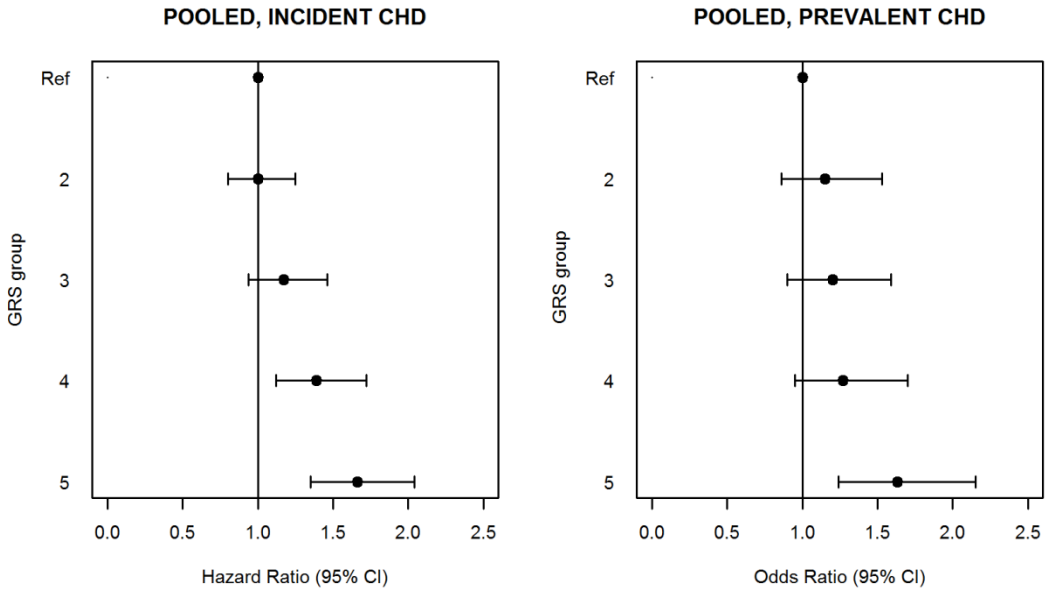


Figure 13. Pooled effects in GRS groups. The Cox proportional hazards models (incident CHD) and logistic regression models (prevalent CHD) were fitted in each cohort separately and the estimates were pooled with fixed effects meta-analysis. Abbreviations: CHD, coronary heart disease; GRS, genetic risk score; CI, confidence interval, Ref, reference group

The GRS effects were not dominated by the strongest, common CHD/MI locus reported to date, rs4977574 at 9p21 near *CDKN2B-CDKN2A*. After adjusting for rs4977574, the hazard ratio for the highest versus lowest GRS quintile was 1.51 (95% CI 1.19–1.91) for CHD, 1.40 (95% CI 1.18–1.67) for CVD, and 1.30 (95% CI 1.00–1.71) for MI. Thus, other variants in the GRS were associated with cardiovascular end-points beyond the 9p21 locus.

In Project IV, we extended the 13-SNP GRS with 15 novel CHD SNPs. We observed that the effect estimates for 28-SNP GRS were consistently higher than for 13-SNP GRS (**Table 10**). With the aim of better identifying individuals with large genetic risk load, we estimated genetic effects for the extended GRS (including 28 SNPs) divided into deciles instead of quintiles, and compared the effects between the highest 10% and average 20% of GRS. Individuals in the top 10% of 28-SNP GRS had a 2.1-fold risk (95% CI 1.68–2.56) for CHD relative to the middle 20%. In this data, the corresponding effect estimate for 13-SNP GRS was 1.6 (95% CI 1.26–1.91).

Table 10. Hazard ratios for cardiovascular events by GRSs.

Trait	HR (95% CI)*		Top 10% versus middle 20% of GRS (95% CI)	
	13-SNP GRS	28-SNP GRS	13-SNP GRS	28-SNP GRS
CHD	1.23 (1.16–1.30)	1.27 (1.20–1.35)	1.55 (1.26–1.91)	2.07 (1.68–2.56)
ACS	1.24 (1.15–1.34)	1.27 (1.18–1.37)	1.49 (1.14–1.95)	1.84 (1.42–2.40)
CVD	1.17 (1.11–1.23)	1.18 (1.12–1.24)	1.47 (1.22–1.76)	1.87 (1.56–2.24)

* per SD of GRS. Cox regression models were adjusted for sex, total cholesterol, high-density lipoprotein (HDL) cholesterol, body mass index, systolic blood pressure, antihypertensive treatment, smoking, and type 2 diabetes; age was used as the timescale. Abbreviations: CHD, coronary heart disease; ACS, acute coronary syndrome; CVD, cardiovascular disease; SNP, single-nucleotide polymorphism; GRS, genetic risk score; HR, hazard ratio; CI, confidence interval

We used three FINRISK studies (N=19,001) to examine the relationship between the family history of CVD and the GRS. We evaluated three models for CHD, CVD, and ACS with traditional risk factors and 1) binary family history indicator, 2) the 28-SNP GRS, and 3) both family history and the GRS. Even though family history was an independent risk factor for cardiovascular events by raising the risk by 37–46%, the effect of the GRS changed only marginally after adjusting for family history (**Figure 14**, left). Likewise, the effects of family history slightly diminished when the GRS was included to the model (**Figure 14**, right). These results indicate that family history and genetic data have independent effects on CHD and support the proposal (Do et al. 2012) to use both types of information in prediction models to optimize prediction accuracy.

Based on scaled Schoenfeld residuals, deviance residuals, and martingale residuals, we found no systematic evidence of nonproportionality, case influence, or nonlinearity of the data.

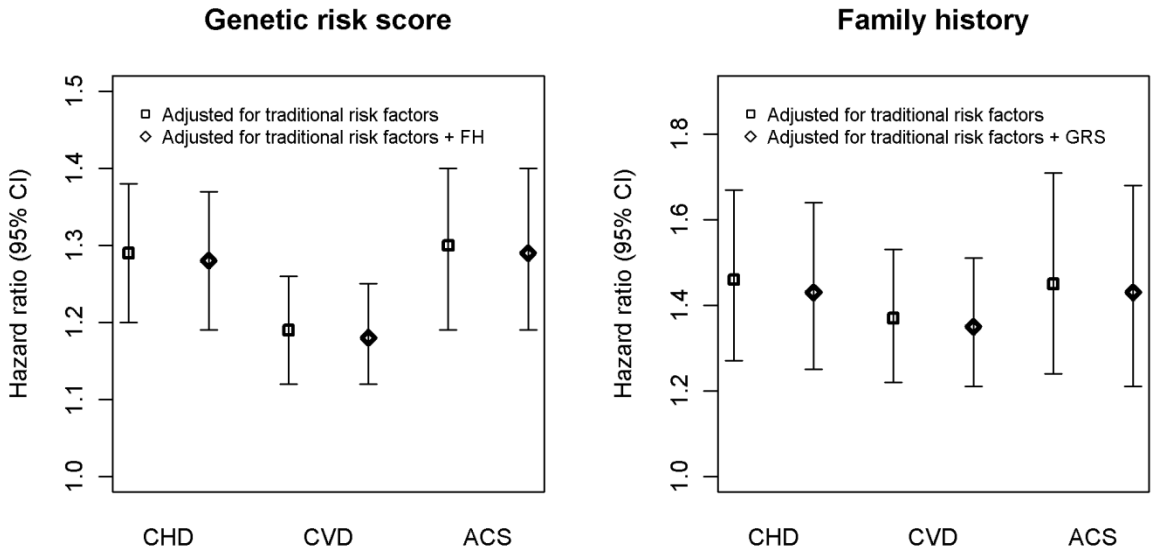


Figure 14. Hazard ratios for cardiovascular events by family history (FH) and genetic risk score (GRS) in FINRISK studies (N=19,001). FH effect: 0 = no family history of CVD, 1 = family history of CVD. GRS effect: per SD of GRS. Traditional risk factors include sex, total cholesterol, high-density lipoprotein (HDL) cholesterol, body mass index, systolic blood pressure, antihypertensive treatment, smoking and type 2 diabetes; age was used as the timescale. Abbreviations: CHD, coronary heart disease; CVD, cardiovascular disease; ACS, acute coronary syndrome; CI, confidence interval

5.5.3 Predictive performance of genetic risk scores

In Project III, the C-indices for CHD, CVD, and MI models with traditional risk factors and the 13-SNP GRS were 0.872, 0.853, and 0.881, respectively, and they were only slightly and nonsignificantly higher than the estimates from the models with only traditional risk factors (0.871, 0.853, and 0.880). Thus, the 13-SNP GRS did not improve the discrimination of cardiovascular end-points. Adding the GRS to the prediction model resulted in overall NRI=0.02 (95% CI -0.01–0.06, P=0.182) and clinical NRI=0.10 (95% CI 0.06–0.14, $P=3\times10^{-6}$).

When comparing C-indices for different risk factors (**Table 11**), gender had the highest discriminatory power for CHD (C-index 0.824, age was used as the time scale in the Cox model). The GRS comprising 28 SNPs had a C-index of 0.803. Thus, the GRS discriminated cases and noncases slightly better than conventional risk factors such as systolic blood pressure (C-index 0.801), baseline diabetes mellitus (C-index 0.801), and total cholesterol (C-index 0.800). Further, the 28-SNP GRS had a better discriminatory power than family history of CVD (0.803 vs. 0.795, P for change=0.01). Discrimination indices were lower in CVD models and higher in ACS models for all individual risk

factors. Gender, HDL, and smoking had the highest, and BMI and family history the lowest discrimination for both CVD and ACS.

Table 11. Risk discrimination of cardiovascular disease by individual risk factors.

Risk factor	C-index		
	CHD	CVD	ACS
Gender	0.824	0.808	0.833
HDL cholesterol	0.821	0.808	0.829
Current smoking	0.809	0.800	0.827
28-SNP GRS	0.803	0.792	0.813
Systolic blood pressure	0.801	0.798	0.812
Diabetes mellitus	0.801	0.795	0.813
13-SNP GRS	0.801	0.792	0.813
Total cholesterol	0.800	0.792	0.811
Body mass index	0.798	0.793	0.810
Family history of CVD	0.795	0.789	0.807

Age was used as the time scale in Cox proportional hazards models. Abbreviations: CHD, coronary heart disease; CVD, cardiovascular disease; ACS, acute coronary syndrome; HDL, high-density lipoprotein; SNP, single-nucleotide polymorphism; GRS, genetic risk score. Unpublished data.

The multivariable models with the traditional risk factors had a C-index of 0.849 for CHD, 0.835 for CVD, and 0.853 for ACS. The GRS slightly improved risk discrimination of CHD, CVD, and ACS over the traditional risk factors and family history. Adding family history to the model with traditional risk factors improved C-index by 0.2% (95% CI 0.1–0.3) for CHD. The addition of the GRS further improved the discrimination by 0.5% (95% CI 0.4–0.6). Corresponding values for CVD were 0.2% (95% CI 0.1–0.3) and 0.3% (95% CI 0.2–0.4), and for ACS 0.2% (95% CI 0.1–0.3) and 0.4% (95% CI 0.3–0.5).

Adding the 28-SNP GRS to the model with traditional risk factors and family history resulted in overall NRI=5% (95% CI 1–9, P=0.01). The GRS also improved reclassification of individuals in the intermediate-risk category (clinical NRI=27%, 95% CI 18–36, $P=1.1\times10^{-8}$). Of this group, 52 CHD cases (27%) were reclassified into the high-risk group and 206 noncases (20%) were correctly reclassified into the lower risk group, when the GRS was added to the model. The change in IDI was significant (Value=0.007, SE=0.002, $P=4.2\times10^{-5}$). The small value of IDI indicates that most of the reclassification occurs adjacent to risk thresholds.

The potential over-fitting of the models was evaluated by external validation. We estimated the effects for two models (with and without the GRS) in FINRISK 2002 and used the estimates to predict the 14-year absolute risk in the test data set (FINRISK 1992

and 1997). Health 2000 lacked the information on family history, thus it was not used as a training dataset. The results from this approach were compared with the method where the effects were estimated directly from the test data. No substantial differences were seen in discrimination or reclassification measures. For example, the C-index for the model with the traditional risk factors was 0.859, when estimated directly from the test data set, and 0.855 in external validation. The overall NRI was 5% in the test dataset and 4% in validation.

In a standard European population of 100,000 individuals, traditional risk factor screening would classify 64,373 subjects into <10%, 18,223 into 10–20% and 17,404 into ≥20% risk category (**Figure 15**). Subjects who had lipid medication or diabetes at baseline were automatically classified into the high-risk group. Additional genetic screening of subjects at intermediate risk for CHD would reclassify 3475 subjects (19%) into the low and 2144 (12%) into the high-risk category. Of those subjects classified into the high-risk category, 676 were expected CHD cases. Assuming that statins reduce the risk by 20% (Baigent et al. 2005), additional GRS screening for individuals in the intermediate-risk category could prevent 135 CHD cases (676×0.2) over 14 years. By contrast, random statin allocation for a comparable number of subjects at intermediate risk (N=2144) is expected to prevent only 54 cases (0.2×272, the expected number of cases).

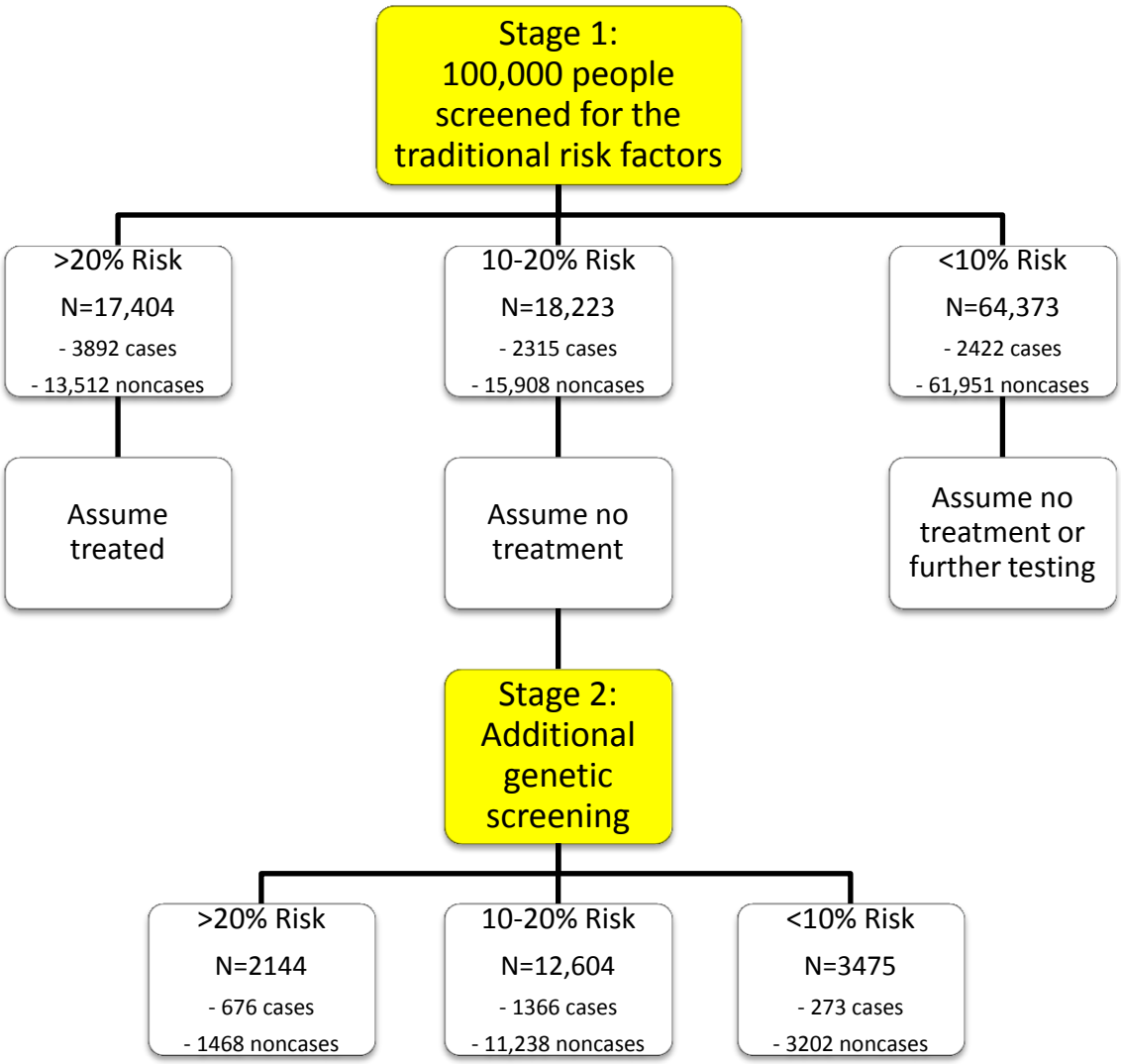


Figure 15. Two-stage risk screening of coronary heart disease in a standard population of 100,000 subjects. Treatment assumptions are based on the guidelines (National Institute for Health and Care Excellence 2008), which recommend that statins are allocated for subjects in the ≥20% risk group. In addition, subjects with baseline lipid treatment and/or diabetes were assumed to be treated.

6 DISCUSSION

6.1 Genetic risk profiles for lipids and blood pressure

In Projects I and II, we studied longitudinal genetic effects of SNPs that have been associated with lipid levels and blood pressure in GWASs. These studies have been conducted in cross-sectional samples of adult populations, and it is not known if the genetic effects are invariant over age, and if the results could be generalized to children and adolescents. By using repeated measurements of lipid and blood pressure levels between the ages of 3 and 45 years and genetic marker data, we observed that the GRSs for HDL, LDL, total cholesterol, triglycerides, and blood pressure were associated with longitudinal measurements of corresponding risk factor traits from childhood to young adulthood. Thus, the results from large-scale GWASs using adult samples seem to generalize to younger age groups.

In Project I, we examined a large number of genetic variants for blood lipid levels. Our data allowed us to study genetic effects for different age groups in more detail. Thus, we tested for interaction between the lipid-specific GRSs and the age groups of 3–6, 9, 12, 15, 18, 21–30, and 33–45 years. We found that the genetic effects for LDL and total cholesterol were stable across age and larger than for HDL and triglycerides in all age groups. Thus, genes have a substantial influence on LDL and total cholesterol levels over age, and this genetic effect is invariant over time. In contrast, we observed different patterns of genetic effects for HDL and triglycerides when different age groups were compared. The estimated GRS effects for triglycerides increased over age and decreased for HDL, especially in males. Moreover, the HDL loci explained two times more HDL variation among children than among adults. This is in line with at least one study that has reported higher heritability estimates for blood lipids in children than adults (Boomsma et al. 1996), and indicates that as environmental variables gain a greater impact with age, they also cause a larger amount of phenotypic variability in adults than in children.

The GRS comprising 32 SNPs associated with triglycerides improved the prediction of hypertriglyceridemia in young adults over and above the childhood lipid measurement. Thus, genetic information might provide useful additional information in initial assessment of the adulthood risk for dyslipidemia and heart disease. However, as we were unable to investigate the prediction of adulthood CHD in our young cohort, more studies are needed to evaluate the relationship between the genetic markers for triglycerides and CHD. Even though fasting triglyceride levels are associated with CHD (Sarwar et al. 2007), to date there is no clear evidence of an association between the genetic variants affecting triglyceride levels and CHD. Two studies that have evaluated the association between a comparable triglyceride GRS to our study and subclinical atherosclerosis have given somewhat contradicting results (Tsao et al. 2012; Isaacs et al. 2013).

In Project II, the blood pressure GRS was shown to be significantly associated with longitudinal measurements of systolic and diastolic blood pressure. However, we observed somewhat stronger associations between the GRS and adult blood pressure measurements taken in 2007. Contrary to lipids, we did not observe significant associations between the childhood blood pressure levels and the GRS. These results indicate that for these 13 SNPs genetic effects for blood pressure become stronger with age. One explanation for this could be that environmental or hormonal effects modify the association between blood pressure and genetic variants. Since dietary salt intake is one of the main environmental modifiers of blood pressure levels, we applied an interaction analysis for genetic variants and dietary salt intake, assessed with a questionnaire. We did not, however, observe significant interactions between salt intake and individual SNPs or the GRS, which might also be due to the high measurement error rate in salt intake variables. Further studies are needed to evaluate the modifying effects of other environmental variables, such as alcohol intake and physical activity.

We found a sex-specific effect for rs11191548 on systolic blood pressure, which is located near *CYP17A1*, a gene known to be involved in steroid hormone metabolism and to cause the rare Mendelian hypertensive disorder. The interaction has also been replicated in an independent population (Wu et al. 2012). As the SNP has been associated with CHD in adult populations (Schunkert et al. 2011), the region is of interest for future research regarding the relationships between hormonal changes, blood pressure and CHD.

We also found that the 13-SNP blood pressure GRS increased the risk for hypertension in adulthood, but did not improve the risk discrimination. An extended GRS with 30 blood pressure SNPs has been shown to increase the risk for CHD, stroke and left ventricular mass (Ehret et al. 2011; Lieb et al. 2013). Further, a 32-SNP blood pressure GRS was associated with incident CVD in a large prospective Finnish study, but it did not improve the risk discrimination or reclassification over and above traditional risk factors (Havulinna et al. 2013). These findings support the pathological role of blood pressure in CVD, but also suggest that the blood pressure GRS might not be useful in predicting CVD events.

6.2 Genetic risk profiles for coronary heart disease

In Projects III and IV, we evaluated the association between a panel of genetic variants and cardiovascular events in case-control and prospective Finnish and Swedish cohorts. We constructed weighted GRSs based on 13 (Project III) and 28 (Project IV) CHD SNPs and estimated the genetic effects on the risk of the first cardiovascular end-point. We found that the magnitude of the effect of the GRS was comparable with that of other existing risk factors such as blood pressure. Also, the genetic effect was independent of traditional risk factors as well as of a family history of CVD. The GRS based on 13 genetic variants did not improve overall risk reclassification or discrimination of CHD.

However, the extended GRS with 28 SNPs improved risk discrimination and reclassification over and above traditional risk factors and family history.

To address the clinical benefit of the 28-SNP GRS in population risk screening, we evaluated the effect of risk reclassification in a standard European population of 100,000 subjects. This concept has been applied in two recent prediction studies evaluating the predictive properties of inflammatory and lipid biomarkers (Di Angelantonio et al. 2012; Kaptoge et al. 2012). Our results suggest that additional GRS screening of individuals classified initially into the intermediate-risk group (14-year risk 10–20%) would reclassify 2144 (12%) of them into the high-risk category (≥20%). We estimate that statin allocation of these subjects could prevent 135 CHD cases over 14 years. In other words, targeted genetic screening could prevent one additional CHD event over a period of 14 years for every 135 people (18,223/135) screened. In comparison, lipoprotein(a) screening resulted in reclassification of 555 subjects from the intermediate- to the high-risk category, and potential prevention of 17 CVD events over 10 years (Di Angelantonio et al. 2012). However, it must be noted that these theoretical estimates of disease prevention should be interpreted with caution as they are based on several assumptions. For example, it is assumed that compliance in statin treatment is complete, which might overestimate the potential benefits. On the other hand, the estimates correspond to the number of cases that could be prevented with additional statin treatment, whereas other preventive strategies have not been taken into account. Modeling scenarios in which statin treatment is combined with lifestyle modification would be more realistic and lead to more accurate estimates for disease prevention.

Many studies have evaluated the predictive properties of comparable CHD SNP scores. Paynter et al. (2010) investigated 19,313 white women with a median follow-up 12.3 years (interquartile range 11.6–12.8 years), and constructed a GRS based on 12 genetic variants that had been found in GWASs between 2005 and 2009. Their GRS showed a modest effect for incident CVD after adjustment for traditional risk factors (HR=1.14, 95% CI 0.94–1.38 for top vs. bottom tertile, P=0.19). At least three other studies (Hughes et al. 2012; Thanassoulis et al. 2012; Vaarhorst et al. 2012) have reported results consistent with ours; although the findings have been slightly weaker. The results of these studies need, however, to be interpreted in light of their possible shortcomings. First, the studies might have been limited in power. For example, Thanassoulis et al. examined only 182 ‘hard CHD’ events (defined as coronary death or MI) and Paynter et al. 199 MI events. Second, event definitions in these cohorts might be more heterogeneous than ours. The main end-point of interest in the study of Paynter et al. included 203 stroke events. These genetic variants have not, however, been associated with strokes, and in our data, we also observed weaker associations for CVD than for CHD events. Finally, the study samples of Paynter et al. and Hughes et al. were limited to health professional females or middle-aged men, respectively, and thus are not representative of the general population.

In Projects III and IV, we applied two novel methods, NRI and IDI (Pencina et al. 2008), for assessing the incremental value of genetic marker information in CHD prediction. The use of these metrics has been motivated by the fact that traditionally used discrimination measure, C-index, is too insensitive a measure for evaluating predictive performance of new risk factors. For example, conventional cardiovascular risk factors, like LDL cholesterol, have modest impacts on C-index individually; in the Women's Health Study, adding LDL cholesterol to the model with age, systolic blood pressure, and smoking only increased the C-index from 0.76 to 0.77 (Cook 2007). However, since first introduced by Pencina et al. (2008), the use and interpretation of IDI and NRI have been under intensive discussion. Particularly, Cook (2008) has argued that the differences in mean predicted probabilities in cases and noncases tend to be small, and thus, the clinical relevance of IDI is modest. On the other hand, Pepe (2008) showed that there is a relationship between IDI and R2 in the model, which gives IDI an interesting alternative interpretation. Thus, IDI might be a useful measure of overall predictive performance, but as a category-free method, it has no impact on treatment decisions. By contrast, NRI is highly dependent on categories (Pencina et al. 2011), and thus, it is important that applied categories are based on valid decision-making thresholds. However, one limitation of NRI is that all reclassification is weighted equally, but actually, some movements between the categories are more relevant than others. For example, movement from intermediate risk to high risk for CHD cases is clinically more important than movement from low risk to intermediate risk (Pepe et al. 2008).

In conclusion, the GRS improved CHD prediction in our study, but more well-powered studies in different populations are needed to evaluate the joint genetic effects of these SNPs on CHD risk. It should be noted that as the methodology for predictive performance is under constant revision this might create fluctuation in results of different studies. It is also important to keep in mind that the identified variants to date explain only a modest part of the CHD heritability. As more genetic variants are identified, the risk estimates are likely to become more accurate.

6.3 Strengths and limitations of the study

In Projects I–IV, we examined genetic effects for CHD and its risk factors in large prospective datasets with longitudinal risk factor data and accurately defined cardiovascular phenotypes. We had the opportunity to evaluate the genetic effects not only on CVD, but also on more precisely defined cardiovascular phenotypes, such as CHD, ACS and MI. The diagnoses of these events are based on ICD-9 and ICD-10 codes from the validated national registries, which cover every death and hospitalization in the country (Pajunen et al. 2005; Tolonen et al. 2007; Sund 2012). We believe that the accuracy in event definitions and comprehensive coverage are the major contributors to the good statistical power of our study.

Our results should also be interpreted in light of the potential limitations of the study. In each project, we constructed GRSs based on the lead SNPs identified in GWASs. The SNPs have been selected on the basis of a stringent significance filter, and thus, the GRSs might lack many causal variants that have not yet achieved a genome-wide significance level. Further, the selected SNPs might not be biologically meaningful, and fine-mapping studies are needed to refine these association signals and locate the functional variants. The SNP panels used in this study are incomplete in the sense that they capture only a fraction of the phenotypic variability. For more precise genetic effect estimation, the gap between the heritability estimates and the explained genetic variation needs to be narrowed.

Beta coefficients (Projects I and II) and odds ratios (Projects III and IV) from the reference GWASs were used as weights when constructing the GRSs. The choice of weights might, however, affect the results. In Project IV, we observed comparable results for the GRSs computed by using odds ratios and log odds ratios as weights. One option is to construct GRSs simply by summing the number of risk alleles for each person and not to use weights at all. The difference between the weighted approach and allele counting has been evaluated in some studies. While some did not observe substantial differences between the two methods (Thanassoulis et al. 2012), others have shown that the weighting outperforms allele counting in discriminatory power (Davies et al. 2010).

We did not include gene-gene and gene-environment interaction effects to the GRSs, as current large-scale GWASs provide limited evidence of such interactions. Detection of interaction effects requires high statistical power, and many true effects might not reach the significance levels of GWASs. Thus, incorporating interaction effects into the genetic profiles would require a different approach for genetic marker selection. More advanced methods for filtering genetic markers for prediction studies include machine-learning techniques and network-based analysis (Okser et al. 2013).

In Projects I and II, we evaluated the genetic risk for dyslipidemia and hypertension and found that the GRSs were associated with the risk of these conditions. We did not, however, have an opportunity (due to young age of study participants) to evaluate the prediction of cardiovascular end-points in the same data, which would be of considerable interest. Even though dyslipidemia and hypertension are risk factors for cardiovascular events, the observed associations between the genetic variants and these risk factors do not necessarily imply a link between the same variants and CVD. Further studies are therefore needed to evaluate the predictive power of lipid and blood pressure GRSs on cardiovascular events.

We assessed the improvement in prediction of dyslipidemia and hypertension by using risk discrimination (e.g. ROC curves), but not reclassification methods (NRI). However, NRI is highly influenced by the choice of categories (Pencina et al. 2011), and because there are no established risk thresholds for dyslipidemia and hypertension that influence

treatment decisions, the category-based NRI is less meaningful. However, it has been suggested that if no established risk categories exist, one could use a single cut-off at 0.20, or calculate category-free NRI and visualize the results with reclassification plots (Pencina et al. 2011).

In Projects III and IV, we applied reclassification analysis in joint FINRISK 1992 and FINRISK 1997 data with 12- and 14-year follow-ups. However, the established CVD risk categories (0–5%, 5–10%, 10–20%, >20%) are usually applied for 10-year risk estimation. To address this, we applied a sensitivity analysis in Project IV with 2% higher risk thresholds. The choice of using four risk categories instead of three (e.g. 0–6%, 6–20%, and >20%) might also affect the reclassification results.

In Project IV, we estimated that 135 CHD events could be prevented with additional statin treatment for subject with a high GRS. It is not known, however, whether individuals with a high genetic risk would benefit from statins. Further studies are needed to examine the relationship between the genetic risk and lipid medication, which is becoming more frequent. For example, in FINRISK 1997 only 3% of study subject had baseline lipid medication, whereas in FINRISK 2002 the proportion was 7%.

Finally, these studies have been conducted in individuals of Finnish and Swedish decent, and the results might not be generalizable to other, especially nonEuropean, populations.

7 CONCLUSIONS AND FUTURE PROSPECTS

GWASs have been undeniably successful in identifying common genetic variants for common, complex diseases. Since 2005, over 10,000 genetic loci have been mapped to over 700 diseases and traits (Hindorff et al. 2013). These efforts have required huge investments, but for most traits, the loci identified to date explain only a fraction of the heritability. For example, even if not limited to genetic variants that have achieved a stringent genome-wide significance level, the identified variants for CHD explain approximately 10.6% of the disease heritability (Deloukas et al. 2012). Thus, the relevance of these findings for public health can be questioned.

GWASs serve as a starting point for several types of additional analyses. First, refining the identified GWAS loci is required to pinpoint the causal variants. Functional studies on causal variants are needed to unravel the novel molecular mechanisms underlying CHD. These analyses can be expected to highly increase our biological understanding of the disease, as most of the GWAS signals are not located near previous candidate genes. Second, the genetic variants can be used to assess the causal role of biomarkers in Mendelian randomization studies. With the exception of LDL cholesterol, it is not clear which of the traditional or novel CHD risk factors cause the disease and which are simply markers of some other causal processes. For example, the causal role of HDL cholesterol has recently been questioned due to lack of an association between HDL SNPs and CHD (Voight et al. 2012). Although noncausal risk factors might be useful in prediction, only causal risk factors are appropriate drug targets (Kathiresan and Srivastava 2012). Third, with the exception of 9p21, the genetic loci identified for CHD and (ischemic) stroke are completely nonoverlapping (Wahlstrand et al. 2009; Deloukas et al. 2012; Traylor et al. 2012), suggesting different molecular backgrounds for these subtypes of CVD. Moreover, while some genetic variants that have been mapped for CHD play a role in lipid metabolism, others seem to be specific for acute infarctions by affecting plaque instability (Reilly et al. 2011). Thus, GWAS findings could have an essential role in refining disease classifications and consequently, better targeting of treatment. Finally, the utility of genetic testing in CHD prediction is under debate and certainly requires further research.

Little progress has been made in preventive cardiovascular medicine since identification of traditional risk factors. Thus, there is room for improvement in CHD prediction, and genetic markers that have emerged during the last decade serve as an attractive tool for refining prediction algorithms. In this study, we have shown that the genetic variants identified in GWASs could be useful in early identification of individuals with increased CHD risk. Since atherosclerotic changes initiate already in childhood and the disease might be symptom-free for years, there is a need for better risk evaluation methods for young people, for whom 10-year predictions are not useful. It has been suggested that the lifetime risk estimation would overcome the limitations of 10-year prediction models for young adults with a substantial risk factor burden (Lloyd-Jones 2010). When evaluating

the risk burden for young people, genetic data have an advantage over other risk factors, as while the genetic risk is determined at conception, the high levels of traditional risk factors are typically not prevalent in children and adolescents.

Due to the wealth of genetic findings made in a short time frame, there is a large gap in our understanding of how these findings are involved in CHD pathophysiology. With the arrival of large-scale omics-data – genomics, transcriptomics, proteomics and metabolomics – new pathways, drug targets and predictive biomarkers are likely to be found in the future. So far, the CHD prediction studies (including Projects III and IV in this thesis) have been limited for studying individual biomarkers or restricted SNP panels, but with the availability of omics-data, a more holistic approach could be adopted also for prediction models. However, prospective study cohorts with comprehensive omics-data are still largely lacking, and the integration of different large-scale data levels requires methods development of efficient computational and modeling strategies.

8 ACKNOWLEDGEMENTS

This study was carried out at the Institute for Molecular Medicine Finland (FIMM), at the Hjelt Institute, University of Helsinki, and at the Public Health Genomics Unit, National Institute for Health and Welfare (THL). I thank the Director of FIMM, Professor Olli Kallioniemi, the head of the Hjelt Institute, Ritva Halila, the director of THL, Professor Pekka Puska, and the head of the Public Health Genomics Unit, Anu Jalanko, for providing excellent research facilities and an inspiring scientific working environment.

I would like to acknowledge financial support for this project provided by the BioSHaRE-EU Consortium, the ENGAGE consortium, the Helsinki Biomedical Graduate Program, the Hjelt Institute, and the Biomedicum Helsinki Foundation.

My deepest gratitude is owed to Dr. Daniel MacArthur for accepting the invitation to be my opponent at the defense. Professors Esa Läärä and Hannes Lohi are acknowledged for the pre-examination of this thesis. Your supportive and thought-provoking comments were extremely valuable, and I learned a lot while revising the thesis manuscript. I also thank my thesis committee members, Iiris Hovatta and Johan Lundin, for support and constructive feedback over the years. Carol Ann Pelli is greatly acknowledged for language editing of this thesis.

I warmly thank my supervisor, Professor Samuli Ripatti, who always finds time for his students. My first summer in the group as a trainee was probably the greatest learning experience of my life. Thank you for your guidance and also for the freedom given to me during the past year, which was essential for me to complete this project. I will forever be grateful to the late Academician of Science, Leena Peltonen, for accepting me as a part of her human genomics research group, even though I knew nothing about genetics. Leena was a world-class scientist and it was my privilege to work for her.

This work would not have been possible without numerous collaborations. I sincerely wish to acknowledge Professors Veikko Salomaa and Aarno Palotie for their expertise given to this project. I am grateful to Professors Olli Raitakari, Terho Lehtimäki, and others from the Young Finns Study cohort for a fruitful collaboration that resulted in two publications of my thesis. I warmly thank other senior scientists, Marju Orho-Melander, Elisabeth Widén, Kaisa Silander, Juha Sinisalo, Markus Perola, and many others for their guidance and contribution to this work. My gratitude is owed to all personnel and study participants in the study cohorts: Cardiovascular Risk in Young Finns Study, Corogene, Health 2000, FINRISK 1992–2002, Malmö Diet and Cancer Study, and Malmö Preventive Project. I am indebted to all co-authors, especially Dr. Aki Havulinna, Tarja Tuovinen, and Dr. Mervi Oikonen. Your input in the projects made this thesis possible.

I am grateful to all administrative personnel and IT support for the practical issues. Special thanks to Sari Kivikko and Huei-yi Shen, who have the answers for everything.

Many thanks to the FIMM IT guys, Teemu Perheentupa, Timo Miettinen, Jani Heikkinen, and Tomi Simonen. I also want to thank Mari Kaunisto for her help with the press release and for making my scientific jargon more understandable.

I have been fortunate to be a part of an exceptionally supportive and fun working environment. My long-term roommates Ida and Antti are thanked for friendship, for computational and statistical support, and for sharing all those essential YouTube videos (and unfortunately, "The Ascension") with me that have enriched my life. Thanks also to the rest of Samuli's poppoo, you have all helped me with various things. Special thanks to P-P for helping me with the Mendelian CHD chapter. Thanks Tiia, you have been a good friend of mine since "Hertsikka" times, and it has been great to have you as a colleague as well. I think that doctoral dissertation somehow resembles black belt graduation, and I have been lucky to have your support and company in both of these great challenges. For all memorable events, conference trips, lots of skumppa (and some Jamppa), I thank all other current and former colleagues: Alfredo, Anni, Annu, Ansku, Anu, Christian, Diana, Hanski, Helena, Himanshu, Jaakko, Jarkko, Joel, Johannes, Jonas, Juho, Jussi, Karola, Kati, Liisa, Luca, Mari, Marine, Marjis, Markus, Matti, Meri, Mikko, Minna, Minttu, Mitja, Nora, Olli, Outi, Perttu, Peter, Pietari, Priit, Saana, Taru, Tero, Verneri, Will, Virpi, and many others not mentioned here.

I warmly thank my family, the Salomäki family, relatives, and friends for counterbalancing my life. Thanks Jonna, Liisa, Mirva, and Siru for the spirit of Jamppa. Thanks Hinni, "Ms. Neanderthal", who is the only one in my extended family with whom I can play with R. Thanks Heini, with whom one day we will break the chin-up records. Thank you also to my other relatives. My warmest thanks go to my great and weird family: Dad, Mom, Tommi, Saana, and Salla. I do not idolize rock stars, actors, or even some science gurus; I admire my family. You have been the most important role models for me. Finally, I thank my husband, Jari, for everyday support and the highlights of my life, our adventures, which have taken place all over the world. Thank you for showing me how goals can be achieved.

Helsinki, September 29th, 2013

Emmi Tikkanen

9 REFERENCES

Abbott R. D., Wilson P. W., Kannel W. B., Castelli W. P. (1988). High density lipoprotein cholesterol, total cholesterol screening, and myocardial infarction. The Framingham Study. *Arteriosclerosis* **8**(3): 207-211.

Abecasis G. R., Auton A., Brooks L. D., DePristo M. A., Durbin R. M., Handsaker R. E., Kang H. M., Marth G. T., McVean G. A. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**(7422): 56-65.

Abifadel M., Varret M., Rabes J. P., Allard D., Ouguerram K., Devillers M., Cruaud C., Benjannet S., Wickham L., Erlich D., Derre A., Villeger L., Farnier M., Beucler I., Bruckert E., Chambaz J., Chanu B., Lecerf J. M., Luc G., Moulin P., Weissenbach J., Prat A., Krempf M., Junien C., Seidah N. G., Boileau C. (2003). Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* **34**(2): 154-156.

Akaike H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**(6): 716-723.

Alberti K. G., Zimmet P., Shaw J. (2005). The metabolic syndrome--a new worldwide definition. *Lancet* **366**(9491): 1059-1062.

Allan G. M., Nouri F., Korownyk C., Kolber M. R., Vandermeer B., McCormack J. (2013). Agreement among Cardiovascular Disease Risk Calculators. *Circulation* **127**(19): 1948-1956.

Ambrose J. A., Barua R. S. (2004). The pathophysiology of cigarette smoking and cardiovascular disease: an update. *J Am Coll Cardiol* **43**(10): 1731-1737.

Antolini L., Nam B., D'Agostico R. (2004). Inference on correlated discrimination measures in survival analysis: A nonparametric approach. *Commun Stat Theory Methods* **33**(9): 2117-2135.

Assmann G., Cullen P., Schulte H. (2002). Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Munster (PROCAM) study. *Circulation* **105**(3): 310-315.

Aulchenko Y. S., Ripatti S., Lindqvist I., Boomsma D., Heid I. M., Pramstaller P. P., Penninx B. W., Janssens A. C., Wilson J. F., Spector T., Martin N. G., Pedersen N. L., Kyvik K. O., Kaprio J., Hofman A., Freimer N. B., Jarvelin M. R., Gyllensten U., Campbell H., Rudan I., Johansson A., Marroni F., Hayward C., Vitart V., Jonasson I., Pattaro C., Wright A., Hastie N., Pichler I., Hicks A. A., Falchi M., Willemsen G., Hottenga J. J., de Geus E. J., Montgomery G. W., Whitfield J., Magnusson P., Saharinen J., Perola M., Silander K., Isaacs A., Sijbrands E. J., Uitterlinden A. G., Witteman J. C., Oostra B. A., Elliott P., Ruokonen A., Sabatti C., Gieger C., Meitinger T., Kronenberg F., Doring A., Wichmann H. E., Smit J. H., McCarthy M. I., van Duijn C. M., Peltonen L. (2009). Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet* **41**(1): 47-55.

Baigent C., Keech A., Kearney P. M., Blackwell L., Buck G., Pollicino C., Kirby A., Sourjina T., Peto R., Collins R., Simes R. (2005). Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet* **366**(9493): 1267-1278.

Barter P. J., Caulfield M., Eriksson M., Grundy S. M., Kastelein J. J., Komajda M., Lopez-Sendon J., Mosca L., Tardif J. C., Waters D. D., Shear C. L., Revkin J. H., Buhr K. A., Fisher M. R., Tall A. R., Brewer B. (2007). Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med* **357**(21): 2109-2122.

Barter P. J., Nicholls S., Rye K. A., Anantharamaiah G. M., Navab M., Fogelman A. M. (2004). Antiinflammatory properties of HDL. *Circ Res* **95**(8): 764-772.

Bayturan O., Kapadia S., Nicholls S. J., Tuzcu E. M., Shao M., Uno K., Shreevatsa A., Lavoie A. J., Wolski K., Schoenhagen P., Nissen S. E. (2010). Clinical predictors of plaque

progression despite very low levels of low-density lipoprotein cholesterol. *J Am Coll Cardiol* **55**(24): 2736-2742.

Berenson G. S., Srinivasan S. R., Bao W., Newman W. P., 3rd, Tracy R. E., Wattigney W. A. (1998). Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. The Bogalusa Heart Study. *N Engl J Med* **338**(23): 1650-1656.

Berglund G., Elmstahl S., Janzon L., Larsson S. A. (1993). The Malmo Diet and Cancer Study. Design and feasibility. *J Intern Med* **233**(1): 45-51.

Blankenberg S., Zeller T., Saarela O., Havulinna A. S., Kee F., Tunstall-Pedoe H., Kuulasmaa K., Yarnell J., Schnabel R. B., Wild P. S., Munzel T. F., Lackner K. J., Tiret L., Evans A., Salomaa V. (2010). Contribution of 30 biomarkers to 10-year cardiovascular risk estimation in 2 population cohorts: the MONICA, risk, genetics, archiving, and monograph (MORGAM) biomarker project. *Circulation* **121**(22): 2388-2397.

Boomsma D. I., Kempen H. J., Gevers Leuven J. A., Havekes L., de Knijff P., Frants R. R. (1996). Genetic analysis of sex and generation differences in plasma lipid, lipoprotein, and apolipoprotein levels in adolescent twins and their parents. *Genet Epidemiol* **13**(1): 49-60.

Brown M. S., Goldstein J. L. (1986). A receptor-mediated pathway for cholesterol homeostasis. *Science* **232**(4746): 34-47.

Canoy D., Boekholdt S. M., Wareham N., Luben R., Welch A., Bingham S., Buchan I., Day N., Khaw K. T. (2007). Body fat distribution and risk of coronary heart disease in men and women in the European Prospective Investigation Into Cancer and Nutrition in Norfolk cohort: a population-based prospective study. *Circulation* **116**(25): 2933-2943.

Chow C. K., Islam S., Bautista L., Rumboldt Z., Yusufali A., Xie C., Anand S. S., Engert J. C., Rangarajan S., Yusuf S. (2011). Parental history and myocardial infarction risk across the world: the INTERHEART Study. *J Am Coll Cardiol* **57**(5): 619-627.

Clamp M., Fry B., Kamal M., Xie X., Cuff J., Lin M. F., Kellis M., Lindblad-Toh K., Lander E. S. (2007). Distinguishing protein-coding and noncoding genes in the human genome. *Proc Natl Acad Sci U S A* **104**(49): 19428-19433.

Clarke R., Peden J. F., Hopewell J. C., Kyriakou T., Goel A., Heath S. C., Parish S., Barlera S., Franzosi M. G., Rust S., Bennett D., Silveira A., Malarstig A., Green F. R., Lathrop M., Gigante B., Leander K., de Faire U., Seedorf U., Hamsten A., Collins R., Watkins H., Farrall M. (2009). Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N Engl J Med* **361**(26): 2518-2528.

Conrad D. F., Jakobsson M., Coop G., Wen X., Wall J. D., Rosenberg N. A., Pritchard J. K. (2006). A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nat Genet* **38**(11): 1251-1260.

Conroy R. M., Pyorala K., Fitzgerald A. P., Sans S., Menotti A., De Backer G., De Bacquer D., Ducimetiere P., Jousilahti P., Keil U., Njolstad I., Oganov R. G., Thomsen T., Tunstall-Pedoe H., Tverdal A., Wedel H., Whincup P., Wilhelmsen L., Graham I. M. (2003). Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* **24**(11): 987-1003.

Cook N. R. (2007). Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* **115**(7): 928-935.

Cook N. R. (2008). Comments on 'Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond' by M. J. Pencina et al., Statistics in Medicine (DOI: 10.1002/sim.2929). *Stat Med* **27**(2): 191-195.

Coronary Artery Disease Genetics Consortium (2011). A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet* **43**(4): 339-344.

Corrao G., Bagnardi V., Zambon A., La Vecchia C. (2004). A meta-analysis of alcohol consumption and the risk of 15 diseases. *Prev Med* **38**(5): 613-619.

D'Agostino R. B., Sr., Vasan R. S., Pencina M. J., Wolf P. A., Cobain M., Massaro J. M., Kannel W. B. (2008). General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation* **117**(6): 743-753.

Danesh J., Lewington S., Thompson S. G., Lowe G. D., Collins R., Kostis J. B., Wilson A. C., Folsom A. R., Wu K., Benderly M., Goldbourt U., Willeit J., Kiechl S., Yarnell J. W., Sweetnam P. M., Elwood P. C., Cushman M., Psaty B. M., Tracy R. P., Tybjaerg-Hansen A., Haverkate F., de Maat M. P., Fowkes F. G., Lee A. J., Smith F. B., Salomaa V., Harald K., Rasi R., Vahtera E., Jousilahti P., Pekkanen J., D'Agostino R., Kannel W. B., Wilson P. W., Tofler G., Arocha-Pinango C. L., Rodriguez-Larralde A., Nagy E., Mijares M., Espinosa R., Rodriquez-Roa E., Ryder E., Diez-Ewald M. P., Campos G., Fernandez V., Torres E., Marchioli R., Valagussa F., Rosengren A., Wilhelmsen L., Lappas G., Eriksson H., Cremer P., Nagel D., Curb J. D., Rodriguez B., Yano K., Salonen J. T., Nyyssonen K., Tuomainen T. P., Hedblad B., Lind P., Loewel H., Koenig W., Meade T. W., Cooper J. A., De Stavola B., Knottenbelt C., Miller G. J., Bauer K. A., Rosenberg R. D., Sato S., Kitamura A., Naito Y., Palosuo T., Ducimetiere P., Amouyel P., Arveiler D., Evans A. E., Ferrieres J., Juhan-Vague I., Bingham A., Schulte H., Assmann G., Cantin B., Lamarche B., Despres J. P., Dagenais G. R., Tunstall-Pedoe H., Woodward M., Ben-Shlomo Y., Davey Smith G., Palmieri V., Yeh J. L., Rudnicka A., Ridker P., Rodeghiero F., Tosetto A., Shepherd J., Ford I., Robertson M., Brunner E., Shipley M., Feskens E. J., Kromhout D., Dickinson A., Ireland B., Juzwishin K., Kaptoge S., Memon A., Sarwar N., Walker M., Wheeler J., White I., Wood A. (2005). Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *JAMA* **294**(14): 1799-1809.

Danesh J., Wheeler J. G., Hirschfield G. M., Eda S., Eiriksdottir G., Rumley A., Lowe G. D., Pepys M. B., Gudnason V. (2004). C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease. *N Engl J Med* **350**(14): 1387-1397.

Dauchet L., Amouyel P., Hercberg S., Dallongeville J. (2006). Fruit and vegetable consumption and risk of coronary heart disease: a meta-analysis of cohort studies. *J Nutr* **136**(10): 2588-2593.

Davies R. W., Dandona S., Stewart A. F., Chen L., Ellis S. G., Tang W. H., Hazen S. L., Roberts R., McPherson R., Wells G. A. (2010). Improved prediction of cardiovascular disease based on a panel of single nucleotide polymorphisms identified through genome-wide association studies. *Circ Cardiovasc Genet* **3**(5): 468-474.

Davis P. H., Dawson J. D., Riley W. A., Lauer R. M. (2001). Carotid intimal-medial thickness is related to cardiovascular risk factors measured from childhood through middle age: The Muscatine Study. *Circulation* **104**(23): 2815-2819.

Deloukas P., Kanoni S., Willenborg C., Farrall M., Assimes T. L., Thompson J. R., Ingelsson E., Saleheen D., Erdmann J., Goldstein B. A., Stirrups K., Konig I. R., Cazier J. B., Johansson A., Hall A. S., Lee J. Y., Willer C. J., Chambers J. C., Esko T., Folkersen L., Goel A., Grundberg E., Havulinna A. S., Ho W. K., Hopewell J. C., Eriksson N., Kleber M. E., Kristiansson K., Lundmark P., Lyytikainen L. P., Rafelt S., Shungin D., Strawbridge R. J., Thorleifsson G., Tikkanen E., Van Zuydam N., Voight B. F., Waite L. L., Zhang W., Ziegler A., Absher D., Altshuler D., Balmforth A. J., Barroso I., Braund P. S., Burgdorf C., Claudi-Boehm S., Cox D., Dimitriou M., Do R., Doney A. S., Mokhtari N. E., Eriksson P., Fischer K., Fontanillas P., Franco-Cereceda A., Gigante B., Groop L., Gustafsson S., Hager J., Hallmans G., Han B. G., Hunt S. E., Kang H. M., Illig T., Kessler T., Knowles J. W., Kolovou G., Kuusisto J., Langenberg C., Langford C., Leander K., Lokki M. L., Lundmark A., McCarthy M. I., Meisinger C., Melander O., Mihailov E., Maouche S., Morris A. D., Muller-Nurasyid M., Nikus K., Peden J. F., Rayner N. W., Rasheed A., Rosinger S., Rubin D., Rumpf M. P., Schafer A., Sivananthan M., Song C., Stewart A. F., Tan S. T., Thorgeirsson G., Schoot C. E., Wagner P. J., Wells G. A., Wild P. S., Yang T. P., Amouyel P., Arveiler D., Basart H., Boehnke M.,

Boerwinkle E., Brambilla P., Cambien F., Cupples A. L., de Faire U., Dehghan A., Diemert P., Epstein S. E., Evans A., Ferrario M. M., Ferrieres J., Gauguier D., Go A. S., Goodall A. H., Gudnason V., Hazen S. L., Holm H., Iribarren C., Jang Y., Kahonen M., Kee F., Kim H. S., Klopp N., Koenig W., Kratzer W., Kuulasmaa K., Laakso M., Laaksonen R., Lind L., Ouwehand W. H., Parish S., Park J. E., Pedersen N. L., Peters A., Quertermous T., Rader D. J., Salomaa V., Schadt E., Shah S. H., Sinisalo J., Stark K., Stefansson K., Tregouet D. A., Virtamo J., Wallentin L., Wareham N., Zimmermann M. E., Nieminen M. S., Hengstenberg C., Sandhu M. S., Pastinen T., Syvanen A. C., Hovingh G. K., Dedoussis G., Franks P. W., Lehtimaki T., Metspalu A., Zalloua P. A., Siegbahn A., Schreiber S., Ripatti S., Blankenberg S. S., Perola M., Clarke R., Boehm B. O., O'Donnell C., Reilly M. P., Marz W., Collins R., Kathiresan S., Hamsten A., Kooner J. S., Thorsteinsdottir U., Danesh J., Palmer C. N., Roberts R., Watkins H., Schunkert H., Samani N. J. (2012). Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet* **45**(1): 25-33.

Den Ruijter H. M., Peters S. A., Anderson T. J., Britton A. R., Dekker J. M., Eijkemans M. J., Engstrom G., Evans G. W., de Graaf J., Grobbee D. E., Hedblad B., Hofman A., Holewijn S., Ikeda A., Kavousi M., Kitagawa K., Kitamura A., Koffijberg H., Lonn E. M., Lorenz M. W., Mathiesen E. B., Nijpels G., Okazaki S., O'Leary D. H., Polak J. F., Price J. F., Robertson C., Rembold C. M., Rosvall M., Rundek T., Salonen J. T., Sitzer M., Stehouwer C. D., Witteman J. C., Moons K. G., Bots M. L. (2012). Common carotid intima-media thickness measurements in cardiovascular risk prediction: a meta-analysis. *JAMA* **308**(8): 796-803.

Di Angelantonio E., Gao P., Pennells L., Kaptoge S., Caslake M., Thompson A., Butterworth A. S., Sarwar N., Wormser D., Saleheen D., Ballantyne C. M., Psaty B. M., Sundstrom J., Ridker P. M., Nagel D., Gillum R. F., Ford I., Ducimetiere P., Kiechl S., Koenig W., Dullaart R. P., Assmann G., D'Agostino R. B., Sr., Dagenais G. R., Cooper J. A., Kromhout D., Onat A., Tipping R. W., Gomez-de-la-Camara A., Rosengren A., Sutherland S. E., Gallacher J., Fowkes F. G., Casiglia E., Hofman A., Salomaa V., Barrett-Connor E., Clarke R., Brunner E., Jukema J. W., Simons L. A., Sandhu M., Wareham N. J., Khaw K. T., Kauhanen J., Salonen J. T., Howard W. J., Nordestgaard B. G., Wood A. M., Thompson S. G., Boekholdt S. M., Sattar N., Packard C., Gudnason V., Danesh J. (2012). Lipid-related markers and cardiovascular disease prediction. *JAMA* **307**(23): 2499-2506.

Di Angelantonio E., Sarwar N., Perry P., Kaptoge S., Ray K. K., Thompson A., Wood A. M., Lewington S., Sattar N., Packard C. J., Collins R., Thompson S. G., Danesh J. (2009). Major lipids, apolipoproteins, and risk of vascular disease. *JAMA* **302**(18): 1993-2000.

Do C. B., Hinds D. A., Francke U., Eriksson N. (2012). Comparison of family history and SNPs for predicting risk of complex disease. *PLoS Genet* **8**(10): e1002973.

Dunham I., Kundaje A., Aldred S. F., Collins P. J., Davis C. A., Doyle F., Epstein C. B., Frietze S., Harrow J., Kaul R., Khatun J., Lajoie B. R., Landt S. G., Lee B. K., Pauli F., Rosenbloom K. R., Sabo P., Safi A., Sanyal A., Shoresh N., Simon J. M., Song L., Trinklein N. D., Altshuler R. C., Birney E., Brown J. B., Cheng C., Djebali S., Dong X., Ernst J., Furey T. S., Gerstein M., Giardine B., Greven M., Hardison R. C., Harris R. S., Herrero J., Hoffman M. M., Iyer S., Kelllis M., Kheradpour P., Lassman T., Li Q., Lin X., Marinov G. K., Merkel A., Mortazavi A., Parker S. C., Reddy T. E., Rozowsky J., Schlesinger F., Thurman R. E., Wang J., Ward L. D., Whitfield T. W., Wilder S. P., Wu W., Xi H. S., Yip K. Y., Zhuang J., Bernstein B. E., Green E. D., Gunter C., Snyder M., Pazin M. J., Lowdon R. F., Dillon L. A., Adams L. B., Kelly C. J., Zhang J., Wexler J. R., Good P. J., Feingold E. A., Crawford G. E., Dekker J., Elinitski L., Farnham P. J., Giddings M. C., Gingeras T. R., Guigo R., Hubbard T. J., Kellis M., Kent W. J., Lieb J. D., Margulies E. H., Myers R. M., Starnatoyannopoulos J. A., Tennebaum S. A., Weng Z., White K. P., Wold B., Yu Y., Wrobel J., Risk B. A., Gunawardena H. P., Kuiper H. C., Maier C. W., Xie L., Chen X., Mikkelsen T. S., Gillespie S., Goren A., Ram O.,

Zhang X., Wang L., Issner R., Coyne M. J., Durham T., Ku M., Truong T., Eaton M. L., Dobin A., Lassmann T., Tanzer A., Lagarde J., Lin W., Xue C., Williams B. A., Zaleski C., Roder M., Kokocinski F., Abdelhamid R. F., Alioto T., Antoshechkin I., Baer M. T., Batut P., Bell I., Bell K., Chakrabortty S., Chrast J., Curado J., Derrien T., Drenkow J., Dumais E., Dumais J., Duttagupta R., Fastuca M., Fejes-Toth K., Ferreira P., Foissac S., Fullwood M. J., Gao H., Gonzalez D., Gordon A., Howald C., Jha S., Johnson R., Kapranov P., King B., Kingswood C., Li G., Luo O. J., Park E., Preall J. B., Presaud K., Ribeca P., Robyr D., Ruan X., Sammeth M., Sandu K. S., Schaeffer L., See L. H., Shahab A., Skancke J., Suzuki A. M., Takahashi H., Tilgner H., Trout D., Walters N., Wang H., Hayashizaki Y., Reymond A., Antonarakis S. E., Hannon G. J., Ruan Y., Carninci P., Sloan C. A., Learned K., Malladi V. S., Wong M. C., Barber G. P., Cline M. S., Dreszer T. R., Heitner S. G., Karolchik D., Kirkup V. M., Meyer L. R., Long J. C., Maddren M., Raney B. J., Grasfeder L. L., Giresi P. G., Battenhouse A., Sheffield N. C., Showers K. A., London D., Bhinge A. A., Shestak C., Schaner M. R., Kim S. K., Zhang Z. Z., Mieczkowski P. A., Mieczkowska J. O., Liu Z., McDaniell R. M., Ni Y., Rashid N. U., Kim M. J., Adar S., Zhang Z., Wang T., Winter D., Keefe D., Iyer V. R., Sandhu K. S., Zheng M., Wang P., Gertz J., Vielmetter J., Partridge E. C., Varley K. E., Gasper C., Bansal A., Pepke S., Jain P., Amrhein H., Bowling K. M., Anaya M., Cross M. K., Muratet M. A., Newberry K. M., McCue K., Nesmith A. S., Fisher-Aylor K. I., Pusey B., DeSalvo G., Parker S. L., Balasubramanian S., Davis N. S., Meadows S. K., Eggleston T., Newberry J. S., Levy S. E., Absher D. M., Wong W. H., Blow M. J., Visel A., Pennachio L. A., Elnitski L., Petrykowska H. M., Abyzov A., Aken B., Barrell D., Barson G., Berry A., Bignell A., Boychenko V., Bussotti G., Davidson C., Despacio-Reyes G., Diekhans M., Ezkurdia I., Frankish A., Gilbert J., Gonzalez J. M., Griffiths E., Harte R., Hendrix D. A., Hunt T., Jungreis I., Kay M., Khurana E., Leng J., Lin M. F., Loveland J., Lu Z., Manthravadi D., Mariotti M., Mudge J., Mukherjee G., Notredame C., Pei B., Rodriguez J. M., Saunders G., Sboner A., Searle S., Sisu C., Snow C., Steward C., Tapanan E., Tress M. L., van Baren M. J., Washieti S., Wilming L., Zadissa A., Zhengdong Z., Brent M., Haussler D., Valencia A., Raymond A., Addleman N., Alexander R. P., Auerbach R. K., Bettinger K., Bhardwaj N., Boyle A. P., Cao A. R., Cayting P., Charos A., Cheng Y., Eastman C., Euskirchen G., Fleming J. D., Grubert F., Habegger L., Hariharan M., Harmanci A., Iyenger S., Jin V. X., Karczewski K. J., Kasowski M., Lacroute P., Lam H., Larnarre-Vincent N., Lian J., Lindahl-Allen M., Min R., Miotto B., Monahan H., Moqtaderi Z., Mu X. J., O'Geen H., Ouyang Z., Patacsil D., Raha D., Ramirez L., Reed B., Shi M., Slifer T., Witt H., Wu L., Xu X., Yan K. K., Yang X., Struhl K., Weissman S. M., Tenebaum S. A., Penalva L. O., Karmakar S., Bhanvadia R. R., Choudhury A., Domanus M., Ma L., Moran J., Victorsen A., Auer T., Centarin L., Eichenlaub M., Gruhl F., Heerman S., Hoeckendorf B., Inoue D., Kellner T., Kirchmaier S., Mueller C., Reinhardt R., Schertel L., Schneider S., Sinn R., Wittbrodt B., Wittbrodt J., Jain G., Balasundaram G., Bates D. L., Byron R., Canfield T. K., Diegel M. J., Dunn D., Ebersol A. K., Frum T., Garg K., Gist E., Hansen R. S., Boatman L., Haugen E., Humbert R., Johnson A. K., Johnson E. M., Kutyavin T. M., Lee K., Lotakis D., Maurano M. T., Neph S. J., Neri F. V., Nguyen E. D., Qu H., Reynolds A. P., Roach V., Rynes E., Sanchez M. E., Sandstrom R. S., Shafer A. O., Stergachis A. B., Thomas S., Vernot B., Vierstra J., Vong S., Weaver M. A., Yan Y., Zhang M., Akey J. A., Bender M., Dorschner M. O., Groudine M., MacCoss M. J., Navas P., Stamatoyannopoulos G., Stamatoyannopoulos J. A., Beal K., Brazma A., Flicek P., Johnson N., Lukk M., Luscombe N. M., Sobral D., Vaquerizas J. M., Batzoglou S., Sidow A., Hussami N., Kyriazopoulou-Panagiotopoulou S., Libbrecht M. W., Schaub M. A., Miller W., Bickel P. J., Banfai B., Boley N. P., Huang H., Li J. J., Noble W. S., Bilmes J. A., Buske O. J., Sahu A. O., Kharchenko P. V., Park P. J., Baker D., Taylor J., Lochovsky L. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**(7414): 57-74.

Ehret G. B., Caulfield M. J. (2013). Genes for blood pressure: an opportunity to understand hypertension. *Eur Heart J* **34**(13): 951-961.

Ehret G. B., Munroe P. B., Rice K. M., Bochud M., Johnson A. D., Chasman D. I., Smith A. V., Tobin M. D., Verwoert G. C., Hwang S. J., Pihur V., Vollenweider P., O'Reilly P. F., Amin N., Bragg-Gresham J. L., Teumer A., Glazer N. L., Launer L., Zhao J. H., Aulchenko Y., Heath S., Sober S., Parsa A., Luan J., Arora P., Dehghan A., Zhang F., Lucas G., Hicks A. A., Jackson A. U., Peden J. F., Tanaka T., Wild S. H., Rudan I., Igl W., Milaneschi Y., Parker A. N., Fava C., Chambers J. C., Fox E. R., Kumari M., Go M. J., van der Harst P., Kao W. H., Sjogren M., Vinay D. G., Alexander M., Tabara Y., Shaw-Hawkins S., Whincup P. H., Liu Y., Shi G., Kuusisto J., Tayo B., Seielstad M., Sim X., Nguyen K. D., Lehtimaki T., Matullo G., Wu Y., Gaunt T. R., Onland-Moret N. C., Cooper M. N., Platou C. G., Org E., Hardy R., Dahgam S., Palmen J., Vitart V., Braund P. S., Kuznetsova T., Uiterwaal C. S., Adeyemo A., Palmas W., Campbell H., Ludwig B., Tomaszewski M., Tzoulaki I., Palmer N. D., Aspelund T., Garcia M., Chang Y. P., O'Connell J. R., Steinle N. I., Grobbee D. E., Arking D. E., Kardia S. L., Morrison A. C., Hernandez D., Najjar S., McArdle W. L., Hadley D., Brown M. J., Connell J. M., Hingorani A. D., Day I. N., Lawlor D. A., Beilby J. P., Lawrence R. W., Clarke R., Hopewell J. C., Ongen H., Dreisbach A. W., Li Y., Young J. H., Bis J. C., Kahonen M., Viikari J., Adair L. S., Lee N. R., Chen M. H., Olden M., Pattaro C., Bolton J. A., Kottgen A., Bergmann S., Mooser V., Chaturvedi N., Frayling T. M., Islam M., Jafar T. H., Erdmann J., Kulkarni S. R., Bornstein S. R., Grassler J., Groop L., Voight B. F., Kettunen J., Howard P., Taylor A., Guarrera S., Ricceri F., Emilsson V., Plump A., Barroso I., Khaw K. T., Weder A. B., Hunt S. C., Sun Y. V., Bergman R. N., Collins F. S., Bonnycastle L. L., Scott L. J., Stringham H. M., Peltonen L., Perola M., Vartiainen E., Brand S. M., Staessen J. A., Wang T. J., Burton P. R., Soler Artigas M., Dong Y., Snieder H., Wang X., Zhu H., Lohman K. K., Rudock M. E., Heckbert S. R., Smith N. L., Wiggins K. L., Doumatey A., Shriner D., Veldre G., Viigimaa M., Kinra S., Prabhakaran D., Tripathy V., Langefeld C. D., Rosengren A., Thelle D. S., Corsi A. M., Singleton A., Forrester T., Hilton G., McKenzie C. A., Salako T., Iwai N., Kita Y., Ogihara T., Ohkubo T., Okamura T., Ueshima H., Umemura S., Eyheramendy S., Meitinger T., Wichmann H. E., Cho Y. S., Kim H. L., Lee J. Y., Scott J., Sehmi J. S., Zhang W., Hedblad B., Nilsson P., Smith G. D., Wong A., Narisu N., Stancakova A., Raffel L. J., Yao J., Kathiresan S., O'Donnell C. J., Schwartz S. M., Ikram M. A., Longstreth W. T., Jr., Mosley T. H., Seshadri S., Shrine N. R., Wain L. V., Morken M. A., Swift A. J., Laitinen J., Prokopenko I., Zitting P., Cooper J. A., Humphries S. E., Danesh J., Rasheed A., Goel A., Hamsten A., Watkins H., Bakker S. J., van Gilst W. H., Janipalli C. S., Mani K. R., Yajnik C. S., Hofman A., Mattace-Raso F. U., Oostra B. A., Demirkan A., Isaacs A., Rivadeneira F., Lakatta E. G., Orru M., Scuteri A., Ala-Korpela M., Kangas A. J., Lyytikainen L. P., Soininen P., Tukiainen T., Wurtz P., Ong R. T., Dorr M., Kroemer H. K., Volker U., Volzke H., Galan P., Hercberg S., Lathrop M., Zelenika D., Deloukas P., Mangino M., Spector T. D., Zhai G., Meschia J. F., Nalls M. A., Sharma P., Terzic J., Kumar M. V., Denniff M., Zukowska-Szczechowska E., Wagenknecht L. E., Fowkes F. G., Charchar F. J., Schwarz P. E., Hayward C., Guo X., Rotimi C., Bots M. L., Brand E., Samani N. J., Polasek O., Talmud P. J., Nyberg F., Kuh D., Laan M., Hveem K., Palmer L. J., van der Schouw Y. T., Casas J. P., Mohlke K. L., Vineis P., Raitakari O., Ganesh S. K., Wong T. Y., Tai E. S., Cooper R. S., Laakso M., Rao D. C., Harris T. B., Morris R. W., Dominiczak A. F., Kivimaki M., Marmot M. G., Miki T., Saleheen D., Chandak G. R., Coresh J., Navis G., Salomaa V., Han B. G., Zhu X., Kooner J. S., Melander O., Ridker P. M., Bandinelli S., Gyllensten U. B., Wright A. F., Wilson J. F., Ferrucci L., Farrall M., Tuomilehto J., Pramstaller P. P., Elosua R., Soranzo N., Sijbrands E. J., Altshuler D., Loos R. J., Shuldiner A. R., Gieger C., Meneton P., Uitterlinden A. G., Wareham N. J., Gudnason V., Rotter J. I., Rettig R., Uda M., Strachan D. P., Witteman J. C., Hartikainen A. L., Beckmann J. S., Boerwinkle E., Vasan R. S., Boehnke M., Larson M. G., Jarvelin

M. R., Psaty B. M., Abecasis G. R., Chakravarti A., Elliott P., van Duijn C. M., Newton-Cheh C., Levy D., Caulfield M. J., Johnson T. (2011). Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**(7367): 103-109.

Endo A., Kuroda M., Tsujita Y. (1976). ML-236A, ML-236B, and ML-236C, new inhibitors of cholesterogenesis produced by Penicillium citrinium. *J Antibiot (Tokyo)* **29**(12): 1346-1348.

Enos W. F., Holmes R. H., Beyer J. (1953). Coronary disease among United States soldiers killed in action in Korea; preliminary report. *J Am Med Assoc* **152**(12): 1090-1093.

Erdmann J., Grosshennig A., Braund P. S., Konig I. R., Hengstenberg C., Hall A. S., Linsel-Nitschke P., Kathiresan S., Wright B., Tregouet D. A., Cambien F., Bruse P., Aherrahrou Z., Wagner A. K., Stark K., Schwartz S. M., Salomaa V., Elosua R., Melander O., Voight B. F., O'Donnell C. J., Peltonen L., Siscovick D. S., Altshuler D., Merlini P. A., Peyvandi F., Bernardinelli L., Ardissino D., Schillert A., Blankenberg S., Zeller T., Wild P., Schwarz D. F., Tiret L., Perret C., Schreiber S., El Mokhtari N. E., Schafer A., Marz W., Renner W., Bugert P., Kluter H., Schrezenmeir J., Rubin D., Ball S. G., Balmforth A. J., Wichmann H. E., Meitinger T., Fischer M., Meisinger C., Baumert J., Peters A., Ouwehand W. H., Deloukas P., Thompson J. R., Ziegler A., Samani N. J., Schunkert H. (2009). New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat Genet* **41**(3): 280-282.

Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (2001). 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* **285**(19): 2486-2497.

Framingham Heart Study. (2013). Framingham Heart Study. A Project of the National Heart, Lung and Blood Institute and Boston University. Retrieved March 18, 2013, from http://www.framinghamheartstudy.org/index.html.

Frayling T. M., Timpson N. J., Weedon M. N., Zeggini E., Freathy R. M., Lindgren C. M., Perry J. R., Elliott K. S., Lango H., Rayner N. W., Shields B., Harries L. W., Barrett J. C., Ellard S., Groves C. J., Knight B., Patch A. M., Ness A. R., Ebrahim S., Lawlor D. A., Ring S. M., Ben-Shlomo Y., Jarvelin M. R., Sovio U., Bennett A. J., Melzer D., Ferrucci L., Loos R. J., Barroso I., Wareham N. J., Karpe F., Owen K. R., Cardon L. R., Walker M., Hitman G. A., Palmer C. N., Doney A. S., Morris A. D., Smith G. D., Hattersley A. T., McCarthy M. I. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**(5826): 889-894.

Frisoli T. M., Schmieder R. E., Grodzicki T., Messerli F. H. (2011). Beyond salt: lifestyle modifications and blood pressure. *Eur Heart J* **32**(24): 3081-3087.

Garces C., Oya I., Lasuncion M. A., Lopez-Simon L., Cano B., de Oya M. (2010). Sex hormone-binding globulin and lipid profile in pubertal children. *Metabolism* **59**(2): 166-171.

Garcia C. K., Wilund K., Arca M., Zuliani G., Fellin R., Maioli M., Calandra S., Bertolini S., Cossu F., Grishin N., Barnes R., Cohen J. C., Hobbs H. H. (2001). Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. *Science* **292**(5520): 1394-1398.

Gieger C., Radhakrishnan A., Cvejic A., Tang W., Porcu E., Pistis G., Serbanovic-Canic J., Elling U., Goodall A. H., Labrune Y., Lopez L. M., Magi R., Meacham S., Okada Y., Pirastu N., Sorice R., Teumer A., Voss K., Zhang W., Ramirez-Solis R., Bis J. C., Ellinghaus D., Gogele M., Hottenga J. J., Langenberg C., Kovacs P., O'Reilly P. F., Shin S. Y., Esko T., Hartiala J., Kanoni S., Murgia F., Parsa A., Stephens J., van der Harst P., Ellen van der Schoot C., Allayee H., Attwood A., Balkau B., Bastardot F., Basu S., Baumeister S. E., Biino G., Bomba L., Bonnefond A., Cambien F., Chambers J. C., Cucca F., D'Adamo P., Davies G., de Boer R. A., de Geus E. J., Doring A., Elliott P., Erdmann J., Evans D. M., Falchi M., Feng W., Folsom A. R., Frazer I. H., Gibson Q. D., Glazer N. L., Hammond C., Hartikainen A. L., Heckbert S. R., Hengstenberg C., Hersch M., Illig T., Loos R. J., Jolley J., Khaw K. T., Kuhnel B., Kyrtsonis M. C., Lagou V., Lloyd-Jones H., Lumley T.,

Mangino M., Maschio A., Mateo Leach I., McKnight B., Memari Y., Mitchell B. D., Montgomery G. W., Nakamura Y., Nauck M., Navis G., Nothlings U., Nolte I. M., Porteous D. J., Pouta A., Pramstaller P. P., Pullat J., Ring S. M., Rotter J. I., Ruggiero D., Ruokonen A., Sala C., Samani N. J., Sambrook J., Schlessinger D., Schreiber S., Schunkert H., Scott J., Smith N. L., Snieder H., Starr J. M., Stumvoll M., Takahashi A., Tang W. H., Taylor K., Tenesa A., Lay Thein S., Tonjes A., Uda M., Ulivi S., van Veldhuisen D. J., Visscher P. M., Volker U., Wichmann H. E., Wiggins K. L., Willemsen G., Yang T. P., Hua Zhao J., Zitting P., Bradley J. R., Dedoussis G. V., Gasparini P., Hazen S. L., Metspalu A., Pirastu M., Shuldiner A. R., Joost van Pelt L., Zwaginga J. J., Boomsma D. I., Deary I. J., Franke A., Froguel P., Ganesh S. K., Jarvelin M. R., Martin N. G., Meisinger C., Psaty B. M., Spector T. D., Wareham N. J., Akkerman J. W., Ciullo M., Deloukas P., Greinacher A., Jupe S., Kamatani N., Khadake J., Kooner J. S., Penninger J., Prokopenko I., Stemple D., Toniolo D., Wernisch L., Sanna S., Hicks A. A., Rendon A., Ferreira M. A., Ouwehand W. H., Soranzo N. (2011). New gene functions in megakaryopoiesis and platelet formation. *Nature* **480**(7376): 201-208.

Gradman A. H., Alfayoumi F. (2006). From left ventricular hypertrophy to congestive heart failure: management of hypertensive heart disease. *Prog Cardiovasc Dis* **48**(5): 326-341.

Grundy S. M., Benjamin I. J., Burke G. L., Chait A., Eckel R. H., Howard B. V., Mitch W., Smith S. C., Jr., Sowers J. R. (1999). Diabetes and cardiovascular disease: a statement for healthcare professionals from the American Heart Association. *Circulation* **100**(10): 1134-1146.

Gudbjartsson D. F., Bjornsdottir U. S., Halapi E., Helgadottir A., Sulem P., Jonsdottir G. M., Thorleifsson G., Helgadottir H., Steinthorsdottir V., Stefansson H., Williams C., Hui J., Beilby J., Warrington N. M., James A., Palmer L. J., Koppelman G. H., Heinzmann A., Krueger M., Boezen H. M., Wheatley A., Altmuller J., Shin H. D., Uh S. T., Cheong H. S., Jonsdottir B., Gislason D., Park C. S., Rasmussen L. M., Porsbjerg C., Hansen J. W., Backer V., Werge T., Janson C., Jonsson U. B., Ng M. C., Chan J., So W. Y., Ma R., Shah S. H., Granger C. B., Quyyumi A. A., Levey A. I., Vaccarino V., Reilly M. P., Rader D. J., Williams M. J., van Rij A. M., Jones G. T., Trabetti E., Malerba G., Pignatti P. F., Boner A., Pescollderungg L., Girelli D., Olivieri O., Martinelli N., Ludviksson B. R., Ludviksdottir D., Eyjolfsson G. I., Arnar D., Thorgeirsson G., Deichmann K., Thompson P. J., Wjst M., Hall I. P., Postma D. S., Gislason T., Gulcher J., Kong A., Jonsdottir I., Thorsteinsdottir U., Stefansson K. (2009). Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet* **41**(3): 342-347.

Hansson G. K. (2005). Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med* **352**(16): 1685-1695.

Harismendy O., Notani D., Song X., Rahim N. G., Tanasa B., Heintzman N., Ren B., Fu X. D., Topol E. J., Rosenfeld M. G., Frazer K. A. (2011). 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. *Nature* **470**(7333): 264-268.

Hawe E., Talmud P. J., Miller G. J., Humphries S. E. (2003). Family history is a coronary heart disease risk factor in the Second Northwick Park Heart Study. *Ann Hum Genet* **67**(Pt 2): 97-106.

Havulinna A. S., Kettunen J., Ukkola O., Osmond C., Eriksson J. G., Kesaniemi Y. A., Jula A., Peltonen L., Kontula K., Salomaa V., Newton-Cheh C. (2013). A Blood Pressure Genetic Risk Score Is a Significant Predictor of Incident Cardiovascular Events in 32 669 Individuals. *Hypertension* **61**(5): 987-994.

Havulinna A. S., Paakkonen R., Karvonen M., Salomaa V. (2008). Geographic patterns of incidence of ischemic stroke and acute myocardial infarction in Finland during 1991-2003. *Ann Epidemiol* **18**(3): 206-213.

Heart Protection Study Collaborative Group (2002). MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* **360**(9326): 7-22.

Helgadottir A., Thorleifsson G., Manolescu A., Gretarsdottir S., Blondal T., Jonasdottir A., Sigurdsson A., Baker A., Palsson A., Masson G., Gudbjartsson D. F., Magnusson K. P., Andersen K., Levey A. I., Backman V. M., Matthiasdottir S., Jonsdottir T., Palsson S., Einarsdottir H., Gunnarsdottir S., Gylfason A., Vaccarino V., Hooper W. C., Reilly M. P., Granger C. B., Austin H., Rader D. J., Shah S. H., Quyyumi A. A., Gulcher J. R., Thorgeirsson G., Thorsteinsdottir U., Kong A., Stefansson K. (2007). A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* **316**(5830): 1491-1493.

Hindorff L. A., MacArthur J., Morales J., Junkins H. A., Hall P. N., Klemm A. K., Manolio T. A. (2013). A Catalog of Published Genome-Wide Association Studies. Retrieved September 10, 2013, from www.genome.gov/gwastudies.

Hippisley-Cox J., Coupland C., Vinogradova Y., Robson J., May M., Brindle P. (2007). Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ* **335**(7611): 136.

Holman R. L., Mc G. H., Jr., Strong J. P., Geer J. C. (1958). The natural history of atherosclerosis: the early aortic lesions as seen in New Orleans in the middle of the of the 20th century. *Am J Pathol* **34**(2): 209-235.

Holmes M. V., Harrison S., Talmud P. J., Hingorani A. D., Humphries S. E. (2011). Utility of genetic determinants of lipids and cardiovascular events in assessing risk. *Nat Rev Cardiol* **8**(4): 207-221.

Hosmer D., Lemeshow S. (1999). Applied Survival Analysis: Regression Modelling of Time to Event Data. New York, Wiley.

Hughes M. F., Saarela O., Stritzke J., Kee F., Silander K., Klopp N., Kontto J., Karvanen J., Willenborg C., Salomaa V., Virtamo J., Amouyel P., Arveiler D., Ferrieres J., Wiklund P. G., Baumert J., Thorand B., Diemert P., Tregouet D. A., Hengstenberg C., Peters A., Evans A., Koenig W., Erdmann J., Samani N. J., Kuulasmaa K., Schunkert H. (2012). Genetic markers enhance coronary risk prediction in men: the MORGAM prospective cohorts. *PLoS One* **7**(7): e40922.

Huxley R. R., Woodward M. (2011). Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. *Lancet* **378**(9799): 1297-1305.

International HapMap Consortium (2003). The International HapMap Project. *Nature* **426**(6968): 789-796.

Ioannidis J. P., Thomas G., Daly M. J. (2009). Validating, augmenting and refining genome-wide association signals. *Nat Rev Genet* **10**(5): 318-329.

Isaacs A., Willems S. M., Bos D., Dehghan A., Hofman A., Ikram M. A., Uitterlinden A. G., Oostra B. A., Franco O. H., Witteman J. C., van Duijn C. M. (2013). Risk scores of common genetic variants for lipid levels influence atherosclerosis and incident coronary heart disease. *Arterioscler Thromb Vasc Biol* **33**(9): 2233-2239.

Jerntorp P., Berglund G. (1992). Stroke registry in Malmo, Sweden. *Stroke* **23**(3): 357-361.

Johansen C. T., Wang J., Lanktree M. B., Cao H., McIntyre A. D., Ban M. R., Martins R. A., Kennedy B. A., Hassell R. G., Visser M. E., Schwartz S. M., Voight B. F., Elosua R., Salomaa V., O'Donnell C. J., Dallinga-Thie G. M., Anand S. S., Yusuf S., Huff M. W., Kathiresan S., Hegele R. A. (2010). Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet* **42**(8): 684-687.

Jousilahti P., Puska P., Vartiainen E., Pekkanen J., Tuomilehto J. (1996). Parental history of premature coronary heart disease: an independent risk factor of myocardial infarction. *J Clin Epidemiol* **49**(5): 497-503.

Juonala M., Viikari J. S., Hutri-Kahonen N., Pietikainen M., Jokinen E., Taittonen L., Marniemi J., Ronnemaa T., Raitakari O. T. (2004). The 21-year follow-up of the Cardiovascular

Risk in Young Finns Study: risk factor levels, secular trends and east-west difference. *J Intern Med* **255**(4): 457-468.

Kannel W. B., Dawber T. R., Kagan A., Revotskie N., Stokes J., 3rd (1961). Factors of risk in the development of coronary heart disease--six year follow-up experience. The Framingham Study. *Ann Intern Med* **55**: 33-50.

Kannel W. B., LeBauer E. J., Dawber T. R., McNamara P. M. (1967). Relation of body weight to development of coronary heart disease. The Framingham study. *Circulation* **35**(4): 734-744.

Kaptoge S., Di Angelantonio E., Pennells L., Wood A. M., White I. R., Gao P., Walker M., Thompson A., Sarwar N., Caslake M., Butterworth A. S., Amouyel P., Assmann G., Bakker S. J., Barr E. L., Barrett-Connor E., Benjamin E. J., Bjorkelund C., Brenner H., Brunner E., Clarke R., Cooper J. A., Cremer P., Cushman M., Dagenais G. R., D'Agostino R. B., Sr., Dankner R., Davey-Smith G., Deeg D., Dekker J. M., Engstrom G., Folsom A. R., Fowkes F. G., Gallacher J., Gaziano J. M., Giampaoli S., Gillum R. F., Hofman A., Howard B. V., Ingelsson E., Iso H., Jorgensen T., Kiechl S., Kitamura A., Kiyohara Y., Koenig W., Kromhout D., Kuller L. H., Lawlor D. A., Meade T. W., Nissinen A., Nordestgaard B. G., Onat A., Panagiotakos D. B., Psaty B. M., Rodriguez B., Rosengren A., Salomaa V., Kauhanen J., Salonen J. T., Shaffer J. A., Shea S., Ford I., Stehouwer C. D., Strandberg T. E., Tipping R. W., Tosetto A., Wassertheil-Smoller S., Wennberg P., Westendorp R. G., Whincup P. H., Wilhelmsen L., Woodward M., Lowe G. D., Wareham N. J., Khaw K. T., Sattar N., Packard C. J., Gudnason V., Ridker P. M., Pepys M. B., Thompson S. G., Danesh J. (2012). C-reactive protein, fibrinogen, and cardiovascular disease prediction. *N Engl J Med* **367**(14): 1310-1320.

Kathiresan S., Melander O., Guiducci C., Surti A., Burtt N. P., Rieder M. J., Cooper G. M., Roos C., Voight B. F., Havulinna A. S., Wahlstrand B., Hedner T., Corella D., Tai E. S., Ordovas J. M., Berglund G., Vartiainen E., Jousilahti P., Hedblad B., Taskinen M. R., Newton-Cheh C., Salomaa V., Peltonen L., Groop L., Altshuler D. M., Orho-Melander M. (2008). Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* **40**(2): 189-197.

Kathiresan S., Srivastava D. (2012). Genetics of human cardiovascular disease. *Cell* **148**(6): 1242-1257.

Kathiresan S., Willer C. J., Peloso G. M., Demissie S., Musunuru K., Schadt E. E., Kaplan L., Bennett D., Li Y., Tanaka T., Voight B. F., Bonnycastle L. L., Jackson A. U., Crawford G., Surti A., Guiducci C., Burtt N. P., Parish S., Clarke R., Zelenika D., Kubalanza K. A., Morken M. A., Scott L. J., Stringham H. M., Galan P., Swift A. J., Kuusisto J., Bergman R. N., Sundvall J., Laakso M., Ferrucci L., Scheet P., Sanna S., Uda M., Yang Q., Lunetta K. L., Dupuis J., de Bakker P. I., O'Donnell C. J., Chambers J. C., Kooner J. S., Hercberg S., Meneton P., Lakatta E. G., Scuteri A., Schlessinger D., Tuomilehto J., Collins F. S., Groop L., Altshuler D., Collins R., Lathrop G. M., Melander O., Salomaa V., Peltonen L., Orho-Melander M., Ordovas J. M., Boehnke M., Abecasis G. R., Mohlke K. L., Cupples L. A. (2009a). Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* **41**(1): 56-65.

Kathiresan S., Voight B. F., Purcell S., Musunuru K., Ardissino D., Mannucci P. M., Anand S., Engert J. C., Samani N. J., Schunkert H., Erdmann J., Reilly M. P., Rader D. J., Morgan T., Spertus J. A., Stoll M., Girelli D., McKeown P. P., Patterson C. C., Siscovick D. S., O'Donnell C. J., Elosua R., Peltonen L., Salomaa V., Schwartz S. M., Melander O., Altshuler D., Merlini P. A., Berzuini C., Bernardinelli L., Peyvandi F., Tubaro M., Celli P., Ferrario M., Fetiveau R., Marziliano N., Casari G., Galli M., Ribichini F., Rossi M., Bernardi F., Zonzin P., Piazza A., Yee J., Friedlander Y., Marrugat J., Lucas G., Subirana I., Sala J., Ramos R., Meigs J. B., Williams G., Nathan D. M., MacRae C. A., Havulinna A. S., Berglund G., Hirschhorn J. N., Asselta R., Duga S., Spreafico M., Daly M. J., Nemesh J., Korn J. M., McCarroll S. A., Surti A., Guiducci C., Gianniny L., Mirel D., Parkin M., Burtt N., Gabriel S. B., Thompson J. R., Braund P. S., Wright B. J., Balmforth

A. J., Ball S. G., Hall A., Linsel-Nitschke P., Lieb W., Ziegler A., Konig I., Hengstenberg C., Fischer M., Stark K., Grosshennig A., Preuss M., Wichmann H. E., Schreiber S., Ouwehand W., Deloukas P., Scholz M., Cambien F., Li M., Chen Z., Wilensky R., Matthai W., Qasim A., Hakonarson H. H., Devaney J., Burnett M. S., Pichard A. D., Kent K. M., Satler L., Lindsay J. M., Waksman R., Knouff C. W., Waterworth D. M., Walker M. C., Mooser V., Epstein S. E., Scheffold T., Berger K., Huge A., Martinelli N., Olivieri O., Corrocher R., McKeown P., Erdmann E., Konig I. R., Holm H., Thorleifsson G., Thorsteinsdottir U., Stefansson K., Do R., Xie C., Siscovick D. (2009b). Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet* **41**(3): 334-341.

Kavousi M., Elias-Smale S., Rutten J. H., Leening M. J., Vliegenthart R., Verwoert G. C., Krestin G. P., Oudkerk M., de Maat M. P., Leebeek F. W., Mattace-Raso F. U., Lindemans J., Hofman A., Steyerberg E. W., van der Lugt A., van den Meiracker A. H., Witteman J. C. (2012). Evaluation of newer risk markers for coronary heart disease risk classification: a cohort study. *Ann Intern Med* **156**(6): 438-444.

Kearney P. M., Blackwell L., Collins R., Keech A., Simes J., Peto R., Armitage J., Baigent C. (2008). Efficacy of cholesterol-lowering therapy in 18,686 people with diabetes in 14 randomised trials of statins: a meta-analysis. *Lancet* **371**(9607): 117-125.

Kettunen J., Tukiainen T., Sarin A. P., Ortega-Alonso A., Tikkanen E., Lyytikainen L. P., Kangas A. J., Soininen P., Wurtz P., Silander K., Dick D. M., Rose R. J., Savolainen M. J., Viikari J., Kahonen M., Lehtimaki T., Pietilainen K. H., Inouye M., McCarthy M. I., Jula A., Eriksson J., Raitakari O. T., Salomaa V., Kaprio J., Jarvelin M. R., Peltonen L., Perola M., Freimer N. B., Ala-Korpela M., Palotie A., Ripatti S. (2012). Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet* **44**(3): 269-276.

Kilpeläinen T. O., Zillikens M. C., Stancakova A., Finucane F. M., Ried J. S., Langenberg C., Zhang W., Beckmann J. S., Luan J., Vandenput L., Styrkarsdottir U., Zhou Y., Smith A. V., Zhao J. H., Amin N., Vedantam S., Shin S. Y., Haritunians T., Fu M., Feitosa M. F., Kumari M., Halldorsson B. V., Tikkanen E., Mangino M., Hayward C., Song C., Arnold A. M., Aulchenko Y. S., Oostra B. A., Campbell H., Cupples L. A., Davis K. E., Doring A., Eiriksdottir G., Estrada K., Fernandez-Real J. M., Garcia M., Gieger C., Glazer N. L., Guiducci C., Hofman A., Humphries S. E., Isomaa B., Jacobs L. C., Jula A., Karasik D., Karlsson M. K., Khaw K. T., Kim L. J., Kivimaki M., Klopp N., Kuhnel B., Kuusisto J., Liu Y., Ljunggren O., Lorentzon M., Luben R. N., McKnight B., Mellstrom D., Mitchell B. D., Mooser V., Moreno J. M., Mannisto S., O'Connell J. R., Pascoe L., Peltonen L., Peral B., Perola M., Psaty B. M., Salomaa V., Savage D. B., Semple R. K., Skaric-Juric T., Sigurdsson G., Song K. S., Spector T. D., Syvanen A. C., Talmud P. J., Thorleifsson G., Thorsteinsdottir U., Uitterlinden A. G., van Duijn C. M., Vidal-Puig A., Wild S. H., Wright A. F., Clegg D. J., Schadt E., Wilson J. F., Rudan I., Ripatti S., Borecki I. B., Shuldiner A. R., Ingelsson E., Jansson J. O., Kaplan R. C., Gudnason V., Harris T. B., Groop L., Kiel D. P., Rivadeneira F., Walker M., Barroso I., Vollenweider P., Waeber G., Chambers J. C., Kooner J. S., Soranzo N., Hirschhorn J. N., Stefansson K., Wichmann H. E., Ohlsson C., O'Rahilly S., Wareham N. J., Speliotes E. K., Fox C. S., Laakso M., Loos R. J. (2011). Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. *Nat Genet* **43**(8): 753-760.

Kivimäki M., Nyberg S. T., Batty G. D., Fransson E. I., Heikkila K., Alfredsson L., Bjorner J. B., Borritz M., Burr H., Casini A., Clays E., De Bacquer D., Dragano N., Ferrie J. E., Geuskens G. A., Goldberg M., Hamer M., Hooftman W. E., Houtman I. L., Joensuu M., Jokela M., Kittel F., Knutsson A., Koskenvuo M., Koskinen A., Kouvonen A., Kumari M., Madsen I. E., Marmot M. G., Nielsen M. L., Nordin M., Oksanen T., Pentti J., Rugulies R., Salo P., Siegrist J., Singh-Manoux A., Suominen S. B., Vaananen A., Vahtera J., Virtanen M., Westerholm P. J., Westerlund H., Zins M., Steptoe A., Theorell

T. (2012). Job strain as a risk factor for coronary heart disease: a collaborative meta-analysis of individual participant data. *Lancet* **380**(9852): 1491-1497.

Klein R. J., Zeiss C., Chew E. Y., Tsai J. Y., Sackler R. S., Haynes C., Henning A. K., SanGiovanni J. P., Mane S. M., Mayne S. T., Bracken M. B., Ferris F. L., Ott J., Barnstable C., Hoh J. (2005). Complement factor H polymorphism in age-related macular degeneration. *Science* **308**(5720): 385-389.

Laatikainen T., Critchley J., Vartiainen E., Salomaa V., Ketonen M., Capewell S. (2005). Explaining the decline in coronary heart disease mortality in Finland between 1982 and 1997. *Am J Epidemiol* **162**(8): 764-773.

Laatikainen T., Pajunen P., Pääkkönen R., Keskimäki I., Hämäläinen H., Rintanen H., Niemi M., Moltchanov V., Salomaa V. (2009). National Cardiovascular Disease Register, statistical database. Retrieved December 27, 2012, from http://www.ktl.fi/portal/english/public_health_monitoring___promotion/registers/cardiovascular_diseases.

Lagor W. R., Millar J. S. (2010). Overview of the LDL receptor: relevance to cholesterol metabolism and future approaches for the treatment of coronary heart disease. *J Receptor Ligand Channel Res* **3**: 1-14.

Lander E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D., Harris K., Heaford A., Howland J., Kann L., Lehoczky J., LeVine R., McEwan P., McKernan K., Meldrim J., Mesirov J. P., Miranda C., Morris W., Naylor J., Raymond C., Rosetti M., Santos R., Sheridan A., Sougnez C., Stange-Thomann N., Stojanovic N., Subramanian A., Wyman D., Rogers J., Sulston J., Ainscough R., Beck S., Bentley D., Burton J., Clee C., Carter N., Coulson A., Deadman R., Deloukas P., Dunham A., Dunham I., Durbin R., French L., Grafham D., Gregory S., Hubbard T., Humphray S., Hunt A., Jones M., Lloyd C., McMurray A., Matthews L., Mercer S., Milne S., Mullikin J. C., Mungall A., Plumb R., Ross M., Shownkeen R., Sims S., Waterston R. H., Wilson R. K., Hillier L. W., McPherson J. D., Marra M. A., Mardis E. R., Fulton L. A., Chinwalla A. T., Pepin K. H., Gish W. R., Chissoe S. L., Wendl M. C., Delehaunty K. D., Miner T. L., Delehaunty A., Kramer J. B., Cook L. L., Fulton R. S., Johnson D. L., Minx P. J., Clifton S. W., Hawkins T., Branscomb E., Predki P., Richardson P., Wenning S., Slezak T., Doggett N., Cheng J. F., Olsen A., Lucas S., Elkin C., Uberbacher E., Frazier M., Gibbs R. A., Muzny D. M., Scherer S. E., Bouck J. B., Sodergren E. J., Worley K. C., Rives C. M., Gorrell J. H., Metzker M. L., Naylor S. L., Kucherlapati R. S., Nelson D. L., Weinstock G. M., Sakaki Y., Fujiyama A., Hattori M., Yada T., Toyoda A., Itoh T., Kawagoe C., Watanabe H., Totoki Y., Taylor T., Weissenbach J., Heilig R., Saurin W., Artiguenave F., Brottier P., Bruls T., Pelletier E., Robert C., Wincker P., Smith D. R., Doucette-Stamm L., Rubenfield M., Weinstock K., Lee H. M., Dubois J., Rosenthal A., Platzer M., Nyakatura G., Taudien S., Rump A., Yang H., Yu J., Wang J., Huang G., Gu J., Hood L., Rowen L., Madan A., Qin S., Davis R. W., Federspiel N. A., Abola A. P., Proctor M. J., Myers R. M., Schmutz J., Dickson M., Grimwood J., Cox D. R., Olson M. V., Kaul R., Shimizu N., Kawasaki K., Minoshima S., Evans G. A., Athanasiou M., Schultz R., Roe B. A., Chen F., Pan H., Ramser J., Lehrach H., Reinhardt R., McCombie W. R., de la Bastide M., Dedhia N., Blocker H., Hornischer K., Nordsiek G., Agarwala R., Aravind L., Bailey J. A., Bateman A., Batzoglou S., Birney E., Bork P., Brown D. G., Burge C. B., Cerutti L., Chen H. C., Church D., Clamp M., Copley R. R., Doerks T., Eddy S. R., Eichler E. E., Furey T. S., Galagan J., Gilbert J. G., Harmon C., Hayashizaki Y., Haussler D., Hermjakob H., Hokamp K., Jang W., Johnson L. S., Jones T. A., Kasif S., Kaspryzk A., Kennedy S., Kent W. J., Kitts P., Koonin E. V., Korf I., Kulp D., Lancet D., Lowe T. M., McLysaght A., Mikkelsen T., Moran J. V., Mulder N., Pollara V. J., Ponting C. P., Schuler G., Schultz J., Slater G., Smit A. F., Stupka E., Szustakowski J., Thierry-Mieg D., Thierry-Mieg J., Wagner L., Wallis J., Wheeler R., Williams A., Wolf Y. I., Wolfe K. H., Yang S. P., Yeh R. F., Collins F., Guyer M. S., Peterson J., Felsenfeld A., Wetterstrand K. A.,

Patrinos A., Morgan M. J., de Jong P., Catanese J. J., Osoegawa K., Shizuya H., Choi S., Chen Y. J. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**(6822): 860-921.

Lauer R. M., Connor W. E., Leaverton P. E., Reiter M. A., Clarke W. R. (1975). Coronary heart disease risk factors in school children: the Muscatine study. *J Pediatr* **86**(5): 697-706.

Laurila P. P., Surakka I., Sarin A. P., Yetukuri L., Hyotylainen T., Soderlund S., Naukkarinen J., Tang J., Kettunen J., Mirel D. B., Soronen J., Lehtimaki T., Ruokonen A., Ehnholm C., Eriksson J. G., Salomaa V., Jula A., Raitakari O. T., Jarvelin M. R., Palotie A., Peltonen L., Oresic M., Jauhiainen M., Taskinen M. R., Ripatti S. (2013). Genomic, Transcriptomic, and Lipidomic Profiling Highlights the Role of Inflammation in Individuals With Low High-density Lipoprotein Cholesterol. *Arterioscler Thromb Vasc Biol* **33**(4): 847-857.

Law M. R., Morris J. K., Wald N. J. (1997). Environmental tobacco smoke exposure and ischaemic heart disease: an evaluation of the evidence. *BMJ* **315**(7114): 973-980.

Levy D., DeStefano A. L., Larson M. G., O'Donnell C. J., Lifton R. P., Gavras H., Cupples L. A., Myers R. H. (2000). Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the framingham heart study. *Hypertension* **36**(4): 477-483.

Levy D., Ehret G. B., Rice K., Verwoert G. C., Launer L. J., Dehghan A., Glazer N. L., Morrison A. C., Johnson A. D., Aspelund T., Aulchenko Y., Lumley T., Kottgen A., Vasan R. S., Rivadeneira F., Eiriksdottir G., Guo X., Arking D. E., Mitchell G. F., Mattace-Raso F. U., Smith A. V., Taylor K., Scharpf R. B., Hwang S. J., Sijbrands E. J., Bis J., Harris T. B., Ganesh S. K., O'Donnell C. J., Hofman A., Rotter J. I., Coresh J., Benjamin E. J., Uitterlinden A. G., Heiss G., Fox C. S., Witteman J. C., Boerwinkle E., Wang T. J., Gudnason V., Larson M. G., Chakravarti A., Psaty B. M., van Duijn C. M. (2009). Genome-wide association study of blood pressure and hypertension. *Nat Genet* **41**(6): 677-687.

Li S., Chen W., Srinivasan S. R., Bond M. G., Tang R., Urbina E. M., Berenson G. S. (2003). Childhood cardiovascular risk factors and carotid vascular changes in adulthood: the Bogalusa Heart Study. *JAMA* **290**(17): 2271-2276.

Libby P., Ridker P. M., Hansson G. K. (2011). Progress and challenges in translating the biology of atherosclerosis. *Nature* **473**(7347): 317-325.

Lieb W., Jansen H., Loley C., Pencina M. J., Nelson C. P., Newton-Cheh C., Kathiresan S., Reilly M. P., Assimes T. L., Boerwinkle E., Hall A. S., Hengstenberg C., Laaksonen R., McPherson R., Thorsteinsdottir U., Ziegler A., Peters A., Thompson J. R., Konig I. R., Erdmann J., Samani N. J., Vasan R. S., Schunkert H. (2013). Genetic Predisposition to Higher Blood Pressure Increases Coronary Artery Disease Risk. *Hypertension* **61**(5): 995-1001.

Littell R. C., Pendergast J., Natarajan R. (2000). Modelling covariance structure in the analysis of repeated measures data. *Stat Med* **19**(13): 1793-1819.

Lloyd-Jones D. M. (2010). Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation* **121**(15): 1768-1777.

Lloyd-Jones D. M., Larson M. G., Beiser A., Levy D. (1999). Lifetime risk of developing coronary heart disease. *Lancet* **353**(9147): 89-92.

Lloyd-Jones D. M., Nam B. H., D'Agostino R. B., Sr., Levy D., Murabito J. M., Wang T. J., Wilson P. W., O'Donnell C. J. (2004). Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA* **291**(18): 2204-2211.

Lorell B. H., Carabello B. A. (2000). Left ventricular hypertrophy: pathogenesis, detection, and prognosis. *Circulation* **102**(4): 470-479.

Lu X., Wang L., Chen S., He L., Yang X., Shi Y., Cheng J., Zhang L., Gu C. C., Huang J., Wu T., Ma Y., Li J., Cao J., Chen J., Ge D., Fan Z., Li Y., Zhao L., Li H., Zhou X., Chen L., Liu D., Duan X., Hao Y., Lu F., Liu Z., Yao C., Shen C., Pu X., Yu L., Fang X., Xu L., Mu J.,

Wu X., Zheng R., Wu N., Zhao Q., Liu X., Wang M., Yu D., Hu D., Ji X., Guo D., Sun D., Wang Q., Yang Y., Liu F., Mao Q., Liang X., Ji J., Chen P., Mo X., Li D., Chai G., Tang Y., Li X., Du Z., Dou C., Yang Z., Meng Q., Wang D., Wang R., Yang J., Schunkert H., Samani N. J., Kathiresan S., Reilly M. P., Erdmann J., Peng X., Chen R., Qiang B., Gu D. (2012). Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. *Nat Genet* **44**(8): 890-894.

Luepker R. V., Apple F. S., Christenson R. H., Crow R. S., Fortmann S. P., Goff D., Goldberg R. J., Hand M. M., Jaffe A. S., Julian D. G., Levy D., Manolio T., Mendis S., Mensah G., Pajak A., Prineas R. J., Reddy K. S., Roger V. L., Rosamond W. D., Shahar E., Sharrett A. R., Sorlie P., Tunstall-Pedoe H. (2003). Case definitions for acute coronary heart disease in epidemiology and clinical research studies: a statement from the AHA Council on Epidemiology and Prevention; AHA Statistics Committee; World Heart Federation Council on Epidemiology and Prevention; the European Society of Cardiology Working Group on Epidemiology and Prevention; Centers for Disease Control and Prevention; and the National Heart, Lung, and Blood Institute. *Circulation* **108**(20): 2543-2549.

Lusis A. J. (2000). Atherosclerosis. *Nature* **407**(6801): 233-241.

Mahoney L. T., Burns T. L., Stanford W., Thompson B. H., Witt J. D., Rost C. A., Lauer R. M. (1996). Coronary risk factors measured in childhood and young adult life are associated with coronary artery calcification in young adults: the Muscatine Study. *J Am Coll Cardiol* **27**(2): 277-284.

Manolio T. A., Collins F. S., Cox N. J., Goldstein D. B., Hindorff L. A., Hunter D. J., McCarthy M. I., Ramos E. M., Cardon L. R., Chakravarti A., Cho J. H., Guttmacher A. E., Kong A., Kruglyak L., Mardis E., Rotimi C. N., Slatkin M., Valle D., Whittemore A. S., Boehnke M., Clark A. G., Eichler E. E., Gibson G., Haines J. L., Mackay T. F., McCarroll S. A., Visscher P. M. (2009). Finding the missing heritability of complex diseases. *Nature* **461**(7265): 747-753.

Maraganore D. M., de Andrade M., Lesnick T. G., Strain K. J., Farrer M. J., Rocca W. A., Pant P. V., Frazer K. A., Cox D. R., Ballinger D. G. (2005). High-resolution whole-genome association study of Parkinson disease. *Am J Hum Genet* **77**(5): 685-693.

Marchini J., Howie B. (2010). Genotype imputation for genome-wide association studies. *Nat Rev Genet* **11**(7): 499-511.

McNamara J. J., Molot M. A., Stremple J. F., Cutting R. T. (1971). Coronary artery disease in combat casualties in Vietnam. *JAMA* **216**(7): 1185-1187.

McPherson R., Pertsemlidis A., Kavaslar N., Stewart A., Roberts R., Cox D. R., Hinds D. A., Pennacchio L. A., Tybjaerg-Hansen A., Folsom A. R., Boerwinkle E., Hobbs H. H., Cohen J. C. (2007). A common allele on chromosome 9 associated with coronary heart disease. *Science* **316**(5830): 1488-1491.

Melander O., Newton-Cheh C., Almgren P., Hedblad B., Berglund G., Engstrom G., Persson M., Smith J. G., Magnusson M., Christensson A., Struck J., Morgenthaler N. G., Bergmann A., Pencina M. J., Wang T. J. (2009). Novel and conventional biomarkers for prediction of incident cardiovascular events in the community. *JAMA* **302**(1): 49-57.

Mirzaei M., Truswell A. S., Taylor R., Leeder S. R. (2009). Coronary heart disease epidemics: not all the same. *Heart* **95**(9): 740-746.

Moore K. J., Tabas I. (2011). Macrophages in the pathogenesis of atherosclerosis. *Cell* **145**(3): 341-355.

Morrow D. A., de Lemos J. A. (2007). Benchmarks for the assessment of novel cardiovascular biomarkers. *Circulation* **115**(8): 949-952.

Muller-Nordhorn J., Binting S., Roll S., Willich S. N. (2008). An update on regional variation in cardiovascular mortality within Europe. *Eur Heart J* **29**(10): 1316-1326.

National Institute for Health and Care Excellence. (2008). Statins for the prevention of cardiovascular events. Retrieved August 7, 2012, from http://www.nice.org.uk/nicemedia/pdf/TA094guidance.pdf.

National Institutes of Health. National Human Genome Research Institute. (2012). Talking Glossary of Genetic Terms. Retrieved November 15, 2012, from http://www.genome.gov/Glossary/index.cfm?id=1.

National Public Health Institute. (2008). METHODOLOGY REPORT. Health 2000 Survey. Retrieved November 26, 2012, from http://www.terveys2000.fi/doc/methodologyrep.pdf.

Newton-Cheh C., Johnson T., Gateva V., Tobin M. D., Bochud M., Coin L., Najjar S. S., Zhao J. H., Heath S. C., Eyheramendy S., Papadakis K., Voight B. F., Scott L. J., Zhang F., Farrall M., Tanaka T., Wallace C., Chambers J. C., Khaw K. T., Nilsson P., van der Harst P., Polidoro S., Grobbee D. E., Onland-Moret N. C., Bots M. L., Wain L. V., Elliott K. S., Teumer A., Luan J., Lucas G., Kuusisto J., Burton P. R., Hadley D., McArdle W. L., Brown M., Dominiczak A., Newhouse S. J., Samani N. J., Webster J., Zeggini E., Beckmann J. S., Bergmann S., Lim N., Song K., Vollenweider P., Waeber G., Waterworth D. M., Yuan X., Groop L., Orho-Melander M., Allione A., Di Gregorio A., Guarrera S., Panico S., Ricceri F., Romanazzi V., Sacerdote C., Vineis P., Barroso I., Sandhu M. S., Luben R. N., Crawford G. J., Jousilahti P., Perola M., Boehnke M., Bonnycastle L. L., Collins F. S., Jackson A. U., Mohlke K. L., Stringham H. M., Valle T. T., Willer C. J., Bergman R. N., Morken M. A., Doring A., Gieger C., Illig T., Meitinger T., Org E., Pfeufer A., Wichmann H. E., Kathiresan S., Marrugat J., O'Donnell C. J., Schwartz S. M., Siscovick D. S., Subirana I., Freimer N. B., Hartikainen A. L., McCarthy M. I., O'Reilly P. F., Peltonen L., Pouta A., de Jong P. E., Snieder H., van Gilst W. H., Clarke R., Goel A., Hamsten A., Peden J. F., Seedorf U., Syvanen A. C., Tognoni G., Lakatta E. G., Sanna S., Scheet P., Schlessinger D., Scuteri A., Dorr M., Ernst F., Felix S. B., Homuth G., Lorbeer R., Reffelmann T., Rettig R., Volker U., Galan P., Gut I. G., Hercberg S., Lathrop G. M., Zelenika D., Deloukas P., Soranzo N., Williams F. M., Zhai G., Salomaa V., Laakso M., Elosua R., Forouhi N. G., Volzke H., Uiterwaal C. S., van der Schouw Y. T., Numans M. E., Matullo G., Navis G., Berglund G., Bingham S. A., Kooner J. S., Connell J. M., Bandinelli S., Ferrucci L., Watkins H., Spector T. D., Tuomilehto J., Altshuler D., Strachan D. P., Laan M., Meneton P., Wareham N. J., Uda M., Jarvelin M. R., Mooser V., Melander O., Loos R. J., Elliott P., Abecasis G. R., Caulfield M., Munroe P. B. (2009). Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet* **41**(6): 666-676.

Nichols M., Townsend N., Luengo-Fernandez R., Leal J., Gray A., Scarborough P., Rayner M. (2012). European Cardiovascular Disease Statistics 2012, European Heart Network, Brussels, European Society of Cardiology, Sophia Antipolis.

Okrainec K., Banerjee D. K., Eisenberg M. J. (2004). Coronary artery disease in the developing world. *Am Heart J* **148**(1): 7-15.

Okser S., Pahikkala T., Aittokallio T. (2013). Genetic variants and their interactions in disease risk prediction - machine learning and network perspectives. *BioData Min* **6**(1): 5.

Orchard T. J., Costacou T., Kretowski A., Nesto R. W. (2006). Type 1 diabetes and coronary artery disease. *Diabetes Care* **29**(11): 2528-2538.

Pajunen P., Koukkunen H., Ketonen M., Jerkkola T., Immonen-Raiha P., Karja-Koskenkari P., Mahonen M., Niemela M., Kuulasmaa K., Palomaki P., Mustonen J., Lehtonen A., Arstila M., Vuorenmaa T., Lehto S., Miettinen H., Torppa J., Tuomilehto J., Kesaniemi Y. A., Pyorala K., Salomaa V. (2005). The validity of the Finnish Hospital Discharge Register and Causes of Death Register data on coronary heart disease. *Eur J Cardiovasc Prev Rehabil* **12**(2): 132-137.

Paynter N. P., Chasman D. I., Pare G., Buring J. E., Cook N. R., Miletich J. P., Ridker P. M. (2010). Association between a literature-based genetic risk score and cardiovascular events in women. *JAMA* **303**(7): 631-637.

Pencina M. J., D'Agostino R. B., Sr., D'Agostino R. B., Jr., Vasan R. S. (2008). Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* **27**(2): 157-172; discussion 207-112.

Pencina M. J., D'Agostino R. B., Sr., Steyerberg E. W. (2011). Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* **30**(1): 11-21.

Pepe M. S., Feng Z., Gu J. W. (2008). Comments on 'Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond' by M. J. Pencina et al., Statistics in Medicine (DOI: 10.1002/sim.2929). *Stat Med* **27**(2): 173-181.

Perk J., De Backer G., Gohlke H., Graham I., Reiner Z., Verschuren M., Albus C., Benlian P., Boysen G., Cifkova R., Deaton C., Ebrahim S., Fisher M., Germano G., Hobbs R., Hoes A., Karadeniz S., Mezzani A., Prescott E., Ryden L., Scherer M., Syvanne M., Scholte Op Reimer W. J., Vrints C., Wood D., Zamorano J. L., Zannad F., Cooney M. T., Bax J., Baumgartner H., Ceconi C., Dean V., Fagard R., Funck-Brentano C., Hasdai D., Kirchhof P., Knuuti J., Kolh P., McDonagh T., Moulin C., Popescu B. A., Sechtem U., Sirnes P. A., Tendera M., Torbicki A., Vahanian A., Windecker S., Aboyans V., Ezquerra E. A., Baigent C., Brotons C., Burell G., Ceriello A., De Sutter J., Deckers J., Del Prato S., Diener H. C., Fitzsimons D., Fras Z., Hambrecht R., Jankowski P., Keil U., Kirby M., Larsen M. L., Mancia G., Manolis A. J., McMurray J., Pajak A., Parkhomenko A., Rallidis L., Rigo F., Rocha E., Ruilope L. M., van der Velde E., Vanuzzo D., Viigimaa M., Volpe M., Wiklund O., Wolpert C. (2012). European Guidelines on cardiovascular disease prevention in clinical practice (version 2012): The Fifth 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) * Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur Heart J* **33**(13): 1637-1701.

Persson M., Hedblad B., Nelson J. J., Berglund G. (2007). Elevated Lp-PLA2 levels add prognostic information to the metabolic syndrome on incidence of cardiovascular events among middle-aged nondiabetic subjects. *Arterioscler Thromb Vasc Biol* **27**(6): 1411-1416.

Porkka K. V., Viikari J. S., Taimela S., Dahl M., Akerblom H. K. (1994). Tracking and predictiveness of serum lipid and lipoprotein measurements in childhood: a 12-year follow-up. The Cardiovascular Risk in Young Finns study. *Am J Epidemiol* **140**(12): 1096-1110.

Puska P., Koskela K., Pakarinen H., Puumalainen P., Soininen V., Tuomilehto J. (1976). The North Karelia Project: a programme for community control of cardiovascular diseases. *Scand J Soc Med* **4**(2): 57-60.

Puska P., Vartiainen E., Tuomilehto J., Salomaa V., Nissinen A. (1998). Changes in premature deaths in Finland: successful long-term prevention of cardiovascular diseases. *Bull World Health Organ* **76**(4): 419-425.

Raiko J. R., Viikari J. S., Ilmanen A., Hutri-Kahonen N., Taittonen L., Jokinen E., Pietikainen M., Jula A., Loo B. M., Marniemi J., Lehtimaki T., Kahonen M., Ronnemaa T., Raitakari O. T., Juonala M. (2010). Follow-ups of the Cardiovascular Risk in Young Finns Study in 2001 and 2007: levels and 6-year changes in risk factors. *J Intern Med* **267**(4): 370-384.

Raitakari O. T., Juonala M., Kahonen M., Taittonen L., Laitinen T., Maki-Torkko N., Jarvisalo M. J., Uhari M., Jokinen E., Ronnemaa T., Akerblom H. K., Viikari J. S. (2003). Cardiovascular risk factors in childhood and carotid artery intima-media thickness in adulthood: the Cardiovascular Risk in Young Finns Study. *JAMA* **290**(17): 2277-2283.

Raitakari O. T., Juonala M., Ronnemaa T., Keltikangas-Jarvinen L., Rasanen L., Pietikainen M., Hutri-Kahonen N., Taittonen L., Jokinen E., Marniemi J., Jula A., Telama R., Kahonen M., Lehtimaki T., Akerblom H. K., Viikari J. S. (2008). Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol* **37**(6): 1220-1226.

Reilly M. P., Li M., He J., Ferguson J. F., Stylianou I. M., Mehta N. N., Burnett M. S., Devaney J. M., Knouff C. W., Thompson J. R., Horne B. D., Stewart A. F., Assimes T. L., Wild P. S., Allayee H., Nitschke P. L., Patel R. S., Martinelli N., Girelli D., Quyyumi A. A.,

Anderson J. L., Erdmann J., Hall A. S., Schunkert H., Quertermous T., Blankenberg S., Hazen S. L., Roberts R., Kathiresan S., Samani N. J., Epstein S. E., Rader D. J. (2011). Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet* **377**(9763): 383-392.

Ridker P. M., Buring J. E., Rifai N., Cook N. R. (2007). Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. *JAMA* **297**(6): 611-619.

Ridker P. M., Hennekens C. H., Buring J. E., Rifai N. (2000). C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women. *N Engl J Med* **342**(12): 836-843.

Ridker P. M., MacFadyen J. G., Fonseca F. A., Genest J., Gotto A. M., Kastelein J. J., Koenig W., Libby P., Lorenzatti A. J., Nordestgaard B. G., Shepherd J., Willerson J. T., Glynn R. J. (2009). Number needed to treat with rosuvastatin to prevent first cardiovascular events and death among men and women with low low-density lipoprotein cholesterol and elevated high-sensitivity C-reactive protein: justification for the use of statins in prevention: an intervention trial evaluating rosuvastatin (JUPITER). *Circ Cardiovasc Qual Outcomes* **2**(6): 616-623.

Ridker P. M., Paynter N. P., Rifai N., Gaziano J. M., Cook N. R. (2008). C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. *Circulation* **118**(22): 2243-2251, 2244p following 2251.

Roberts R., McNally E. (2011). Genetic Basis for Cardiovascular Disease. Hurst's The Heart. V. Fuster, R. Walsh and R. Harrington, The McGraw-Hill Companies.

Roest A. M., Martens E. J., de Jonge P., Denollet J. (2010). Anxiety and risk of incident coronary heart disease: a meta-analysis. *J Am Coll Cardiol* **56**(1): 38-46.

Roger V. L., Go A. S., Lloyd-Jones D. M., Benjamin E. J., Berry J. D., Borden W. B., Bravata D. M., Dai S., Ford E. S., Fox C. S., Fullerton H. J., Gillespie C., Hailpern S. M., Heit J. A., Howard V. J., Kissela B. M., Kittner S. J., Lackland D. T., Lichtman J. H., Lisabeth L. D., Makuc D. M., Marcus G. M., Marelli A., Matchar D. B., Moy C. S., Mozaffarian D., Mussolino M. E., Nichol G., Paynter N. P., Soliman E. Z., Sorlie P. D., Sotoodehnia N., Turan T. N., Virani S. S., Wong N. D., Woo D., Turner M. B. (2012). Heart disease and stroke statistics--2012 update: a report from the American Heart Association. *Circulation* **125**(1): e2-e220.

Romero-Corral A., Montori V. M., Somers V. K., Korinek J., Thomas R. J., Allison T. G., Mookadam F., Lopez-Jimenez F. (2006). Association of bodyweight with total mortality and with cardiovascular events in coronary artery disease: a systematic review of cohort studies. *Lancet* **368**(9536): 666-678.

Rose G. (1981). Strategy of prevention: lessons from cardiovascular disease. *Br Med J (Clin Res Ed)* **282**(6279): 1847-1851.

Rose G. (2001). Sick individuals and sick populations. *Int J Epidemiol* **30**(3): 427-432; discussion 433-424.

Räsänen L., Wilska M., Kantero R. L., Nanto V., Ahlstrom A., Hallman N. (1978). Nutrition survey of Finnish rural children. IV. Serum cholesterol values in relation to dietary variables. *Am J Clin Nutr* **31**(6): 1050-1056.

Sacks F. M., Pfeffer M. A., Moye L. A., Rouleau J. L., Rutherford J. D., Cole T. G., Brown L., Warnica J. W., Arnold J. M., Wun C. C., Davis B. R., Braunwald E. (1996). The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events Trial investigators. *N Engl J Med* **335**(14): 1001-1009.

Sarwar N., Danesh J., Eiriksdottir G., Sigurdsson G., Wareham N., Bingham S., Boekholdt S. M., Khaw K. T., Gudnason V. (2007). Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 Western prospective studies. *Circulation* **115**(4): 450-458.

Saxena R., Voight B. F., Lyssenko V., Burtt N. P., de Bakker P. I., Chen H., Roix J. J., Kathiresan S., Hirschhorn J. N., Daly M. J., Hughes T. E., Groop L., Altshuler D., Almgren P., Florez J. C., Meyer J., Ardlie K., Bengtsson Bostrom K., Isomaa B., Lettre G., Lindblad U., Lyon H. N., Melander O., Newton-Cheh C., Nilsson P., Orho-Melander M., Rastam L., Speliotes E. K., Taskinen M. R., Tuomi T., Guiducci C., Berglund A., Carlson J., Gianniny L., Hackett R., Hall L., Holmkvist J., Laurila E., Sjogren M., Sterner M., Surti A., Svensson M., Tewhey R., Blumenstiel B., Parkin M., Defelice M., Barry R., Brodeur W., Camarata J., Chia N., Fava M., Gibbons J., Handsaker B., Healy C., Nguyen K., Gates C., Sougnez C., Gage D., Nizzari M., Gabriel S. B., Chirn G. W., Ma Q., Parikh H., Richardson D., Ricke D., Purcell S. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**(5829): 1331-1336.

Scallan C., Doonan R. J., Daskalopoulou S. S. (2010). The combined effect of hypertension and smoking on arterial stiffness. *Clin Exp Hypertens* **32**(6): 319-328.

Scandinavian Simvastatin Survival Study Group (1994). Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* **344**(8934): 1383-1389.

Schaub M. A., Boyle A. P., Kundaje A., Batzoglou S., Snyder M. (2012). Linking disease associations with regulatory information in the human genome. *Genome Res* **22**(9): 1748-1759.

Schunkert H., Konig I. R., Kathiresan S., Reilly M. P., Assimes T. L., Holm H., Preuss M., Stewart A. F., Barbalic M., Gieger C., Absher D., Aherrahrou Z., Allayee H., Altshuler D., Anand S. S., Andersen K., Anderson J. L., Ardissino D., Ball S. G., Balmforth A. J., Barnes T. A., Becker D. M., Becker L. C., Berger K., Bis J. C., Boekholdt S. M., Boerwinkle E., Braund P. S., Brown M. J., Burnett M. S., Buysschaert I., Carlquist J. F., Chen L., Cichon S., Codd V., Davies R. W., Dedoussis G., Dehghan A., Demissie S., Devaney J. M., Diemert P., Do R., Doering A., Eifert S., Mokhtari N. E., Ellis S. G., Elosua R., Engert J. C., Epstein S. E., de Faire U., Fischer M., Folsom A. R., Freyer J., Gigante B., Girelli D., Gretarsdottir S., Gudnason V., Gulcher J. R., Halperin E., Hammond N., Hazen S. L., Hofman A., Horne B. D., Illig T., Iribarren C., Jones G. T., Jukema J. W., Kaiser M. A., Kaplan L. M., Kastelein J. J., Khaw K. T., Knowles J. W., Kolovou G., Kong A., Laaksonen R., Lambrechts D., Leander K., Lettre G., Li M., Lieb W., Loley C., Lotery A. J., Mannucci P. M., Maouche S., Martinelli N., McKeown P. P., Meisinger C., Meitinger T., Melander O., Merlini P. A., Mooser V., Morgan T., Muhleisen T. W., Muhlestein J. B., Munzel T., Musunuru K., Nahrstaedt J., Nelson C. P., Nothen M. M., Olivieri O., Patel R. S., Patterson C. C., Peters A., Peyvandi F., Qu L., Quyyumi A. A., Rader D. J., Rallidis L. S., Rice C., Rosendaal F. R., Rubin D., Salomaa V., Sampietro M. L., Sandhu M. S., Schadt E., Schafer A., Schillert A., Schreiber S., Schrezenmeir J., Schwartz S. M., Siscovick D. S., Sivananthan M., Sivapalaratnam S., Smith A., Smith T. B., Snoep J. D., Soranzo N., Spertus J. A., Stark K., Stirrups K., Stoll M., Tang W. H., Tennstedt S., Thorgeirsson G., Thorleifsson G., Tomaszewski M., Uitterlinden A. G., van Rij A. M., Voight B. F., Wareham N. J., Wells G. A., Wichmann H. E., Wild P. S., Willenborg C., Witteman J. C., Wright B. J., Ye S., Zeller T., Ziegler A., Cambien F., Goodall A. H., Cupples L. A., Quertermous T., Marz W., Hengstenberg C., Blankenberg S., Ouwehand W. H., Hall A. S., Deloukas P., Thompson J. R., Stefansson K., Roberts R., Thorsteinsdottir U., O'Donnell C. J., McPherson R., Erdmann J., Samani N. J. (2011). Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* **43**(4): 333-338.

Scott L. J., Mohlke K. L., Bonnycastle L. L., Willer C. J., Li Y., Duren W. L., Erdos M. R., Stringham H. M., Chines P. S., Jackson A. U., Prokunina-Olsson L., Ding C. J., Swift A. J., Narisu N., Hu T., Pruim R., Xiao R., Li X. Y., Conneely K. N., Riebow N. L., Sprau A. G., Tong M., White P. P., Hetrick K. N., Barnhart M. W., Bark C. W., Goldstein J. L., Watkins L., Xiang F., Saramies J., Buchanan T. A., Watanabe R. M., Valle T. T., Kinnunen L., Abecasis G. R., Pugh E. W., Doheny K. F., Bergman R. N., Tuomilehto J.,

Collins F. S., Boehnke M. (2007). A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**(5829): 1341-1345.

Shepherd J., Cobbe S. M., Ford I., Isles C. G., Lorimer A. R., MacFarlane P. W., McKillop J. H., Packard C. J. (1995). Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med* **333**(20): 1301-1307.

Smith E. N., Chen W., Kahonen M., Kettunen J., Lehtimaki T., Peltonen L., Raitakari O. T., Salem R. M., Schork N. J., Shaw M., Srinivasan S. R., Topol E. J., Viikari J. S., Berenson G. S., Murray S. S. (2010). Longitudinal genome-wide association of cardiovascular disease risk factors in the Bogalusa heart study. *PLoS Genet* **6**(9).

Soria L. F., Ludwig E. H., Clarke H. R., Vega G. L., Grundy S. M., McCarthy B. J. (1989). Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. *Proc Natl Acad Sci U S A* **86**(2): 587-591.

Speed D., Hemani G., Johnson M. R., Balding D. J. (2012). Improved heritability estimation from genome-wide SNPs. *Am J Hum Genet* **91**(6): 1011-1021.

Srinivasan S. R., Frerichs R. R., Webber L. S., Berenson G. S. (1976). Serum lipoprotein profile in children from a biracial community: the Bogalusa Heart Study. *Circulation* **54**(2): 309-318.

Steyerberg E. W., Pencina M. J. (2010). Reclassification calculations for persons with incomplete follow-up. *Ann Intern Med* **152**(3): 195-196; author reply 196-197.

Sund R. (2012). Quality of the Finnish Hospital Discharge Register: A systematic review. *Scand J Public Health* **40**(6): 505-515.

Superko H. R., King S., 3rd (2008). Lipid management to reduce cardiovascular risk: a new strategy is required. *Circulation* **117**(4): 560-568; discussion 568.

Surakka I., Kristiansson K., Anttila V., Inouye M., Barnes C., Moutsianas L., Salomaa V., Daly M., Palotie A., Peltonen L., Ripatti S. (2010). Founder population-specific HapMap panel increases power in GWA studies through improved imputation accuracy and CNV tagging. *Genome Res* **20**(10): 1344-1351.

Takeuchi F., Yokota M., Yamamoto K., Nakashima E., Katsuya T., Asano H., Isono M., Nabika T., Sugiyama T., Fujioka A., Awata N., Ohnaka K., Nakatochi M., Kitajima H., Rakugi H., Nakamura J., Ohkubo T., Imai Y., Shimamoto K., Yamori Y., Yamaguchi S., Kobayashi S., Takayanagi R., Ogihara T., Kato N. (2012). Genome-wide association study of coronary artery disease in the Japanese. *Eur J Hum Genet* **20**(3): 333-340.

Teslovich T. M., Musunuru K., Smith A. V., Edmondson A. C., Stylianou I. M., Koseki M., Pirruccello J. P., Ripatti S., Chasman D. I., Willer C. J., Johansen C. T., Fouchier S. W., Isaacs A., Peloso G. M., Barbalic M., Ricketts S. L., Bis J. C., Aulchenko Y. S., Thorleifsson G., Feitosa M. F., Chambers J., Orho-Melander M., Melander O., Johnson T., Li X., Guo X., Li M., Shin Cho Y., Jin Go M., Jin Kim Y., Lee J. Y., Park T., Kim K., Sim X., Twee-Hee Ong R., Croteau-Chonka D. C., Lange L. A., Smith J. D., Song K., Hua Zhao J., Yuan X., Luan J., Lamina C., Ziegler A., Zhang W., Zee R. Y., Wright A. F., Witteman J. C., Wilson J. F., Willemsen G., Wichmann H. E., Whitfield J. B., Waterworth D. M., Wareham N. J., Waeber G., Vollenweider P., Voight B. F., Vitart V., Uitterlinden A. G., Uda M., Tuomilehto J., Thompson J. R., Tanaka T., Surakka I., Stringham H. M., Spector T. D., Soranzo N., Smit J. H., Sinisalo J., Silander K., Sijbrands E. J., Scuteri A., Scott J., Schlessinger D., Sanna S., Salomaa V., Saharinen J., Sabatti C., Ruokonen A., Rudan I., Rose L. M., Roberts R., Rieder M., Psaty B. M., Pramstaller P. P., Pichler I., Perola M., Penninx B. W., Pedersen N. L., Pattaro C., Parker A. N., Pare G., Oostra B. A., O'Donnell C. J., Nieminen M. S., Nickerson D. A., Montgomery G. W., Meitinger T., McPherson R., McCarthy M. I., McArdle W., Masson D., Martin N. G., Marroni F., Mangino M., Magnusson P. K., Lucas G., Luben R., Loos R. J., Lokki M. L., Lettre G., Langenberg C., Launer L. J., Lakatta E. G., Laaksonen R., Kyvik K. O., Kronenberg F., Konig I. R., Khaw K. T., Kaprio J., Kaplan L. M., Johansson A., Jarvelin M. R., Janssens A. C., Ingelsson E., Igl W., Kees Hovingh G., Hottenga J. J., Hofman A.,

Hicks A. A., Hengstenberg C., Heid I. M., Hayward C., Havulinna A. S., Hastie N. D., Harris T. B., Haritunians T., Hall A. S., Gyllensten U., Guiducci C., Groop L. C., Gonzalez E., Gieger C., Freimer N. B., Ferrucci L., Erdmann J., Elliott P., Ejebe K. G., Doring A., Dominiczak A. F., Demissie S., Deloukas P., de Geus E. J., de Faire U., Crawford G., Collins F. S., Chen Y. D., Caulfield M. J., Campbell H., Burtt N. P., Bonnycastle L. L., Boomsma D. I., Boekholdt S. M., Bergman R. N., Barroso I., Bandinelli S., Ballantyne C. M., Assimes T. L., Quertermous T., Altshuler D., Seielstad M., Wong T. Y., Tai E. S., Feranil A. B., Kuzawa C. W., Adair L. S., Taylor H. A., Jr., Borecki I. B., Gabriel S. B., Wilson J. G., Holm H., Thorsteinsdottir U., Gudnason V., Krauss R. M., Mohlke K. L., Ordovas J. M., Munroe P. B., Kooner J. S., Tall A. R., Hegele R. A., Kastelein J. J., Schadt E. E., Rotter J. I., Boerwinkle E., Strachan D. P., Mooser V., Stefansson K., Reilly M. P., Samani N. J., Schunkert H., Cupples L. A., Sandhu M. S., Ridker P. M., Rader D. J., van Duijn C. M., Peltonen L., Abecasis G. R., Boehnke M., Kathiresan S. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**(7307): 707-713.

Thanassoulis G., Peloso G. M., Pencina M. J., Hoffmann U., Fox C. S., Cupples L. A., Levy D., D'Agostino R. B., Hwang S. J., O'Donnell C. J. (2012). A genetic risk score is associated with incident cardiovascular disease and coronary artery calcium: the Framingham Heart Study. *Circ Cardiovasc Genet* **5**(1): 113-121.

The InterAct Consortium (2012). Long-Term Risk of Incident Type 2 Diabetes and Measures of Overall and Regional Obesity: The EPIC-InterAct Case-Cohort Study. *PLoS Med* **9**(6): e1001230.

Tobin M. D., Sheehan N. A., Scurrah K. J., Burton P. R. (2005). Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med* **24**(19): 2911-2935.

Tolonen H., Salomaa V., Torppa J., Sivenius J., Immonen-Raiha P., Lehtonen A. (2007). The validation of the Finnish Hospital Discharge Register and Causes of Death Register data on stroke diagnoses. *Eur J Cardiovasc Prev Rehabil* **14**(3): 380-385.

Traylor M., Farrall M., Holliday E. G., Sudlow C., Hopewell J. C., Cheng Y. C., Fornage M., Ikram M. A., Malik R., Bevan S., Thorsteinsdottir U., Nalls M. A., Longstreth W., Wiggins K. L., Yadav S., Parati E. A., Destefano A. L., Worrall B. B., Kittner S. J., Khan M. S., Reiner A. P., Helgadottir A., Achterberg S., Fernandez-Cadenas I., Abboud S., Schmidt R., Walters M., Chen W. M., Ringelstein E. B., O'Donnell M., Ho W. K., Pera J., Lemmens R., Norrving B., Higgins P., Benn M., Sale M., Kuhlenbaumer G., Doney A. S., Vicente A. M., Delavaran H., Algra A., Davies G., Oliveira S. A., Palmer C. N., Deary I., Schmidt H., Pandolfo M., Montaner J., Carty C., de Bakker P. I., Kostulas K., Ferro J. M., van Zuydam N. R., Valdimarsson E., Nordestgaard B. G., Lindgren A., Thijs V., Slowik A., Saleheen D., Pare G., Berger K., Thorleifsson G., Hofman A., Mosley T. H., Mitchell B. D., Furie K., Clarke R., Levi C., Seshadri S., Gschwendtner A., Boncoraglio G. B., Sharma P., Bis J. C., Gretarsdottir S., Psaty B. M., Rothwell P. M., Rosand J., Meschia J. F., Stefansson K., Dichgans M., Markus H. S. (2012). Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol* **11**(11): 951-962.

Tregouet D. A., Konig I. R., Erdmann J., Munteanu A., Braund P. S., Hall A. S., Grosshennig A., Linsel-Nitschke P., Perret C., DeSuremain M., Meitinger T., Wright B. J., Preuss M., Balmforth A. J., Ball S. G., Meisinger C., Germain C., Evans A., Arveiler D., Luc G., Ruidavets J. B., Morrison C., van der Harst P., Schreiber S., Neureuther K., Schafer A., Bugert P., El Mokhtari N. E., Schrezenmeir J., Stark K., Rubin D., Wichmann H. E., Hengstenberg C., Ouwehand W., Ziegler A., Tiret L., Thompson J. R., Cambien F., Schunkert H., Samani N. J. (2009). Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nat Genet* **41**(3): 283-285.

Tsao C. W., Preis S. R., Peloso G. M., Hwang S. J., Kathiresan S., Fox C. S., Cupples L. A., Hoffmann U., O'Donnell C. J. (2012). Relations of Long-Term and Contemporary Lipid Levels and Lipid Genetic Risk Scores With Coronary Artery Calcium in the Framingham Heart Study. *J Am Coll Cardiol* **60**(23): 2364-2371.

Vaara S., Nieminen M. S., Lokki M. L., Perola M., Pussinen P. J., Allonen J., Parkkonen O., Sinisalo J. (2012). Cohort Profile: The Corogene study. *Int J Epidemiol* **41**(5): 1265-1271.

Vaarhorst A. A., Lu Y., Heijmans B. T., Dolle M. E., Bohringer S., Putter H., Imholz S., Merry A. H., van Greevenbroek M. M., Jukema J. W., Gorgels A. P., van den Brandt P. A., Muller M., Schouten L. J., Feskens E. J., Boer J. M., Slagboom P. E. (2012). Literature-based genetic risk scores for coronary heart disease: the Cardiovascular Registry Maastricht (CAREMA) prospective cohort study. *Circ Cardiovasc Genet* **5**(2): 202-209.

Wahlstrand B., Orho-Melander M., Delling L., Kjeldsen S., Narkiewicz K., Almgren P., Hedner T., Melander O. (2009). The myocardial infarction associated CDKN2A/CDKN2B locus on chromosome 9p21 is associated with stroke independently of coronary events in patients with hypertension. *J Hypertens* **27**(4): 769-773.

Wain L. V., Verwoert G. C., O'Reilly P. F., Shi G., Johnson T., Johnson A. D., Bochud M., Rice K. M., Henneman P., Smith A. V., Ehret G. B., Amin N., Larson M. G., Mooser V., Hadley D., Dorr M., Bis J. C., Aspelund T., Esko T., Janssens A. C., Zhao J. H., Heath S., Laan M., Fu J., Pistis G., Luan J., Arora P., Lucas G., Pirastu N., Pichler I., Jackson A. U., Webster R. J., Zhang F., Peden J. F., Schmidt H., Tanaka T., Campbell H., Igl W., Milaneschi Y., Hottenga J. J., Vitart V., Chasman D. I., Trompet S., Bragg-Gresham J. L., Alizadeh B. Z., Chambers J. C., Guo X., Lehtimaki T., Kuhnel B., Lopez L. M., Polasek O., Boban M., Nelson C. P., Morrison A. C., Pihur V., Ganesh S. K., Hofman A., Kundu S., Mattace-Raso F. U., Rivadeneira F., Sijbrands E. J., Uitterlinden A. G., Hwang S. J., Vasan R. S., Wang T. J., Bergmann S., Vollenweider P., Waeber G., Laitinen J., Pouta A., Zitting P., McArdle W. L., Kroemer H. K., Volker U., Volzke H., Glazer N. L., Taylor K. D., Harris T. B., Alavere H., Haller T., Keis A., Tammesoo M. L., Aulchenko Y., Barroso I., Khaw K. T., Galan P., Hercberg S., Lathrop M., Eyheramendy S., Org E., Sober S., Lu X., Nolte I. M., Penninx B. W., Corre T., Masciullo C., Sala C., Groop L., Voight B. F., Melander O., O'Donnell C. J., Salomaa V., d'Adamo A. P., Fabretto A., Faletra F., Ulivi S., Del Greco F., Facheris M., Collins F. S., Bergman R. N., Beilby J. P., Hung J., Musk A. W., Mangino M., Shin S. Y., Soranzo N., Watkins H., Goel A., Hamsten A., Gider P., Loitfelder M., Zeginigg M., Hernandez D., Najjar S. S., Navarro P., Wild S. H., Corsi A. M., Singleton A., de Geus E. J., Willemsen G., Parker A. N., Rose L. M., Buckley B., Stott D., Orru M., Uda M., van der Klauw M. M., Zhang W., Li X., Scott J., Chen Y. D., Burke G. L., Kahonen M., Viikari J., Doring A., Meitinger T., Davies G., Starr J. M., Emilsson V., Plump A., Lindeman J. H., Hoen P. A., Konig I. R., Felix J. F., Clarke R., Hopewell J. C., Ongen H., Breteler M., Debette S., Destefano A. L., Fornage M., Mitchell G. F., Smith N. L., Holm H., Stefansson K., Thorleifsson G., Thorsteinsdottir U., Samani N. J., Preuss M., Rudan I., Hayward C., Deary I. J., Wichmann H. E., Raitakari O. T., Palmas W., Kooner J. S., Stolk R. P., Jukema J. W., Wright A. F., Boomsma D. I., Bandinelli S., Gyllensten U. B., Wilson J. F., Ferrucci L., Schmidt R., Farrall M., Spector T. D., Palmer L. J., Tuomilehto J., Pfeufer A., Gasparini P., Siscovick D., Altshuler D., Loos R. J., Toniolo D., Snieder H., Gieger C., Meneton P., Wareham N. J., Oostra B. A., Metspalu A., Launer L., Rettig R., Strachan D. P., Beckmann J. S., Witteman J. C., Erdmann J., van Dijk K. W., Boerwinkle E., Boehnke M., Ridker P. M., Jarvelin M. R., Chakravarti A., Abecasis G. R., Gudnason V., Newton-Cheh C., Levy D., Munroe P. B., Psaty B. M., Caulfield M. J., Rao D. C., Tobin M. D., Elliott P., van Duijn C. M. (2011). Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet* **43**(10): 1005-1011.

Wall J. D., Pritchard J. K. (2003). Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet* **4**(8): 587-597.

Vartiainen E., Laatikainen T., Peltonen M., Juolevi A., Mannisto S., Sundvall J., Jousilahti P., Salomaa V., Valsta L., Puska P. (2010). Thirty-five-year trends in cardiovascular risk factors in Finland. *Int J Epidemiol* **39**(2): 504-518.

Vartiainen E., Puska P., Pekkanen J., Tuomilehto J., Jousilahti P. (1994). Changes in risk factors explain changes in mortality from ischaemic heart disease in Finland. *BMJ* **309**(6946): 23-27.

Wauters E., Carruthers K. F., Buysschaert I., Dunbar D. R., Peuteman G., Belmans A., Budaj A., Van de Werf F., Lambrechts D., Fox K. A. (2012). Influence of 23 coronary artery disease variants on recurrent myocardial infarction or cardiac death: the GRACE Genetics Study. *Eur Heart J* **34**(13): 993-1001.

Webber L. S., Srinivasan S. R., Wattigney W. A., Berenson G. S. (1991). Tracking of serum lipids and lipoproteins from childhood to adulthood. The Bogalusa Heart Study. *Am J Epidemiol* **133**(9): 884-899.

Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**(7145): 661-678.

Venter J. C., Adams M. D., Myers E. W., Li P. W., Mural R. J., Sutton G. G., Smith H. O., Yandell M., Evans C. A., Holt R. A., Gocayne J. D., Amanatides P., Ballew R. M., Huson D. H., Wortman J. R., Zhang Q., Kodira C. D., Zheng X. H., Chen L., Skupski M., Subramanian G., Thomas P. D., Zhang J., Gabor Miklos G. L., Nelson C., Broder S., Clark A. G., Nadeau J., McKusick V. A., Zinder N., Levine A. J., Roberts R. J., Simon M., Slayman C., Hunkapiller M., Bolanos R., Delcher A., Dew I., Fasulo D., Flanigan M., Florea L., Halpern A., Hannenhalli S., Kravitz S., Levy S., Mobarry C., Reinert K., Remington K., Abu-Threideh J., Beasley E., Biddick K., Bonazzi V., Brandon R., Cargill M., Chandramouliswaran I., Charlab R., Chaturvedi K., Deng Z., Di Francesco V., Dunn P., Eilbeck K., Evangelista C., Gabrielian A. E., Gan W., Ge W., Gong F., Gu Z., Guan P., Heiman T. J., Higgins M. E., Ji R. R., Ke Z., Ketchum K. A., Lai Z., Lei Y., Li Z., Li J., Liang Y., Lin X., Lu F., Merkulov G. V., Milshina N., Moore H. M., Naik A. K., Narayan V. A., Neelam B., Nusskern D., Rusch D. B., Salzberg S., Shao W., Shue B., Sun J., Wang Z., Wang A., Wang X., Wang J., Wei M., Wides R., Xiao C., Yan C., Yao A., Ye J., Zhan M., Zhang W., Zhang H., Zhao Q., Zheng L., Zhong F., Zhong W., Zhu S., Zhao S., Gilbert D., Baumhueter S., Spier G., Carter C., Cravchik A., Woodage T., Ali F., An H., Awe A., Baldwin D., Baden H., Barnstead M., Barrow I., Beeson K., Busam D., Carver A., Center A., Cheng M. L., Curry L., Danaher S., Davenport L., Desilets R., Dietz S., Dodson K., Doup L., Ferriera S., Garg N., Gluecksmann A., Hart B., Haynes J., Haynes C., Heiner C., Hladun S., Hostin D., Houck J., Howland T., Ibegwam C., Johnson J., Kalush F., Kline L., Koduru S., Love A., Mann F., May D., McCawley S., McIntosh T., McMullen I., Moy M., Moy L., Murphy B., Nelson K., Pfannkoch C., Pratts E., Puri V., Qureshi H., Reardon M., Rodriguez R., Rogers Y. H., Romblad D., Ruhfel B., Scott R., Sitter C., Smallwood M., Stewart E., Strong R., Suh E., Thomas R., Tint N. N., Tse S., Vech C., Wang G., Wetter J., Williams S., Williams M., Windsor S., Winn-Deen E., Wolfe K., Zaveri J., Zaveri K., Abril J. F., Guigo R., Campbell M. J., Sjolander K. V., Karlak B., Kejariwal A., Mi H., Lazareva B., Hatton T., Narechania A., Diemer K., Muruganujan A., Guo N., Sato S., Bafna V., Istrail S., Lippert R., Schwartz R., Walenz B., Yooseph S., Allen D., Basu A., Baxendale J., Blick L., Caminha M., Carnes-Stine J., Caulk P., Chiang Y. H., Coyne M., Dahlke C., Mays A., Dombroski M., Donnelly M., Ely D., Esparham S., Fosler C., Gire H., Glanowski S., Glasser K., Glodek A., Gorokhov M., Graham K., Gropman B., Harris M., Heil J., Henderson S., Hoover J., Jennings D., Jordan C., Jordan J., Kasha J., Kagan L., Kraft C., Levitsky A., Lewis M., Liu X., Lopez J., Ma D., Majoros W., McDaniel J., Murphy S., Newman M., Nguyen T., Nguyen N., Nodell M., Pan S., Peck J., Peterson M., Rowe W., Sanders R., Scott J., Simpson M., Smith T., Sprague A., Stockwell T., Turner R., Venter E., Wang M., Wen M., Wu D., Wu M., Xia A., Zandieh A., Zhu X. (2001). The sequence of the human genome. *Science* **291**(5507): 1304-1351.

Whitlock G., Lewington S., Sherliker P., Clarke R., Emberson J., Halsey J., Qizilbash N., Collins R., Peto R. (2009). Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* **373**(9669): 1083-1096.

Wienke A., Holm N. V., Skytthe A., Yashin A. I. (2001). The heritability of mortality due to heart diseases: a correlated frailty model applied to Danish twins. *Twin Res* **4**(4): 266-274.

Viikari J. S., Juonala M., Raitakari O. T. (2006). Trends in cardiovascular risk factor levels in Finnish children and young adults from the 1970s: The Cardiovascular Risk in Young Finns Study. *Exp Clin Cardiol* **11**(2): 83-88.

Willer C. J., Sanna S., Jackson A. U., Scuteri A., Bonnycastle L. L., Clarke R., Heath S. C., Timpson N. J., Najjar S. S., Stringham H. M., Strait J., Duren W. L., Maschio A., Busonero F., Mulas A., Albai G., Swift A. J., Morken M. A., Narisu N., Bennett D., Parish S., Shen H., Galan P., Meneton P., Hercberg S., Zelenika D., Chen W. M., Li Y., Scott L. J., Scheet P. A., Sundvall J., Watanabe R. M., Nagaraja R., Ebrahim S., Lawlor D. A., Ben-Shlomo Y., Davey-Smith G., Shuldiner A. R., Collins R., Bergman R. N., Uda M., Tuomilehto J., Cao A., Collins F. S., Lakatta E., Lathrop G. M., Boehnke M., Schlessinger D., Mohlke K. L., Abecasis G. R. (2008). Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* **40**(2): 161-169.

Williams P. T. (2001). Physical fitness and activity as separate heart disease risk factors: a meta-analysis. *Med Sci Sports Exerc* **33**(5): 754-761.

Wilson P. W., D'Agostino R. B., Levy D., Belanger A. M., Silbershatz H., Kannel W. B. (1998). Prediction of coronary heart disease using risk factor categories. *Circulation* **97**(18): 1837-1847.

Wilson P. W., D'Agostino R. B., Parise H., Sullivan L., Meigs J. B. (2005). Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. *Circulation* **112**(20): 3066-3072.

Visscher P. M., Hill W. G., Wray N. R. (2008). Heritability in the genomics era--concepts and misconceptions. *Nat Rev Genet* **9**(4): 255-266.

Voight B. F., Peloso G. M., Orho-Melander M., Frikke-Schmidt R., Barbalic M., Jensen M. K., Hindy G., Holm H., Ding E. L., Johnson T., Schunkert H., Samani N. J., Clarke R., Hopewell J. C., Thompson J. F., Li M., Thorleifsson G., Newton-Cheh C., Musunuru K., Pirruccello J. P., Saleheen D., Chen L., Stewart A., Schillert A., Thorsteinsdottir U., Thorgeirsson G., Anand S., Engert J. C., Morgan T., Spertus J., Stoll M., Berger K., Martinelli N., Girelli D., McKeown P. P., Patterson C. C., Epstein S. E., Devaney J., Burnett M. S., Mooser V., Ripatti S., Surakka I., Nieminen M. S., Sinisalo J., Lokki M. L., Perola M., Havulinna A., de Faire U., Gigante B., Ingelsson E., Zeller T., Wild P., de Bakker P. I., Klungel O. H., Maitland-van der Zee A. H., Peters B. J., de Boer A., Grobbee D. E., Kamphuisen P. W., Deneer V. H., Elbers C. C., Onland-Moret N. C., Hofker M. H., Wijmenga C., Verschuren W. M., Boer J. M., van der Schouw Y. T., Rasheed A., Frossard P., Demissie S., Willer C., Do R., Ordovas J. M., Abecasis G. R., Boehnke M., Mohlke K. L., Daly M. J., Guiducci C., Burtt N. P., Surti A., Gonzalez E., Purcell S., Gabriel S., Marrugat J., Peden J., Erdmann J., Diemert P., Willenborg C., Konig I. R., Fischer M., Hengstenberg C., Ziegler A., Buysschaert I., Lambrechts D., Van de Werf F., Fox K. A., El Mokhtari N. E., Rubin D., Schrezenmeir J., Schreiber S., Schafer A., Danesh J., Blankenberg S., Roberts R., McPherson R., Watkins H., Hall A. S., Overvad K., Rimm E., Boerwinkle E., Tybjaerg-Hansen A., Cupples L. A., Reilly M. P., Melander O., Mannucci P. M., Ardissino D., Siscovick D., Elosua R., Stefansson K., O'Donnell C. J., Salomaa V., Rader D. J., Peltonen L., Schwartz S. M., Altshuler D., Kathiresan S. (2012). Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* **380**(9841): 572-580.

Voight B. F., Scott L. J., Steinthorsdottir V., Morris A. P., Dina C., Welch R. P., Zeggini E., Huth C., Aulchenko Y. S., Thorleifsson G., McCulloch L. J., Ferreira T., Grallert H., Amin N., Wu G., Willer C. J., Raychaudhuri S., McCarroll S. A., Langenberg C., Hofmann O. M.,

Dupuis J., Qi L., Segre A. V., van Hoek M., Navarro P., Ardlie K., Balkau B., Benediktsson R., Bennett A. J., Blagieva R., Boerwinkle E., Bonnycastle L. L., Bengtsson Bostrom K., Bravenboer B., Bumpstead S., Burtt N. P., Charpentier G., Chines P. S., Cornelis M., Couper D. J., Crawford G., Doney A. S., Elliott K. S., Elliott A. L., Erdos M. R., Fox C. S., Franklin C. S., Ganser M., Gieger C., Grarup N., Green T., Griffin S., Groves C. J., Guiducci C., Hadjadj S., Hassanali N., Herder C., Isomaa B., Jackson A. U., Johnson P. R., Jorgensen T., Kao W. H., Klopp N., Kong A., Kraft P., Kuusisto J., Lauritzen T., Li M., Lieverse A., Lindgren C. M., Lyssenko V., Marre M., Meitinger T., Midthjell K., Morken M. A., Narisu N., Nilsson P., Owen K. R., Payne F., Perry J. R., Petersen A. K., Platou C., Proenca C., Prokopenko I., Rathmann W., Rayner N. W., Robertson N. R., Rocheleau G., Roden M., Sampson M. J., Saxena R., Shields B. M., Shrader P., Sigurdsson G., Sparso T., Strassburger K., Stringham H. M., Sun Q., Swift A. J., Thorand B., Tichet J., Tuomi T., van Dam R. M., van Haeften T. W., van Herpt T., van Vliet-Ostaptchouk J. V., Walters G. B., Weedon M. N., Wijmenga C., Witteman J., Bergman R. N., Cauchi S., Collins F. S., Gloyn A. L., Gyllensten U., Hansen T., Hide W. A., Hitman G. A., Hofman A., Hunter D. J., Hveem K., Laakso M., Mohlke K. L., Morris A. D., Palmer C. N., Pramstaller P. P., Rudan I., Sijbrands E., Stein L. D., Tuomilehto J., Uitterlinden A., Walker M., Wareham N. J., Watanabe R. M., Abecasis G. R., Boehm B. O., Campbell H., Daly M. J., Hattersley A. T., Hu F. B., Meigs J. B., Pankow J. S., Pedersen O., Wichmann H. E., Barroso I., Florez J. C., Frayling T. M., Groop L., Sladek R., Thorsteinsdottir U., Wilson J. F., Illig T., Froguel P., van Duijn C. M., Stefansson K., Altshuler D., Boehnke M., McCarthy M. I. (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* **42**(7): 579-589.

Working group set up by the Finnish Medical Society Duodecim and the Finnish Respiratory Society. (2011). Diabetes. Current care guideline. Retrieved November 28, 2012, from www.kaypahoito.fi.

World Health Organization. (2012a). Country statistics. Retrieved September 5, 2012, from http://www.who.int/gho/countries/en/.

World Health Organization. (2012b). International Classification of Diseases (ICD). Retrieved November 28, 2012, from http://www.who.int/classifications/icd/en/.

Wray N., Visscher P. (2008). Estimating trait heritability. *Nature Education* **1**(1).

Wu L., Xi B., Zhang M., Shen Y., Zhao X., Wang T., Cheng H., Hou D., Liu G., Wang X., Mi J. (2012). A sex-specific effect of the CYP17A1 SNP rs11191548 on blood pressure in Chinese children. *J Hum Hypertens* **26**(12): 731-736.

Yusuf S., Hawken S., Ounpuu S., Dans T., Avezum A., Lanas F., McQueen M., Budaj A., Pais P., Varigos J., Lisheng L. (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* **364**(9438): 937-952.

Zaitlen N., Kraft P., Patterson N., Pasaniuc B., Bhatia G., Pollack S., Price A. L. (2013). Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genet* **9**(5): e1003520.

Zdravkovic S., Wienke A., Pedersen N. L., Marenberg M. E., Yashin A. I., De Faire U. (2002). Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. *J Intern Med* **252**(3): 247-254.

Zeggini E., Weedon M. N., Lindgren C. M., Frayling T. M., Elliott K. S., Lango H., Timpson N. J., Perry J. R., Rayner N. W., Freathy R. M., Barrett J. C., Shields B., Morris A. P., Ellard S., Groves C. J., Harries L. W., Marchini J. L., Owen K. R., Knight B., Cardon L. R., Walker M., Hitman G. A., Morris A. D., Doney A. S., McCarthy M. I., Hattersley A. T. (2007). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* **316**(5829): 1336-1341.