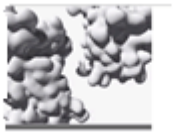
NETSIM

Bloodomics



Department of Public Health and Primary Care

UNIVERSITY OF CAMBRIDGE

Tobacco Use and The Risk of Cardiovascular Diseases In Developed and Developing Countries

Myriam Alexander

Queen's College

University of Cambridge, UK

2008-2012

This dissertation is submitted for the Degree of Doctor in Philosophy

This thesis is dedicated
to my parents
who have taught me to value intellectual knowledge
to my husband
who has supported me daily
and to my two wonderful children
who were born during the course of this PhD
and have taught me that science is like children
and requires patience, endless learning, unconditional love, formidable energy and a willingness to approach the future positively and with confidence.

SUMMARY

Background & objective: The association between cigarette smoking and the risk of cardiovascular diseases (CVD) is well established. However, the effect of other, less common, types of smoking on CVD risk, such as pipes and cigars in developed countries, remains uncertain. By contrast, in developing countries, a large panel of smokeless tobacco products are consumed alongside smoking products, with unknown effects on the risk of CVD. The aim of this thesis is to investigate the association between various forms of tobacco use with the risk of CVD in the setting of developed countries and of a developing country with a large population, Pakistan.

Data Sources: Firstly, for the investigation of cigarette, pipe and cigar smoking, the analysis was based on the Emerging Risk Factor Collaboration database. It included, in April 2011, up to 929,335 individuals with baseline information on smoking status from 135 prospective cohort studies in developed countries, who experienced 40,218 incident coronary artery disease events during an average of 14.2 years. Secondly, for the investigation of the association between chewing and dipping forms of tobacco and the risk of CVD, the analysis was based on the Pakistani Risk of Myocardial Infarction Study, which had recruited, by May 2012, a total of 7,905 first ever myocardial infarction (MI) cases and 7,458 age and sex frequency matched controls.

Results: All forms of tobacco use were significantly associated with excess risk of CHD and CVD. Current cigarette smoking was most strongly associated and produced a hazard ratio of 1.99 (95 % confidence interval: 1.86; 2.13) of MI and 1.64 (1.54; 1.75) of cerebrovascular events, compared to never-smokers, in developed countries. By contrast, compared to never-smokers, the risk of MI was 1.35 (1.20; 1.51) for current cigar use and 1.84 (1.69; 2.01) for current pipe use in developed countries. For smokeless tobacco, which was investigated in the South Asian context, odds ratios of non-fatal MI compared to never tobacco users were 1.71 (1.46; 2.00) when currently chewing *paan*, *supari* or *gutka* products, and 1.46 (1.21; 1.78) when currently dipping South Asian snuff called *naswar*. The risks associated with current tobacco use, and in particular, current smoking, were significant even at low intensities (<5 products a day) in both developed and developing countries. Quitting smoking or stopping the use of smokeless products was associated with a significantly lower excess risk of CVD worldwide. Few individuals in developing countries are ex-tobacco users. The risk of past smokers was investigated in a developed population and was shown to become not significantly different from that of never-smokers 20 years after stopping.

Conclusion: All forms of tobacco use increased CVD risk, and the highest risk was found amongst cigarette users. Neither pipe or cigar smoking nor smokeless tobacco constituted safe alternatives to cigarette smoking and current users carried a higher CHD risk than ex- and never-users. Rather than advocating alternatives to cigarette smoking on the basis that they are less harmful for CVD risk, this thesis emphasizes the need in both developed and developing countries to adopt consistent policies restricting the use of all types of tobacco use, to prevent the very high burden of tobacco related CVD worldwide.

CONTENTS

Note: Relevant tables, figures and references are provided at the end of each chapter. See p.3 for abbreviations.

PREFACE

The aim of this thesis was to investigate, in more detail than has been done before, the association between different types of tobacco use and the risk of cardiovascular diseases in developed and developing countries. The work is presented, following an **Introduction**, in two **Sections** looking at the effect of different types of tobacco use: firstly in developed countries and secondly in a developing country, taking the example of Pakistan. **Section A** contains four **Chapters** and **Section B** contains three **Chapters**. At the end, a **Discussion** reviews the main findings and considers future work.

During my doctoral studies, I have also conducted research on other topics relevant to cardiovascular diseases, including blood pressure, the chromosome 9p21.3, diet and socio-economic status, using the Pakistani Risk of Myocardial Infarction Study dataset. In addition, I have been involved in side projects regarding tobacco use, including a project researching the relationship between access to tobacco outlet and smoking abstinence in the UK. Brief description of these projects and list of publications that arose from these works are presented in the **Appendices**.

This dissertation is the result of my own work except where collaborations are specifically acknowledged in the Acknowledgments section.

ACKNOWLEDGEMENTS

Although this thesis is the result of my own work, it would not have been possible without the precious help of several people. I would like to acknowledge my funders, NETSIM Bloodomics, for offering me an EU funded Marie Curie scholarship, and Dr Willem Ouwehand for supervising and providing and inspirational leadership to all Bloodomics students. I am indebted toward Professor John Danesh for welcoming me into the Cardiovascular Epidemiology Unit initially as an intern, then as an MPhil student, and presently as a PhD student. On a day to day basis, I am grateful to my two co-supervisors, Dr Emanuele Di Angelantonio and Dr Danish Saleheen, who have guided me and encouraged me during my time as a PhD student and especially Dr Emanuele Di Angelantonio for commenting on this thesis. I was privileged to work on datasets which had already been collected, and am grateful to all the people involved in setting up the Emerging Risk Factors Collaboration (ERFC) and the Pakistan Risk of Myocardial Infarction Study (RPOMIS). I would like to thank all the members of the Cardiovascular Epidemiology Unit for offering their support when needed and providing a friendly environment. My gratitude goes in particular to fellow statisticians Dr Stephen Kaptoge and Dr Angela Wood, Dr David Wormser, Dr Rao Kondapally Seshasai, Dr Emma Heydon, Dr Sebhat Erqou, to my fellow PhD students, to the data management team and to the administrative team. I detail below the contribution of colleagues to the analyses presented in this thesis.

Introduction

I performed reviews of the published epidemiological literature using search engines such as PubMed Google Scholar and Web of Science. I organized the **Introduction** into several paragraphs, and synthesized the evidence regarding each aspect of the relationship between tobacco use and CVD risk. Dr E. Heydon and Dr D. Wormser commented helpfully.

Section A:

ERFC was a project which was started before my arrival at the Cardiovascular Epidemiology Unit and was initiated by Professor J. Danesh with the help of others (Dr N. Sarwar, Dr Di Angelantonio, Dr A. Thompson, Dr S. Thompson, Dr I. White, Dr M. Walker and S. Watson). A list of collaborators who contributed individual data is available at www.phpc.cam.ac.uk. I am a member of the Coordinating Centre of ERFC. Actual collation and cleaning of data was done by the data management team including Dr M. Walker and S. Watson. Statistical methods were developed by a team of statisticians who were members of the Coordinating Centre of ERFC to whom I am part, and were published in the *Am J Epidemiol* in 2007. I conducted all the statistical analyses presented in **Chapters 2-5,** except for the estimation of years of life lost as a result of smoking in the last paragraph of Results of **Chapter 4** which was done by Dr S Kaptoge. I wrote appropriate programming code in STATA language, calling programs developed by Dr S Kaptoge and Dr P Perry specifically for the analysis of ERFC dataset. I produced relevant tables and figures and drafted the text, being inspired by previous reports

published by ERFC (*Lancet* 2011 March, *N Engl J Med.* 2011 March, *BMJ* 2011 Feb, *Lancet* 2010, *JAMA* 2009). Dr S. Thompson, Dr A. Wood, Dr S. Kaptoge, P. Gao and Dr D. Wormser provided statistical advice. Dr E. Di Angelantonio was my daily supervisor for the analyses presented in these Chapters. Dr D. Wormser, Dr E. Heydon and S. Warnakula helpfully commented on these Chapters.

Section B:

The second dataset I worked on is PROMIS. PROMIS' principal Investigator is Dr. D. Saleheen and he is supported in Cambridge by Professor J. Danesh, and in Pakistan by a large group dedicated to the enrolment of cases and controls (http://www.cncdpk.com/projects/the-pakistan-risk-of-myocardial-infarction-study-promis.html). Data entry was performed in 6 hospitals located in urban centres across Pakistan. Checks and cleaning of this dataset were done by data manager Dr M. Walker and by myself. I conducted all the statistical analyses, wrote appropriate programming code in STATA language, produced relevant tables and figures and drafted the text. Developing programming code for these analyses was a time consuming task and I shared my codes with other members of the Cardiovascular Epidemiology Unit who were involved in the analysis of PROMIS and other datasets informally and at seminars. Dr. D. Saleheen was my main supervisor for these analyses, seconded by Dr E. Di Angelantonio. Dr D. Wormser, Dr E. Heydon and W Kee Ho commented helpfully on these Chapters.

Appendices 2-3:

Analyses were conducted at the same time as for **Section B**. Dr L. Johnson and Dr N. Naswar provided helpful comment on the dietary analyses.

Appendix 4:

The questionnaire data was provided by Dr. D. Saleheen, after cleaning done by Dr M. Walker and myself. The genetic data was generated at the Wellcome Trust Sanger Institute Cambridge, England. Analysis plan was drafted by Dr. D. Saleheen and by me, with helpful comments from Dr E. Di Angelantonio. I produced all tables and figures. The manuscript was written by me and Dr D. Saleheen, with corrections from Dr E. Di Angelantonio and Professor J. Danesh. Other co-authors and their contributions are detailed in the body of the manuscript.

ABBREVIATIONS

BMI: Body mass index
CAD: Coronary artery disease
CHD: Coronary heart disease
CI: Confidence interval
CO: Carbon monoxide
CPD: Cigarettes per day
CPS: Cancer Prevention Study
CVD: Cardiovascular diseases
DBP: Diastolic blood pressure
DALYs: Disability adjusted life years
ERFC: Emerging Risk Factors Collaboration
EU: European Union
FAV: Floating absolute variance
FCTC: Framework Convention on Tobacco Control
GWA: Genome wide association study
HDL-C: High density lipoprotein cholesterol
HR: Hazard ratios
LDL-C: Low-density lipoprotein cholesterol
MI: Myocardial infarction
OR: Odds ratio
PROMIS: Pakistan Risk of Myocardial Infarction Study
RR: Relative risks
RYO: “Roll your own” cigarettes
SBP: Systolic blood pressure
SNP: Single nucleotide polymorphism
UK: United Kingdom
UN: United Nations
USA: United States of America
WHO: World Health Organization
WHR: Waist to hip ratio

Chapter 1: INTRODUCTION

Summary

Cardiovascular disease (CVD) is the leading cause of death worldwide, responsible for over 17 million deaths globally in 2008. Approximately eight out of ten of these deaths occur in developing countries, which are currently experiencing an epidemic of non-communicable diseases. A rise in tobacco use concomitant with a reduction in infectious disease is thought to contribute to this epidemic, and tobacco has been estimated to directly cause 10% of all CVD worldwide. There are currently more than 1.1 billion smokers worldwide, and while developed countries remain at the top in terms of prevalence, with around one fifth of the total adult population smoking, nine out of ten smokers now reside in developing countries. This is due to especially high and rising prevalence in men in developing countries, with men in lower middle income countries averaging 40% smoking prevalence, whilst prevalence in developed countries has been approximately halved in men and women over the past 50 years. Alternative forms of tobacco use are increasingly being used: cigar smoking in developed countries and smokeless tobacco in developing countries. The epidemiological evidence on cigarette smoking in relation to CVD is compelling, but the relationships between CVD and other smoking types including pipe, cigars, and smokeless types of tobacco, remain uncertain. Biological mechanisms by which tobacco causes CVD are partially understood and are thought to involve nicotine, oxidant chemical, and carbon monoxide. The objective of this thesis is to address remaining epidemiological uncertainties on tobacco in relation to CVD risk using the Emerging Risk Factors Collaboration and the Pakistan Risk of Myocardial Infarction Study.

1.1 Introduction

The aim of this Chapter is to define several types of tobacco used globally and present the current prevalence of tobacco use. In addition, CVD is defined and the potential biological mechanisms of the association between CVD and tobacco use are also described. Finally, a review of the literature on the epidemiological evidence of causal associations between tobacco use and CVD is presented, which motivates the overall objectives of this thesis.

1.2 Definition of tobacco uses

The WHO Framework Convention on Tobacco Control defines tobacco products as "products entirely or partly made of the leaf tobacco as raw material which are manufactured to be used for smoking, sucking, chewing or snuffing" [1]. The main form of tobacco use in developed countries is smoking, while sucking, chewing and snuffing are common in developing countries. There is a wide variety of tobacco products and here I give examples of products found in Pakistan and more generally South Asia.

1.2.1 Smoking tobacco

1.2.1.1 Cigarettes: Cigarettes are the most popular form of tobacco use worldwide. The modern cigarette evolved from a 16th century variant of the cigar [2] and consists of a roll of tobacco wrapped in paper or a substance not containing tobacco [3]. However, tobacco companies have been inclined to blur the difference between cigars and cigarettes to avoid the heavy taxes imposed on cigarettes, and as a result the US code of federal regulation has extended the definition of cigarettes to "any roll of tobacco wrapped in any substance containing tobacco which, because of its appearance, the type of tobacco used in the filler or its packaging and labelling, is likely to be offered to or purchased by consumers as cigarettes [rather than a cigar]" [4]. The tobacco contained in cigarettes is mainly heat- or air-cured, and in a small percentage sun-cured. Air-curing involves hanging the whole tobacco plant or primed tobacco leaves in barns for 30 to 40 days, while heat-curing involves hanging the leaves of tobacco on tiers in barns where the air is gradually warmed to a temperature of 70°C to 75°C over a period of 5 to 7 days. These curing processes make the smoke acidic and its nicotinic content easier to inhale, promoting nicotine addiction [5]. Conventionally, cigarette smoke is divided into two phases: a tar phase and a gas phase [6]. The tar or particulate phase is defined as the material that is

trapped when the smoke stream is passed through the Cambridge glass-fibre filter that retains 99.9% of all particulate material with a size >0.1 μm, while the gas phase is the material that passes through the filter. Cigarette smoke comprises approximately 90% of gaseous components and 10% of tar.

1.2.1.2 Roll-your-own (RYO): Roll-your-own cigarettes are handmade cigarettes made by wrapping paper around tobacco. RYO are very popular in developing countries for cultural reasons and increasingly used in developed countries where they are sometimes used as a cheaper replacement to heavily taxed industrial cigarettes[2]. ***Bidies*** (also written *beedies* or *bidis)* are a form of RYO cigarette popular in Pakistan and other South-East Asian countries (**Figure 1.1**). They are small hand-rolled cigarettes which require more frequent puffing and pulmonary effort than cigarettes due to their non-porous wrapper of *temburini* leaf. Bidies usually contain a smaller amount of tobacco than cigarettes (0.15g to 0.25g versus 1g) and produce a smaller volume of smoke; but more frequent puffing means they produce up to 3 times more carbon monoxide and nicotine, and 2 to 3 times more tar than regular cigarettes [7].

1.2.1.3 Cigars: Cigars are made of a roll of tobacco wrapped in a leaf of tobacco or in any substance which contains tobacco [4]. The main difference between cigar and cigarette resides in the processing of tobacco, which is air cured and fermented in cigars, while it is not fermented in cigarettes [3]. Fermentation entails packing the tobacco leaves and placing them in fermentation rooms for 3 to 5 weeks. They are subsequently removed, repacked, and returned to the fermentation rooms several times to achieve the desired flavour and aroma. The fermentation process is responsible for making cigar tobacco richer than cigarettes in nitrate, carcinogens formed from nornicotine and nicotine, and nitrogen oxides. Cigars also have a higher pH than cigarettes, which increases the amount of free nicotine in the gas and solid phases of the smoke. As a result, the smoke of cigars dissolves more easily in the saliva than the smoke of cigarettes. The desired dose of nicotine is achieved without the need to inhale the smoke into the lung and high levels of dependency can be created even if the smoke is not inhaled.

1.2.1.4 Pipes: Pipes are a device consisting of a tube of wood, clay or other material with a small bowl at one end which is filled with tobacco. Tobacco for smoking in pipes is often carefully treated and blended to achieve flavour nuances not available

in other tobacco products. Similar to cigars, the smoke produced by pipes tends to have a higher pH than cigarette smoke and thus does not need to be directly inhaled, or even lit up, to sustain high levels of nicotine addiction. Owing to the relatively large quantities of tobacco put into pipes, pipe smokers are generally exposed to smoke equivalent to that from several cigarettes [2]. ***Chilum*** is a straight conical pipe, usually made of clay, in use in South and Southeast Asia for smoking tobacco, cannabis and opiates, and requiring deep pulmonary effort. Chilum is entirely filled with tobacco and is held vertically for smoking. To prevent the tobacco from entering the mouth, a pebble or stopper is interested into its top, usually made of a wet piece of cloth which protects the mouth from the heat and serves as a filter [8].

1.2.1.5 Water pipes: called *shisha* or *narghile* in Middle Eastern countries and *Hookah* in India and South Asia. Water pipes consist of a receptacle for water, with an opening on the top, to which a long wooden stem is fixed. At the top of this stem, a small bowl is attached for tobacco [9]. The smoke of tobacco is made to pass through the water before being inhaled, which cools and filters it. In ***hookahs***, around 20g of tobacco is burned over smouldering charcoal at each sitting, 20 times the amount contained in a cigarette. The water filter is not effective in removing tar, nicotine or carbon monoxide. Whereas a cigarette is typically smoked over approximately five minutes with 300-500 ml of smoke inhaled, hookah smoking sessions last from 20 to 60 minutes with volumes of 10 litres or more of smoke inhaled [2].

1.2.2 Smokeless tobacco

Smokeless tobacco is used to describe tobacco that is consumed without burning it [8]. Oral use of smokeless tobacco has existed for thousands of years in South America and has gained popularity in other areas of the world after the discovery of the new continent by Christopher Columbus at the end of the 15th century. Smokeless tobacco is used in two ways. One way is to place the preparations in various parts of the mouth and suck it, which is called “dipping tobacco”. Another way is to chew the preparations placed in the mouth which is called “chewing tobacco”.

1.2.2.1 Chewing tobacco: Tobacco is shredded like short cut grass, generally mildly acidic and intended to be chewed throughout the day as desired. In Pakistan and other parts of South Asia the most popular chewing product is betel quid. There are several ways of preparing these products but the main ingredients are betel, areca

and tobacco. ***Paan*** means “leaves of the betel vine” and is commonly used to refer to a chewing mixture wrapped in betel leaves. Most *paan* fillings are made of areca nuts, slaked lime paste and flavouring agents such as menthol, camphor, sugar, rosewater, aniseed, mint etc (*paan supari, paan masala, sada paan*); but tobacco can also be added to the filling (*tambako paan*). In *tambako paan*, tobacco is sun-dried, roasted, powdered and flavoured; or alternatively boiled, then made into a paste and scented with rosewater or perfume. *Paan* is placed in the mouth and gently sucked and chewed as a palate cleanser, a breath freshener and for digestive purposes, usually at the end of a meal. The speed of nicotine absorption, and hence the strength of the nicotine effect, increases with the pH of the mixture which is raised by the slaked lime [2]. ***Supari*** means “nut” in Indic language and indicates all mixtures from areca nuts, including *paan*. The areca nut can be either broken into pieces and chewed by itself; mixed with slaked lime, flavourings and optionally tobacco as paan; or mixed with sugar, spices, flavourings such as menthol and optionally tobacco [9]. ***Gutka*** (or *gutkha*) is in use in India, Pakistan, South-East Asia, and in the UK. It is a dry mixture of tobacco, betel nut and catechu, which are mixed together with slaked lime, flavouring (menthol, saffron) and sweetening spices (e.g.: cardamom, clove), and is held in the mouth to be chewed. Saliva is generally spit out, but sometimes swallowed. G*utka* was introduced half a century ago as a manufactured and cheaper substitute of hand-prepared products such as *paan* and *supari.* Compared to these products, *gutkha* is industrially prepared and has a longer shelf life. Its growth has been so rapid that it has overtaken the smoking of tobacco in India, being especially popular amongst younger generations and women[9].

1.2.2.2 Snuff dipping refers to tobacco chopped into particles like large coffee grounds, moistened and used by holding between gum and cheek, rather than chewed. ***Naswar*** is also called *nass* or *niswar*. It is used in Pakistan, Central Asia, Iran, Afghanistan, Baluchistan and India [8]. *Naswar* is made with fresh tobacco leaves dried in the sun or hot room, slaked lime, ash from a tree bark, flavouring (cardamom, oil, menthol) and colouring agents (indigo) mixed together. Water is added and the mixture is rolled into balls. It is held in the mouth for 10 to 15 minutes and is sometimes chewed slowly [9].

1.3 Definition of cardiovascular diseases

The term cardiovascular disease encompasses all disorders of the heart, blood vessels and blood circulation [10]. Development of the disease might be clinically silent for years until stenosis impairs the function of the heart or another organ and CVD becomes symptomatic. Coronary heart disease (CHD), also called ischaemic heart diseases and coronary artery disease, is the most common form of CVD. CHD manifests as myocardial infarction (MI) when atherosclerosis leads to narrowing and obstruction of the coronary arteries by formation of a thrombus or embolus, resulting in interruption of the blood supply to the heart and damage of the heart muscle [10].

Cerebrovascular diseases are another common form of CVD caused by either atherosclerosis developing in an artery directly supplying blood to the brain or emboli from a distant artery causing obstruction to the blood flow directed to the brain [10]. Cerebrovascular outcomes can be further classified as ischaemic stroke, haemorrhagic stroke and subarachnoid haemorrhage. Stroke refers to the damage to part of the brain caused by an interruption of its blood supply which can be due to a blockage of cerebral artery by a blood clot (ischaemic stroke) and rupture of a blood vessel in or near the brain (haemorrhagic stroke). Subarachnoid haemorrhage is a type of brain haemorrhage in which a blood vessel ruptures into the cerebrospinal fluid that surrounds the brain and spinal cords caused by burst aneurysm.

Other types of CVD include coronary aortic aneurysm, pulmonary embolism, heart failure, cardiac dysrhythmia and peripheral vascular diseases [10]. Aortic aneurysm refers to an abnormal dilatation of an artery located in the aorta caused by the pressure of the blood flowing through a weakened area as a result of atherosclerosis or MI. Pulmonary embolism occurs when a pulmonary artery or one of its branches becomes obstructed by an embolus, usually after a deep vein thrombosis. Heart failure means an inability of the heart to cope with its workload or pumping blood to the lungs and to the rest of the body. Cardiac dysrhythmia refers to a disturbance of heart rhythm caused by a problem with the electrical impulses in the heart. Peripheral vascular diseases refer to a narrowing of blood vessels in the legs or arms, causing blood flow and pain and, in severe cases, development of gangrene.

CVD are complex diseases mediated through multiple risk factors and pathways including tobacco use, dyslipidaemia, hypertension, inflammation, and insulin

resistance. Atherosclerosis is used to mean the development of fatty plaques on the artery linings, which begins in childhood, and develops usually by middle age into plaques several centimetres across the arterial intima, narrowing the arteries and impairing blood flow [11]. Lipoprotein particles bound to arterial intima have increased susceptibility to oxidative and other chemical modifications [12]. Early after initiation of hypercholesterolemia, leucocytes adhere to the endothelium and enter the intima, where they begin to accumulate lipids and become foam cells. Once macrophages? have taken up residence in the intima and become foam cells, they replicate and serve as a reservoir for excess lipid. Macrophages also provide a source of pro-inflammatory mediators. This is turn promotes the progression of lesions. Whereas the early events in atheroma initiation involve primarily altered endothelial function and recruitment of leukocytes, the subsequent evolution of atheroma involves smooth muscle cells as well. These smooth muscle cells multiply by cell division and also sometimes die, both mechanisms contributing to the complication of the atherosclerotic process. Another mechanism of plaque progression is thought to involve angiogenesis in plaques. Micro-vessels within plaques form in response to angiogenic peptides over-expressed in atheroma. The micro-vessels are known to be friable and prone to rupture [13]. A third mechanism is through development of areas of calcification in plaques and mineralization. In most cases, MI is caused by a fracture of the fibrous cap of the plaque [12]. Mechanisms which lead to the rupture of the fibrous cap include reduced collagen synthesis by smooth muscle cells, increased catabolism of the extra-cellular matrix macromolecules, deaths of smooth muscle cells and the presence of a large lipid pool.

1.4 Biological mechanisms by which tobacco use leads to CVD

Cigarette smoking has been shown to promote inflammation of the vessel wall, reduce oxygen availability and activate the sympathetic nervous system, all factors contributing to an imbalance between supply and demand of myocardial blood, oxygen and nutrients, and ultimately leading to clinical outcomes [14].

Cigarette smoking is associated with a more atherogenic lipid profile; including lower HDL-C and higher LDL-C, apolipoprotein A1, VLDL-C and triglycerides levels compared to non-smokers [15]. Smokers also have higher levels of oxidized LDL,which are taken up by macrophages to become foam cells that are an integral part of the atherosclerotic plaque [16] (**Figure 1.2**).

Inflammation is believed to contribute to atherogenesis; and white blood cell counts, C-reactive protein, and fibrinogen are predictors of future cardiovascular events [17]. Cigarette smoking results in a chronic inflammatory state, and smokers have been shown to have higher levels of reactive oxygen species, circulating leukocytes, C-reactive protein, and acute phase reactants such as fibrinogen [18] (**Figure 1.2**). Cigarette smoking enhances the recruitment and adhesion of leukocytes to blood vessel walls [19] and activates monocytes [20], which cause damage to the vessel walls.

Smoking also induces a hyper-coagulable state in individuals with relatively low levels of atherosclerosis, such as young smokers [21]. In men who experienced a sudden death, pathological findings of acute thrombosis are more likely to be present in smokers than in non-smokers [22]. Thrombosis can result from platelet activation, itself a consequence of endothelial dysfunction. Smoking may also affect the thrombogenecity of atherosclerotic plaques, through higher levels in tissue factor which is known to contribute to thrombosis after plaque disruption, higher levels of vascular cell adhesion molecule 1 and greater numbers of macrophages in atherosclerotic plaques [23].

Smoking not only acts as a major risk factor for CVD, but may also serve to induce or aggravate type 2 diabetes, therefore enhancing another major risk factor [24]. Smokers have been shown to have higher levels of free fatty acids and triglycerides after meals, which has been associated with insulin resistance.

Tobacco contains nicotine and 4,000 other compounds which are partially transferred into the smoke when a cigarette, cigar or pipe is lit up [25]. In pipes, cigars and bidies, tobacco is easily absorbed across the oral mucosa while the paper wrapping up tobacco in cigarettes reduces its absorption, necessitating inhalation of the smoke into larger surfaces of the lungs in order to satisfy the smoker's addiction [14]. The tar phase of the smoke, referring to particles of the smoke which can be measured by machine tests methods, contains a large number of toxic constituents and carcinogenic compounds. The vapour phase of the smoke mainly contains carbon monoxide produced by the combustion. Out of all the known components of cigarette smoke, only a few have been examined in isolation [6] (**Figure 1.3**).

1.4.1 Effects of nicotine

Nicotine has been shown to increase heart rate, and to produce endothelial dysfunction, lipid abnormalities and insulin resistance in smokers [26]. The amount of nicotine recovered from the smoke taken into the mouth varies from less than 0.3 mg to 3 mg per cigarette, depending on the degree of inhalation. Nicotine is then absorbed rapidly, with peak arterial blood levels of 10 to 100 ng/ml after a cigarette, and eliminated within 6 to 8 hours [27].

Intravenous nicotine, nicotine nasal spray, and nicotine chewing gum have been shown to increase heart rate and systolic blood pressure [28-30]. Nicotine stimulates the sympathetic nervous system by increasing plasma levels of norepinephrine and epinephrine, which translates into an acute rise in heart rate (up to 20 beats per minute) after smoking a cigarette [31] (**Figure 1.3**). In contrast, nicotine from chewing tobacco is absorbed slowly and peak arterial levels are much lower than those seen in cigarette smokers [14] (**Figure 1.4**).

Nicotine may also cause endothelial cells dysfunction [25, 32, 33]. Nicotine in concentrations similar to those found in the blood of cigarette smokers alters the structural and functional characteristics of cultured vascular smooth muscle and endothelial cells [34]. In studies of cultured endothelial cells, nicotine enhances the release of basic fibroblast growth factor and inhibits the production of transforming growth factor 1, increases DNA synthesis, mitogenic activity, and endothelial proliferation [35]. Nicotine has also been reported to act on human monocyte-derived dendritic cells involved in adaptive immunity, which have been detected in the wall of arteries and in atherosclerotic lesions and are known to stimulate an inflammatory response [36].

Finally, nicotine has been demonstrated to increase insulin resistance in smokers [37] and long term nicotine users [38] by activating the sympathetic nervous system, increasing release of corticosteroids and growth hormone.

1.4.2 Effects of carbon monoxide (CO)

CO is a major constituent of cigarette smoke, while users of smokeless tobacco remain unexposed. CO has been directly associated with CVD risk in a dose-response manner [22]. Non-smokers have average levels of carboxyhemoglobin

ranging from 0.5% to 2%, compared to levels of 5% to 10% in smokers [39]. CO binds avidly to haemoglobin, reducing the amount of haemoglobin available to carry oxygen, and impedes oxygen release by haemoglobin. CO inhalation in people with CAD has been shown to provoke reduced exercise tolerance and exercise-induced ventricular dysfunction, including ventricular arrhythmias [40]. Long term CO exposure also results in an elevated red cell mass in smokers, to compensate for hypoxemia, which contributes to increased blood viscosity and promotes a hyper-coagulable state in smokers.

1.4.3 Effects of oxidants

Tobacco smoke also delivers a high concentration of oxidizing chemicals to the smoker [41]. These chemicals include oxides of nitrogen and a number of different free-radicals, found both in the gas and tar phases of the smoke. Oxidant stress translates into elevated levels of peroxides and decreased levels of traditional plasma antioxidants such as vitamins A and C in smokers [42]. Oxidant stress provokes a reduction in nitric oxide release amongst smokers [43] and is believed to contribute to a number of the potential mechanisms of CVD, including inflammation, endothelial dysfunction, oxidation of LDL-C and platelet activation. Nitric oxide and prostacyclin are vasodilators and have anti-platelet aggregation effects. Oxidant chemicals also produce hyper-aggregability of platelets through peroxidation of free fatty acids.

1.4.4 Effects of carcinogenic chemicals

Some pollutants found in the tar fraction of cigarette smoke called polycyclic aromatic hydrocarbons, such as benzo(a)pyrene, benzo(a)anthracene [44] and butadiene [45], have been reported to accelerate atherosclerosis in experimental animals, at doses below those that produce tumours. A mechanism of atherogenesis is speculated to be a mutation of smooth muscle or other cells that become the source of an atherosclerotic plaque.

1.5 Prevalence of tobacco use

Amongst all forms of tobacco use, smoking is the most prevalent globally. In 1995, more than 1.1 billion people smoked worldwide, with about 82% of smokers residing in low- and middle-income countries [46]. Worldwide, male smoking far exceeds female smoking, with a smaller gender difference in high-income countries. Smokeless

tobacco use is more common in developing countries, but migration and tourism mean its use is being exported to developed countries, where they are most fashionable in younger age groups.

1.5.1 Past and current prevalence of tobacco use in developed countries

The use and culture of tobacco is old, dating back to around 1,000 BCE, by natives of the Americas who smoked tobacco in pipes for medicinal and ceremonial purposes [47]. Christopher Columbus was the first Westerner to discover tobacco in 1492, bringing back leaves and seeds with him to Europe. However, the use of tobacco did not spread until the mid-16th century, popularized by adventurers, sailors and travellers such as the French Jean Nicot, who gave his name to the world nicotine. Smoking was introduced to France, Portugal, Spain and England in the mid-16th century and in North America at the beginning of the 17th century. At that time, the main forms of tobacco use were pipe smoking, chewing and snuff. Cigars became popular at the beginning of the 19th century while cigarettes did not start to be mass-produced in the USA until the end of the 19th century with the invention of a machine to replace hand-production [48]. With the mass production of cigarettes, the proportion of smokers increased enormously in developed countries during the first part of the 20th century. In the meantime, pipes and cigars became less prevalent and snuffing and chewing became on the verge of extinction. Between 1900 and 1960, the annual number of cigarettes sold in the USA was multiplied by a factor of 80, going from an average of 50 cigarettes per person per year to an average of 3,900 cigarettes per person per year (**Figure 1.5**). During the same period, the sale of pipe tobacco was divided by 3 (from 1.6 to 0.6 pounds per person per year), the number of cigars per person per year was halved, and the consumption of chewing tobacco and snuff, already low, kept decreasing (from 110 to 60 usages per person per year) [2] (**Figure 1.5**).

Cigarette smoking prevalence began to decline in the USA in 1960s, following the publication of reports on the dangers of tobacco by the UK Royal College of Physicians and US Surgeon General [7, 47]. The public became aware of the dangers of cigarette smoking and governments started to regulate cigarette sale and advertisement. In the USA for example, health warnings were imposed in 1965 on cigarettes packs, broadcast advertising was forbidden in 1971, and smoking was banned in buses and domestic flights in 1990; the Food and Drug Administration was created in 1995 with the aim of regulating the sale of tobacco, and its role was

reinforced in 2009 [7]. As a result, the proportion of smokers has been approximately halved over the past 50 years in developed countries. Between 1960 and 2005, the proportion of adult male smokers dropped from 52% to 27% in the USA, from 81% to 43% in Japan, and from 61% to 25% in the UK. In women, the drop in prevalence was less dramatic: from 13% to 12% in Japan, from 42% to 24% in the UK, and from 34% to 19% in the USA [46]. At baseline of the British Doctors Study (BDS) initiated 1951 by Doll and Hill, 87% of UK male doctors reported smoking cigarettes and/or pipes [49] while, in 2000, the proportion of male physicians smoking in the UK was 8% [46]. Concomitant with a reduction in prevalence, smoking intensity also declined. In the USA, in 1965, 56% of smokers were smoking at least 20 cigarettes per day (CPD), while smokers of more than 20 cigarettes a day represented 23% of smokers in California and 40% of smokers in the remaining United States in 2007 [50].

However, the addictive nature of smoking and the peer pressure existing amongst teenagers to initiate smoking have meant that roughly 1 in 5 adults still smoke in developed countries (**Figure 1.6**). Current smokers comprise 1 in 5 adults in the USA, 1 in 4 in the UK and 1 in 3 in Japan. In the European Union, the proportion of people currently smoking varies between 25% and 35% for men and between 15 and 25% for women [51]. With a lesser decline in prevalence amongst women compared to men, the gender gap is narrowing and women are catching up with men. The European Union has the highest worldwide prevalence of women smoking [46]. Teenage girls are at least as likely and sometimes more likely than teenage boys to start smoking. In the USA, more than 20% of both boys and girls aged 13-15 years old were tobacco users in 2000 [46]. A youth survey in Portugal in 2001-2002 found that 26% of girls versus 18% of boys smoked at least once a week [52].

While cigarette smoking has experienced a major decrease, other forms of smoking tobacco remain stable or are on the rise. Roll-your-own (RYO) cigarettes represent a cheaper option than manufactured cigarettes, and are often used to combine tobacco with illegal products such as marijuana. Their use is increasing in the USA and Europe, especially amongst young people [2]. Studies have found that RYO users are heavier smokers, more addicted to nicotine and less likely to consider quitting [53]. Another popular alternative to cigarettes has been cigars, particularly small cigars called “cigarillos” which are made to look like cigarettes in shapes and format. The proportion of cigar users has been increasing in the USA since 1990. Since 1995, initiation of cigar smoking has outnumbered initiation of cigarette smoking in the

USA, being particularly appealing to teenagers and women [54]. In 2003, current cigar smokers made up 5.5% of the population in USA and the highest rates were reported amongst young adults aged 18 to 25, where 11% reported having smoked a cigar in the past month. In 2007, in industrialized countries from the former Commonwealth (Australia, Canada, UK and USA), the prevalence of cigar use varied between 3% and 13%, pipe use between 0.3% and 2.1%, smokeless tobacco was used by less than 2.3% and RYO by 12% to 21% of the population [55].

1.5.2 Past and current prevalence of tobacco use in developing countries

Tobacco use originated in South America, and spread to other parts of the world, including Africa and Asia, by European travellers and merchants, during the 16th and 17th century [2]. It is during this period that mixing tobacco with various chewable mixtures of herbs, spices, areca nut, betel leaf and other substances became popular in the South Asian subcontinent. In the 18th and 19th centuries, dry powdered tobacco to be snuffed into the nose became popular in parts of East-Asia.

The production and consumption of tobacco has expanded rapidly in developing countries during the second half of the 20th century with the introduction of industrially and mass produced tobacco products. Developing countries represented in 1970 around 40% of the world production and consumption of cigarettes, and they now represent 70% (**Figure 1.7**). The proportion of smokers is now highest worldwide in lower-middle income countries, where it attains 40% (**Figure 1.6**). There are regional differences: prevalence is lowest in the African Region as defined by the World Bank (<15% men smoking), and highest in the Western Pacific Region (>45% men smoking). In the South-East Asian region, prevalence of daily smoking is above 30% in men and it is close to 30% in the Eastern Mediterranean region. In particular, in China and the Philippines, above 65% of men currently smoke [56]. One shared aspect of tobacco uses across developing countries is the presence of a gender gap due to cultural and social reasons. On average, fewer than 50% women smoke in developing countries (**Figure 1.6**). However, women are nowadays aggressively targeted by the tobacco industry, which seeks to associate tobacco use with feminism, sophistication, weight control, and Western-style independence. Recent increases in female smoking prevalence have been reported in Cambodia, Malaysia and Bangladesh [7].

Until recently, governments have been showing a weak response to the rise in tobacco use. In terms of per capita public spending, for every $1 dedicated by developed countries to tobacco control, middle income countries dedicate $0.005, and low income countries $0.001 [51]. Philip Morris, the world's biggest cigarette company, was also in 1996 the world's ninth largest advertiser, spending more than $3 billion on promoting its products, mainly in developing countries. Following the adoption in 2003 of the WHO Framework Convention on Tobacco Control, which binds states to ban tobacco promotion, several developing countries have passed anti-smoking laws. However, their impact remains limited. In Pakistan, anti-smoking laws which ban consumption in public areas and storage near educational institutions were passed in 2009, but are not being enforced [57]. In India, a ban on the sale of *gutka* issued in December 2010 has had no practical effects due to a lack of support from the general population, insufficiently educated on the dangers of tobacco use [58]. In China, where the prevalence of smoking among health-care professionals reaches 40%, many (health professionals) smoke in front of their patients and as many as 29% of non-smoking physicians accept cigarettes as gifts, making them ineffective in discouraging smoking initiation and promoting smoking cessation to the wider population [59]. Whilst prevalence of smokers who have quitted their habit is generally above 35% in developed countries, it is less than 20% in developing countries [60]. In Bangladesh, India and China, the prevalence of former smokers is below 10% before age 45 years old.

1.6 Burden and economic impact of tobacco use and cardiovascular diseases

1.6.1 Tobacco overall morbidity and mortality

During the 20th century, 100 million people worldwide died from tobacco-related diseases; and the tobacco burden is predicted to reach 1 billion individuals during the 21st century [7, 46]. While the tobacco epidemic has reached its peak in developed countries, it is only at its early phases in developing countries, with a gap between men and women and an increasing prevalence of smoking over time (**Figure 1.9**) [61]. In 2000, nearly 5 million premature deaths in the world were attributable to smoking, half of them in developing countries and half of them in developed countries [51]. Nowadays, 6 million people die from tobacco use each year, both from direct tobacco and second-hand smoke [62]. Based on current smoking patterns, annual smoking

related deaths will rise to 8.3 million by 2030, with four out of five deaths occurring in developing countries. Smoking represented 3% of all deaths in 2000 for medium and low income countries, and this proportion is expected to rise to 8% by 2030 (**Figure 1.10**) [63].

1.6.2 Cardiovascular burden from any cause

CVD is the main cause of death worldwide and over 80% of CVD deaths now occur in developing countries [51]. CVD were responsible for the largest proportion of non-communicable deaths under 70 years old in 2008, representing 17 million deaths of 48% of deaths from non-communicable disease. Non-communicable diseases were collectively responsible for 36 million deaths, representing 63% of all deaths (57 million individuals in 2008), exceeding in all regions except Africa the burden of communicable, maternal, perinatal and nutritional conditions combined (**Figure 1.11**). Population ageing is a significant trend in most parts of the world and translates into increasing burden of CVD. Whereas annual infectious disease deaths are projected to decline, annual CVD deaths are projected to increase by 6 million worldwide over the next 20 years. The burden of non-communicable disease is expected to increase by 15% globally between 2010 and 2020, to 44 million deaths a year. Greatest increases are predicted to happen in Africa, South-East Asia and the Eastern Mediterranean region, where the burden will increase by over 20%, and lowest increases in the European Region where the annual burden is expected to remain stable. By 2020, the regions that are projected to have the greatest total number of deaths from non-communicable diseases are South-East Asia (10.4 million deaths) and the Western Pacific (12.3 million deaths).

1.6.3 Cardiovascular burden as a result of tobacco use

Tobacco is estimated to cause 10% of CVD worldwide [63]. In 2000, amongst individuals aged ≥30 years old, smoking accounted for 22% of CVD mortality in North America, 13% in Western Europe, 10% in South Asia including India and 4% in South-East Asia including China. Because female smokers are attaining similar levels to male smokers in developed countries, smoking related CVD mortality is now roughly similar across genders, while a gap remains in developing countries. In North America, in 2000, smoking accounted for 23% of CVD mortality in men versus 21% in women. Inversely, CVD represents a leading cause of the smoking related burden.

It represents 51% of smoking related deaths in Japan [64]; compared to 13% in China [65].

CVD represented 1.7 million tobacco related deaths worldwide in 2000 and is projected to reach 1.9 million deaths in 2015 (0.92 million CHD deaths, 0.52 million stroke deaths and 0.24 million other CVD deaths) [66]. In the USA, cigarette smoking is the cause of nearly 500,000 premature deaths a year, about 1 in every 5 deaths, and nearly half of them are due to cardiovascular diseases [67]

1.6.4 Economic impact of cardiovascular diseases

Once thought of as a diseases of the rich, CVD is now the leading cause of death in low- and middle-income countries and adds up to a heavy socio-economic toll worldwide. Nearly 30% of all deaths from non-communicable diseases in developing countries occur before 60 years old, whereas in high-income countries the proportion is only 13% [51]. Each year, 100 million people are pushed into poverty because they have to pay directly for health services and for a large proportion this is as a result of CVD. Developing economies with large population will bear the highest cost of the CVD epidemic. From 2005 to 2012, China and India are projected to lose 1% of the Gross Domestic Product as a result of heart diseases and 1.5% of Gross Domestic Product as a result of strokes and diabetes.

1.6.5 Economic impact of tobacco use

There are multiple economic impacts of tobacco, ranging from health costs and impoverishment of families to damage done to the environment (for example, forest fires caused by lit cigarettes carelessly thrown to the ground). The personal economic cost of buying tobacco impoverishes the most deprived sections of the population, and slows economic development, especially in developing economies [7]. Indeed, tobacco use is more common amongst people in low socio-economic categories, who dedicate an important part of their resources to buying tobacco products and are often unable to cover the medical costs incurred as a result of tobacco related diseases. The poorest households in Bangladesh spend almost 10 times as much on tobacco as on education; and tobacco use has been demonstrated to exacerbate child malnutrition diverting household funds away from food and other necessities [68]. In Pakistan, a pack of imported manufactured cigarettes costs more than half the average daily income [46]. In Indonesia, the lowest income group spends

15% of its total expenditure on tobacco. In China, poor households spend between 7% in cities and 11% in the countryside of their income on cigarettes [69]. Tobacco is also a main cause of disability and death among middle-aged men, often the main breadwinners of their families, causing considerable economic losses and accelerating entry into poverty [63].

At a population level, tobacco health burden represents a sizeable part of total health expenditure in both industrialized countries and developing countries. Developed regions account for 12% of the worldwide burden from all causes of death and disability, but for 90% of health expenditure [7]. Total annual health expenditure related to smoking, including diseases and deaths from passive smoking, runs to $81 billion in the USA, $7 billion in Germany and $1 billion in Australia [46]. In China, the annual cost of health expenditure related to smoking reaches $3.5 billion [51].

1.7 Epidemiological evidence on the cardiovascular risks associated with tobacco use

This section combines synthetizes literature reviews of smoking in relation to CVD with regard to 1) dose-response relationship, 2) the effect of duration, age at risk and starting age, 3) the effect of cessation, 4) the importance of confounding and 5) the potential for effect modification. In addition, the relationship between smoking and other risk factors or markers of CVD is reviewed. Finally, a review of the evidence regarding other types of tobacco use is conducted and epidemiological evidence is compared to cigarette smoking. Studies reviewed in this Chapter are listed at the end of the Chapter in **Table 1.1**.

1.7.1 Landmark epidemiological studies and evaluation of the causal effect of smoking on CVD risk

Until the landmark studies published in 1954 in the UK by Doll and Hill, and in 1958 in the US by Hammond and Horn [70] (**Table 1.1**), epidemiological studies on the effects of tobacco had been small, focused on the risk of lung or mouth cancer, and retrospective. These studies were unable to prove beyond doubt a causal association between smoking and CVD. Therefore, Doll and Hill recognized the need for *"some entirely new approach. That approach I considered should be prospective. It should determine the frequency with which the disease appeared, in the future, among*

groups of persons whose smoking habits were already known". Doll and Hill enrolled 40,000 British doctors in November 1951 who filled in a short questionnaire on their smoking habits. After 3 years of follow-up, the authors found: *"…There is a rise in the mortality from deaths attributed to coronary thrombosis as the amount smoked increases, but the gradient is much less steep than that revealed by cancer of the lung".* After a follow-up of 12 years, the authors were able to conclude more assertively *"In short, that cigarette smoking is a cause of coronary thrombosis is not, I think, proved; but it is the most reasonable interpretation of the available facts"* [71] and after a 20 years follow-up, they classified the burden of coronary heart diseases (CHD) as *"probably partly or wholly attributable to smoking"* [72].

Complementary to the study by Doll and Hill, Hammond and Horn initiated in 1952 a study on a larger scale in the USA, enrolling 188,000 white men age 50 to 69 years old, and published their first results in 1958. During four years of follow-up, 11,870 deaths occurred, including 5297 due to CHD [70]. With a larger data size and longer follow-up than Doll and Hill, the authors had enough statistical power to conclude already in 1958 that *"coronary heart disease and other circulatory diseases showed a high degree of association with cigarette smoking"*, and to estimate a doubling in risk of CHD death in individuals smoking at least 1 pack of cigarettes per day (CPD) compared to never-smokers. Since these two landmark studies, cigarette smoking has been demonstrated to cause myocardial infarction, stroke, aortic aneurysm, sudden cardiac death and peripheral vascular disease [73].

1.7.2 Dose-response relationship

Two prospective studies including each around 7,000 people, the Goteborg Study (GOTO) in Sweden and the British Regional Heart Study (BRHS) in the UK, found a similar effect on the risk of fatal and non-fatal MI independently of the level of exposure to cigarette smoking [77,78] (**Table 1.1)**. However, most studies have observed a dose-response relationship between CPD and the risk of death from CVD [79-81]. In 1976, Doll & Peto published the first significant tests for trends with increasing CPD for the risk of deaths from ischaemic heart disease, myocardial degeneration, arteriosclerosis, aortic aneurysm and cerebral thrombosis [72]. In the Nurses' Health Study (NHS), a prospective cohort study of 100,000 women (**Table 1.1**), the multivariate adjusted HR for CHD deaths was 2.8, 4.8, 7.0, 7.8 in women smoking respectively 1-14, 15-24, 25-34 and ≥35 CPD versus never-smokers [74]. For stroke, the test of linear trend was non-significant but the increased risk was 2.1 (95% CI:

1.6-2.8) in women smoking 1-4 CPD versus 3.3 (2.1-5.4) in women consuming ≥35 CPD. INTERHEART, an international case-control study of the risk of MI conducted in 52 countries (**Table 1.1**) estimated an increase of 6% in risk per additional cigarette [75].

Even low levels of smoking have been associated with a significantly higher risk of CVD. The risk of first ever non-fatal MI was significant in INTERHEART at 3-4 CPD. In the NHS, smoking 1 to 4 CPD was associated with a twofold increase in the risk of MI. In the Copenhagen City Heart Study (CCHS, see **Table 1.1**), individuals consuming 3-5 grams of tobacco per day (corresponding approximately to 3-5 CPD) were shown to carry a significantly increased risk of developing MI with a RR of 2.14 (1.11; 4.13) [76]. In a prospective study conducted in Norway, started in 1970 and followed up until 2002, smoking 1-4 CPD was positively associated with a significantly higher risk of dying from MI [77,78]. A rapid rise in CVD risk even in light smokers is consistent with observations made in passive smokers, who are exposed to relatively low levels of smoke compared to even light current smokers, but who are nevertheless at significant increased risk of CVD [79]. Cotinine is a metabolite of nicotine with a relatively long half-life of 16 hours commonly used as a proxy to exposure to passive smoking and daily cigarette consumption. A prospective study measuring cotinine found that never-smokers exposed to high levels of passive smoking were also at higher risk of MI: the HR for top versus bottom 4th of the distribution was 1.57 (1.08-2.28) after adjustment for conventional risk factors [80]. Law and Wald proposed a shape of association with CHD with an exponential increase up to 5 CPD, and then a linear increase up to 30 CPD. (**Figure 1.12**).

The existence of a plateau effect at high levels of smoking status remains controversial. INTERHEART observed a linear increase in MI risk up to 21 cigarettes per day with no evidence of a threshold [81]. In the Framingham Study (**Table 1.1**), the risk of stroke increased linearly with CPD and the RR was 2 for smoking greater than 40 CPD compared to fewer than 10 CPD [82]. In a prospective study of 325,384 white US males followed for 5 years, age-adjusted mortality rates were linearly increasing with no plateau at consumption >35 CPD [83]. In the Stroke Prevention and Young Women Study (SPYW, see **Table 1.1**), a strong dose-response relationship was observed for the risk of stroke, with an OR of 9.1 (95% CI: 3.2-26) for women smoking 40 or more cigarettes per day compared to never-smokers [84]. However, in the NHS, risk ratios for all CHD were the same in women consuming either 25-34 or

≥35 CPD compared to never-smokers: 3.7 (3.1; 4.4) and 3.7 (1.9; 4.7) respectively after 24 years of follow-up [74, 85]. Similarly for cerebrovascular disease, risk ratios were 2.9 (2.0; 4.2) and 2.9 (1.8; 4.8). The convergence effect found in this study has been attributed to a steeper increase in exposure to smoking at low levels of smoking, as heavier smokers may take lighter and shorter puffs, therefore inhaling less smoke. A study which measured cotinine found that cotinine levels increased more steeply from 0 to 10 CPD than from 20 to 30 and were reaching a plateau above 30 CPD [86].

1.7.3 Duration, age at risk and starting age

The effect of duration of smoking in relation to disease is difficult to disentangle from the effect of intensity, due to the addictive nature of smoking, as amount is likely to increase with increasing duration. In addition, most smokers start during adolescence so age and duration are nearly collinear variables, and the estimation of their separate effects requires large sample sizes [73]. In the Multiple Risk Factors Intervention Trial (MRFIT, see **Table 1.1**), a study conducted among 361,662 men followed up for 10 years, the association of duration with risk of CHD death was not significant, after adjustment for age, cholesterol levels, blood pressure and CPD [87]. However, rate ratios compared to never-smokers were increasing with duration at any level of cigarette consumption and in any age strata, in individuals younger than 70 years old in CPS-1, and in everyone in CPS-2 [85], and this finding has been replicated in other studies [88]. The effect of smoking on disease is often represented by the number of pack-years increase, measured as the number of packs of CPD times the number of years smoking, each pack containing 20 cigarettes, and makes the assumption that an increase in 1 year of duration is equivalent to increasing smoking consumption by 1 pack per day. In 3 City Study (3CS, see **Table 1.1**), women smoking more than 30 pack-years had a hazard ratio of 3.2 (2.4-4.5) compared to never-smokers [89].

Regarding age at risk, studies have consistently shown a decrease in risk ratios with increasing age [71, 90-92]. In the British Doctors Study (BDS, see **Table 1.1**), death rates from CHD were 5.7 times higher among cigarette smokers than among non-smokers at ages 35 to 44 but were approximately equal to those of non-smokers at ages 75 to 84. In a large case-control study of 14,000 cases and 32,000 controls, the rates of MI in smokers were 5 times those in non-smokers at age 30-49 years old, 3 times at age 50-59, and twice at age 60-79 [75]. A similar effect has been observed for most risk factors of CVD, with stronger associations observed at younger ages [85, 93, 94] . In

individuals below 45 years old, CHD was the dominant cause of increased mortality attributable to cigarette smoking in CPS-2 [85].

The age of starting smoking has also been associated with CVD risk, independently of duration, amount, age, and known intermediate risk factors; with young starters experiencing higher risks than late starters. In CPS-2 and in the US Veteran Study, individuals starting earlier were generally at higher risk independently of their number of CPD and their age [102]. In the NHS, at 12 years of follow-up, the adjusted risk ratio of CVD compared to never-smokers was 9.2 (95% CI: 5.3-16.2) in women who had started before age 15, while it was 3.2 (2.1-4.8) for women who had started aged ≥26 years [71].

1.7.4 Smoking cessation

While reducing smoking amount does not reduce CVD risk, smoking cessation is beneficial, although studies differ in their estimation of time necessary to achieve a risk comparable to that of never-smokers [85, 95]. In the BDS, after correcting for reverse causality by excluding the first 5 years of follow-up, CHD mortality decreased relatively slowly in ex-smokers: the rate was 60% that of non-smokers after 5-9 years and 29% the rate of non-smokers after ≥20 years [74]. However, ex-smokers were relatively rare at the time and these results were based on small numbers. In the US Veteran Study, CHD deaths rates depended on the amount formerly smoked, were not reduced until 20 years after stopping smoking, and became comparable to that of never-smokers only 30 years after cessation [75]. In the BRHS, the incidence of major CHD event was significantly raised even 20 years after giving up smoking and increased with the number of years smoked [96]. In INTERHEART, while ORs for MI risk were halved within the first 1-3 years of abstinence, they remained significant after 20 years of cessation (1.31; 95%CI: 1.13-1.51) for ex- versus never-smokers unexposed to passive smoking, independently of the amount previously smoked. By contrast, in the NHS, the multivariate adjusted risk of CHD death was halved within 5 years and reached the level of non-smokers in 10-15 years, with a similar decrease observed for cerebrovascular death [90]. In the Third National Health and Examination Survey (NHANES-3, see **Table 1.1**), inflammatory markers of atherosclerotic disease such as C-reactive protein, white blood cell count, albumin and fibrinogen returned to population levels 5 years after smoking cessation [97].

1.7.5 Confounding

1.7.5.1 Lifestyle risk factors: There has been a suggestion from the tobacco industry that the association of smoking with CHD risk reflects an inadequate control for confounding by lifestyle risk factors. In CPS-2, smokers tended to be less educated, drank more alcohol and ate fewer vegetables. Male smokers in particular were less likely to be employed, more likely to be in a unskilled job, consume a fatty diet and were less physically active [14]. However, adjustment of the risk ratios for race, education level, marital status, unskilled job, weekly consumption of vegetables and citrus fruit, aspirin use, alcohol consumption, body mass index, physical activity and weekly consumption of fatty food decreased the estimate for CHD death for current versus never-smokers by only 9% in men and 5% in women. For the risk of stroke death, the biggest decrease upon adjustment was in men: the age adjusted HR for current versus never was 2.1 (1.9-2.4) while the full adjusted HR was 1.7 (1.5-2.0) (in women it was 2.3; 95%CI: 2.0-2.6 compared to 2.2; 95%CI: 2.0-2.5). In a British study of 7,142 men followed up for 15 years, adding BMI and physical activity as covariates to the Cox model provoked a modest increase in the HR of MI and stroke [98].

1.7.5.2 Intermediate CVD risk factors: Cigarette smoking has been shown to influence several intermediate risk factors of CVD. It has been found to acutely increase blood pressure and heart-rate, and with regular use throughout the day, the increase remains persistent [99]. In the Edinburgh Artery Study (EAS), smoking was associated with reduced dietary antioxidant vitamin intake, HDL-C and diastolic blood pressure (DBP), and with increased alcohol intake, serum triglycerides, blood viscosity, plasma fibrinogen and markers of endothelial disturbance [100]. In the BRHS, smokers had lower BMI, DBP and HDL-C; while they had higher systolic blood pressure (SBP) and serum triglycerides than non-smokers. Ex- versus never-smokers had higher SBP, BMI, total cholesterol and triglycerides. In the Munster Heart Study (MHS-PROCAM), the same observations were made, as well as higher levels of fibrinogen in smokers versus non-smokers [101]. Changes in lipid levels were of a greater magnitude in women compared to men [74]. However, adjusting for intermediate risk factors of CVD only marginally affected RR in these two studies. In the Seven Countries Study (SCS), after adjustment for cohort effect, age, BMI, serum cholesterol levels, SBP and the presence of clinical CVD, the HR for smokers ≥10 CPD remained highly significant at 1.8 (1.6-2.1) for CHD death compared to never-smokers [102].

Overall, both lifestyle and intermediate risk factors fail to explain the association between smoking and CVD, which indicate that smoking independently increases risk and causes CVD by independent biological pathways. It also means that even a tight control of other risk factors such as blood pressure and cholesterol level would not counterbalance the effect of smoking. In NHS, adjustment for hypertension, diabetes, high cholesterol levels, BMI, change in weight between age 18 and baseline, alcohol intake, physical activity, contraceptive and hormone use, menopause status, parental history of MI, diet, daily number of CPD and age of starting smoking slightly raised the association with vascular deaths for current versus never-smokers: 3.0 (2.4;3.2) to 3.3 (3.0; 3.8) [75].

1.7.6 Effect modifications

1.7.6.1 Sex: Early epidemiological studies observed significantly lower risk of all causes of death and CVD for women than for men when they smoked, even after a 20 years follow-up in the case of the BDS [85, 103]. However, women smokers included in these studies were mostly born before the First World War; and these women differed substantially from men of the same era who smoked, and from younger women smokers. In particular, these women were less likely to inhale and had started smoking later than their male counterparts. By contrast, in more recent studies, the risk for women smoking has been shown to be equal [90] or even higher than the risk for men [76, 104], in particular for the risk of early CVD [12, 89]. In the Finmark study, the incidence of MI was increased six-fold in women versus threefold in men who smoked ≥20 CPD compared to never-smokers. In all age groups, the HR for women smoking was 3.3 (2.1-5.1) compared to 1.9 (1.6-2.3) in men [105]. Reasons for this difference are unknown and may include hormonal effects of smoking (female sex hormones may affect CVD risk), synergy with oral contraceptive agents or even arteriolar differences.

1.7.6.2 Intermediate factors and medical history: Whether the effect of smoking is exacerbated by the presence of other known CVD risk factors remains uncertain because of sparse evidence. In ARIC, the association of smoking with fatal and non-fatal CHD was stronger in individuals with higher LDL-C levels [106]. In individuals with LDL-C ≥130 mg/dl and smoking ≥15 CPD, the HR was 2.81 and higher than expected if these factors were independently affecting CHD risk (1.15X1.71=1.97). Nevertheless, confidence intervals were wide and this difference was not significant. In the Asia Pacific Study Collaboration (ASPC), smoking interacted positively with

total cholesterol and HDL-C on CHD risk, but p-values fell between 0.01 and 0.05 and therefore were only modestly significant [107]. In diabetic patients, smoking has been found in a single study to interact with the duration of diabetes to accentuate atherosclerosis [108]. The presence of diabetes or hypertension was also found to strengthen the association between smoking and subclinical atherosclerosis, measured by carotid intima-media thickness [109, 109].

1.7.6.3 Lifestyle factors: It has been suggested that smoking worsens the negative effect of alcohol and offsets the protective effect of low consumption of alcohol on CVD risk. In two prospective studies, the protective effect of alcohol in light and moderate alcohol drinkers on CVD mortality was non-significant in smokers, however there was no statistical evidence of an interaction on a multiplicative scale at high levels of consumption [110, 111].

1.7.7 Smoking and the progression of CVD

As seen in the section above, the risk of acute CVD events decreases rapidly upon cessation of smoking, which may suggest that smoking has a greater impact on plaque rupture and thrombus formation than on the atherosclerotic process of plaque building [83]. In the Atherosclerosis Risk in Community Study (ARIC), the association of smoking was stronger with advanced rather than with early atherosclerosis measured by carotid intima media thickness, regardless of levels of smoking exposure; and smoking was associated with plaque calcification [106]. The Multi-ethnic Study of Atherosclerosis (MSA) found that smoking accelerated plaque progression to thicker, more fibrous lesions which are more vulnerable to rupture [112]. A stronger association between smoking and CVD has also been reported at high compared to low levels of LDL-C, which is consistent with the view that smoking accelerates the progression of cholesterol-filled regions [113].

1.7.8 Other smoking and smokeless types of tobacco use and CVD risk

1.7.8.1 "Filtered", "low tar" and "roll-your own" cigarettes: The tobacco companies have been quick to develop alternatives to the standard cigarette with the claim that these new products are less harmful. However, by changing the way they smoke, for example, inhaling more deeply, or increasing the number of CPD, it is possible to obtain as much nicotine from these new products as from regular cigarettes, in order to satisfy one's addiction, counterbalancing any "protective" effect of these new products [114].

"Filtered" cigarettes, which aim to reduce the amount of toxicants that go into the smoke inhaled by the smoker, have been shown to carry the same CVD risk as "unfiltered" cigarettes [115]. Similarly, "Low-tar" or "light cigarettes", designed to burn more quickly than standard cigarettes, and to produce lower machine-measured yields of tar and nicotine, have not been shown to reduce the risk of non-fatal MI [116] or subclinical atherosclerosis [117].

"Roll your own" and bidies have not been demonstrated to carry either more or less risk than cigarettes. In a case-control study in India, the risk of vascular death associated with cigarette smoking was 1.8 (1.7-1.9) in an urban setting, where manufactured cigarettes are predominant (26% smoked only cigarettes and 69% both cigarettes and bidies); and was nearly identical to the OR in rural areas, where bidies are most popular (38% smoked only bidies and 57% both bidies and cigarettes): 1.7 (1.6-1.9) [118]. In a smaller case-control study of Indians, individuals smoking ≥10 cigarettes or bidies per day had an OR of 6.7 ($p<0.001$) [49]. A systematic review on bidi smoking found that bidi smokers inhale on average 2–3 times more nicotine and tar than smokers of conventional cigarettes, due to the poor combustibility of the bidi wrapper and greater puff frequency needed to keep the bidi alight; with resulting health hazards at least as great as for cigarettes. A small scale case-control study in Bangalore found higher OR for bidi versus cigarette smoking [3].

1.7.8.2 Pipe and cigars: The BDS was the first study to investigate the effect of pipe and cigar on the risk of CVD and found a non-significant association, leading to the early belief that these products were safer alternatives to cigarettes [119]. Smoke from pipes and cigars contains the same toxic substances as cigarette smoke, but those who use a pipe or cigar tend to smoke at lower intensity and tend not to inhale the smoke, thus reducing their exposure to its toxic substances. For example, 2/3 of those who smoke both cigars and cigarettes (>40% cigar smokers) inhale cigar smoke, compared with less than 15% of cigar smokers who have never smoked cigarettes [120].

More recent studies have established that pipe and cigar smokers are not protected even if their risk is smaller than cigarette smokers [121]. In the Zutphen Study, current cigarette smokers experienced a reduced life-expectancy by 6.8 years (disease-free life years decreased by 5.8 years) compared to never cigarette smokers, whereas

current pipe or cigar smokers had on average 4.7 years of life lost (5.2 disease-free years) compared to never pipe or cigar smokers[122]. In the Kaiser Permanente Study (KPS), current cigar smokers were at higher risk of CHD after multiple adjustment (RR compared to never-smokers: 1.27; 95%CI: 1.12 to 1.45) with evidence of a dose-response relationship [123]. In CPS-2, the association between current versus never cigar smoking and CHD death was stronger among younger men and was non-significant in ex-smokers [124]. The association with current versus never pipe smoking in this cohort was similar to that of cigar smoking, with a RR of 1.30 (1.18-1.43) for CHD.

1.7.8.3 Smokeless tobacco: Smokeless tobacco is varied in its forms, and epidemiological evidence on each type is limited and generally derived from small-scale studies. Snuff users have been shown to have levels comparable to never-smokers in terms of inflammation (fibrinogen [125] and C-reactive protein [126]), endothelial dysfunction (levels of carotid intima media thickness measured in the carotid and femoral arteries), and oxidant stress [127]. A case-control study conducted in Bangladesh found that, compared to never users of tobacco, ever cigarette smokers had an OR for CHD of 3.6 (1.5-8.5), ever bidi smokers had an OR of 2.9 (1.3-6.3), while ever use of betel nut with quid carried an excess risk of 3.8 (1.9-7.7) and ever use of dried tobacco leaf (corresponding to Pakistani naswar) carried an OR of 2.8 (1.0-4.5) [128]. In a meta-analysis of smokeless tobacco in industrialized countries, with data mainly coming from CPS conducted in the USA, the current use of snuff and other smokeless tobacco products was significantly associated with fatal MI and stroke risk [129]. Pooled RR were respectively 1.13 (1.06; 1.21) and 1.40 (1.28; 1.54) taking never users as the reference group. The authors estimated that smokeless tobacco contributed to 5.6% of all fatal MI, and to 5.4% of all fatal strokes in Sweden which occurred in 2001. Combining information from both developing and developed countries, INTERHEART obtained an OR of acute MI of 2.23 for only users of smokeless tobacco compared to non-tobacco users; comparable to the excess risk experienced by current cigarette smokers in that study (OR: 2.95 for current cigarette smokers versus never users of tobacco) [81].

1.8 Aims of the thesis and outline

1.8.1 Strengths and weaknesses of the available epidemiological evidence

The evidence of a causal relationship between smoking and CVD risk is compelling and has been gathered over more than 50 years of epidemiological research since seminal papers in the UK and the USA [13]. However, some aspects of this relationship still remain unclear or subject to controversy. A stronger effect of smoking in women compared to men has been observed by some studies but not others, and the interplay between smoking and obesity, raised cholesterol and elevated blood pressure has been rarely studied due to requirements of large sample sizes to detect any significant interaction. The effect of smoking has been mostly investigated in relation to most common CVD including MI and cerebrovascular events, but uncertainty remains regarding the strength of an association between smoking and pulmonary embolism, aortic aneurysm and other rare types of CVD. Review of the literature on CVD and tobacco use showed that most studies investigated the effect of cigarette smoking, which is the most common form of tobacco use in developed countries. Therefore, the effect of pipes and cigar remain relatively unknown, and there have not been any large scale studies conducted in Western European populations on the effect of these other types of smoking. Regarding smokeless forms of tobacco which are relatively uncommon in developed countries but common in developing countries, there have been few studies investigating their effect on cardiovascular health. These studies have been either set up in the USA or Northern Europe and results may not be applicable to developing settings, or characterized by small sample size.

1.8.2 Aims of the thesis

In this context, the aim of this thesis is to investigate the association between several forms of tobacco and the risk of major cardiovascular events using epidemiological data from developed and developing countries. Smoking is the most common form of tobacco use in developed countries and the effect of cigarette smoking on the risk of CVD has already been investigated by several studies, so the objectives are to:

(1) Summarize the evidence on cigarette smoking with the risk of CVD in developed countries using meta-analyses an in a South Asian developing country, Pakistan.
(2) Investigate the effect of pipe and cigar smoking with the risk of CVD, two alternative forms of smoking popular in developed countries.

Regarding smokeless tobacco, its use remains mainly confined to the developing world and here the objectives are to:

(1) Look at the effect of chewing tobacco in relation to MI risk in South Asia.

(2) Look at the effect of dipping tobacco in relation to MI risk in South Asia.

1.8.3 Outline of the thesis

This thesis is organised in two sections. **Section A** focuses on developed countries, and **Section B** focuses on a developing country with a high prevalence of smoking and smokeless tobacco use as well as a substantial population, Pakistan. In **Section A**, the dataset used for my analysis is presented in **Chapter 2,** and is a large collaboration of prospective cohorts with detailed information on smoking. **Chapter 3** presents correlates of smoking with a range of other lifestyle and biochemical risk factors of CVD. **Chapter 4** looks at the associations between cigarette smoking and CVD, which is compared **in Chapter 5** to that of cigar and pipe smoking with CVD. In **Section B** of this thesis, the relationship between smoking and smokeless tobacco use with the risk of myocardial infarction is investigated in the context of Pakistan. The Pakistan Risk of Myocardial Infarction Study is presented in **Chapter 6**. Correlates of tobacco use with other conventional and locally relevant risk factors of MI are presented in **Chapter 7**. In **Chapter 8**, I investigate strength of the association between smoking and smokeless tobacco use with MI, with several levels of adjustment for confounders and tests of effect modifications by other risk factors. The **Discussion** exposes the limits of my results, and discusses their public health relevance. It also envisions future work in the field of tobacco use and CVD risk. Finally, in the **Appendices**, (1) a list of contributions to other analyses and resulting publications is given, (2) the association between socio-economic status and the risk of MI is presented, (3) the association between diet and the risk of MI using principal component analyses is presented, and (4) the association between the chromosome 9p21 and the risk of MI in Pakistanis is shown.

Figures and Tables

Figure 1.1: Main types of tobacco use in Pakistan

Smoking tobacco

Beedies

Cigarettes

Hookah

Chilum

Smokeless tobacco

Gutka

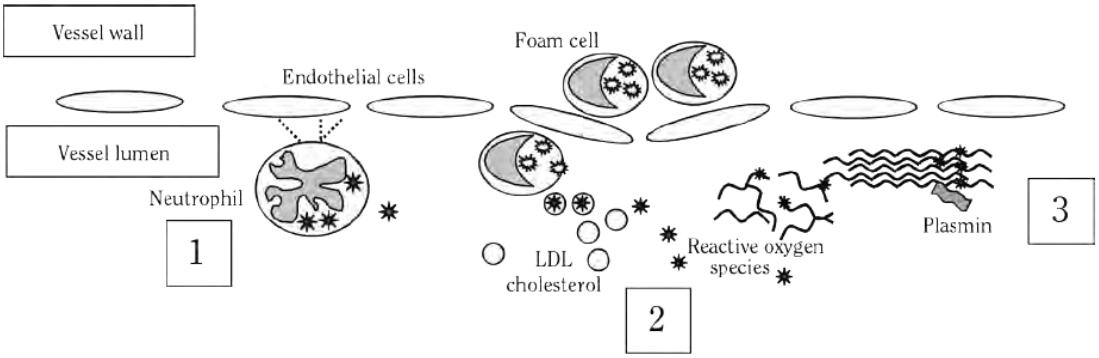
Paan

Supari

Naswar

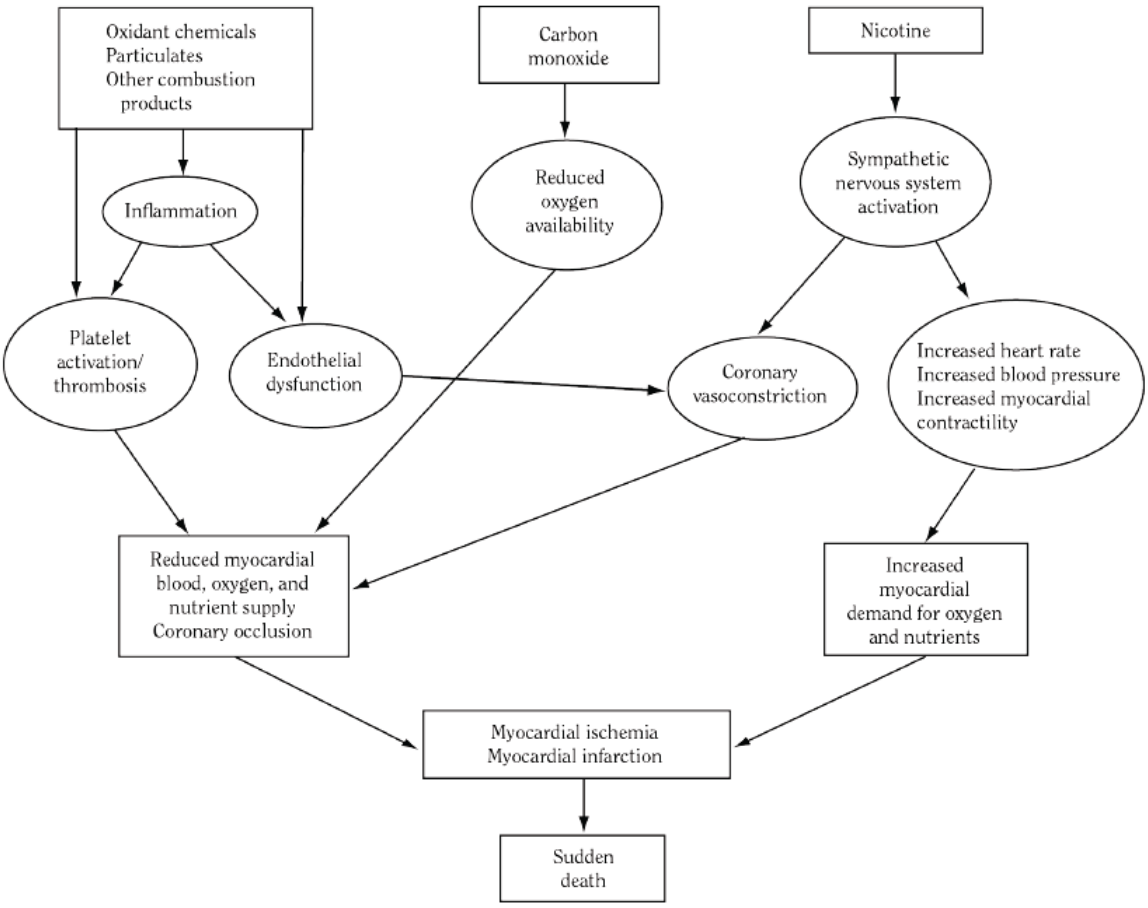
Sources: 3rd International Conference on Smokeless Tobacco 2002 “Fact sheet”.

Figure 1.2: Potential sites of effect of smoking on thrombosis through oxidative stress and other mechanisms



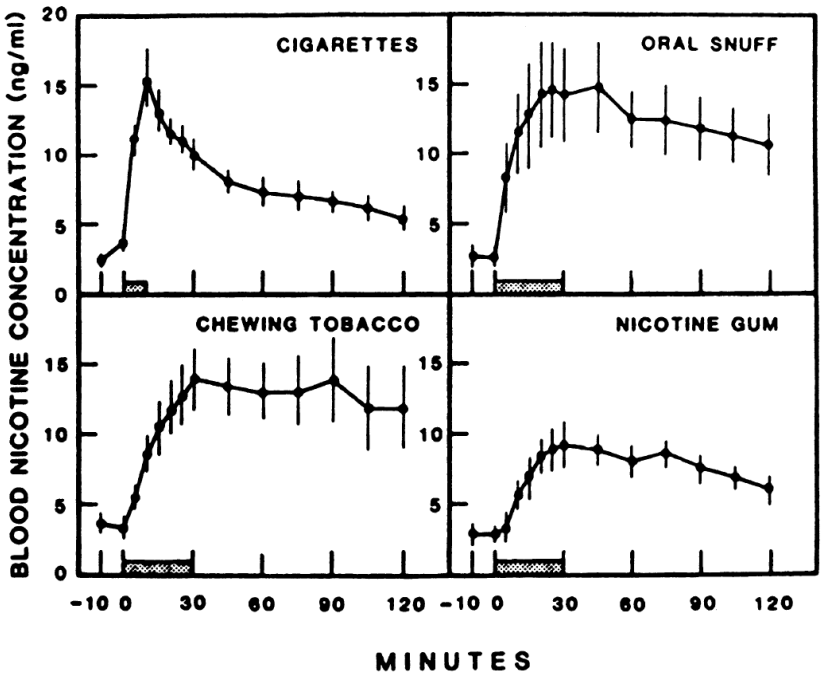
1. Increased number and activation of polymorphonuclear leukocytes; increased production of superoxide radicals; and increased expression of integrins and adhesion molecules on leukocytes and endothelial cells. 2. Increased oxidation of LDL-C; oxidized LDL-C taken up more easily into macrophages to produce foam cells; and increased adhesiveness of monocytes to endothelial cells. 3. Increased levels of fibrinogen; increased nitration of tyrosine residues on fibrinogen, rendering it more thrombogenic; impaired activity of plasmin; and decreased thrombolysis. Source: US Surgeon General Report, 2010 [130].

Figure 1.3: Overview of mechanisms by which cigarette smoking causes an acute cardiovascular event



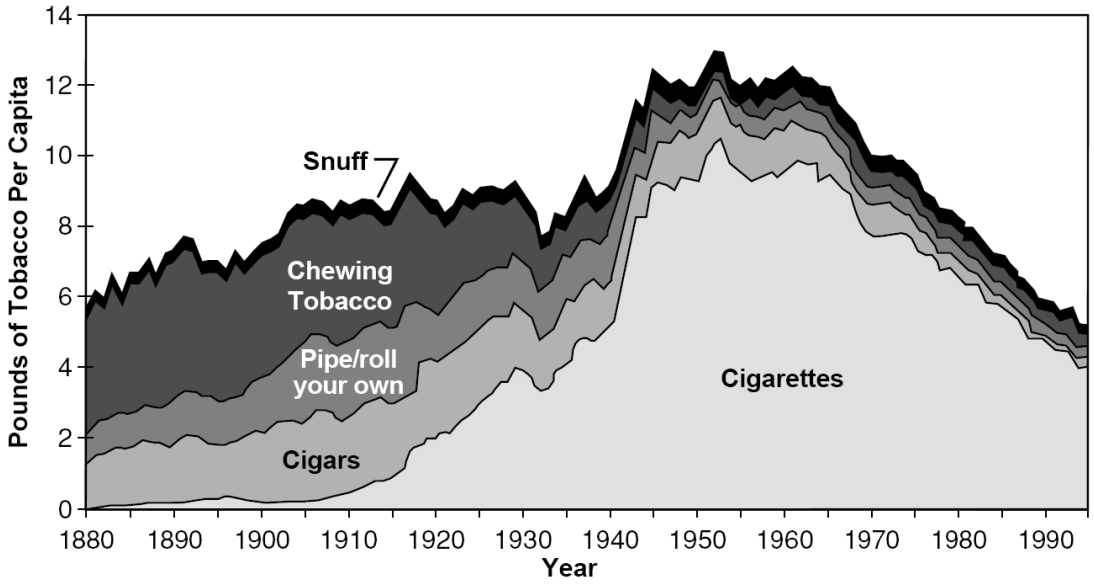
Source: Benowitz, 2003 [131].

Figure 1.4: Rapidity of absorption of nicotine according to types of tobacco use



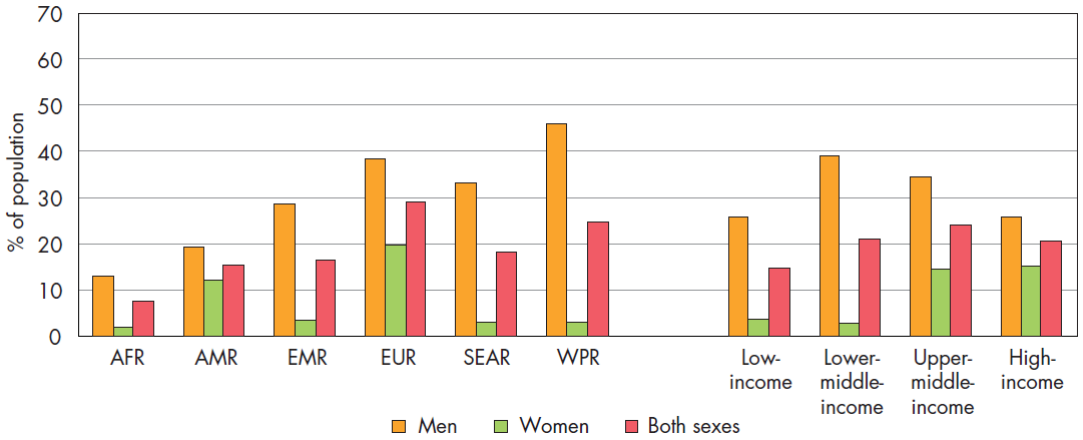
Legend: Mean (standard error) blood concentrations of nicotine in 10 subjects who smoked cigarettes for 9 minutes (1.3 cigarettes), used oral snuff (2.5g), used chewing tobacco (7.9 g) and chewed nicotine gum (4 mg). Shaded bars above the time axis indicate the period of exposure to tobacco or nicotine gum. Source: Benowitz, 1988 [85].

Figure 1.5: Per capita consumption of different forms of tobacco in the United States 1880-1995



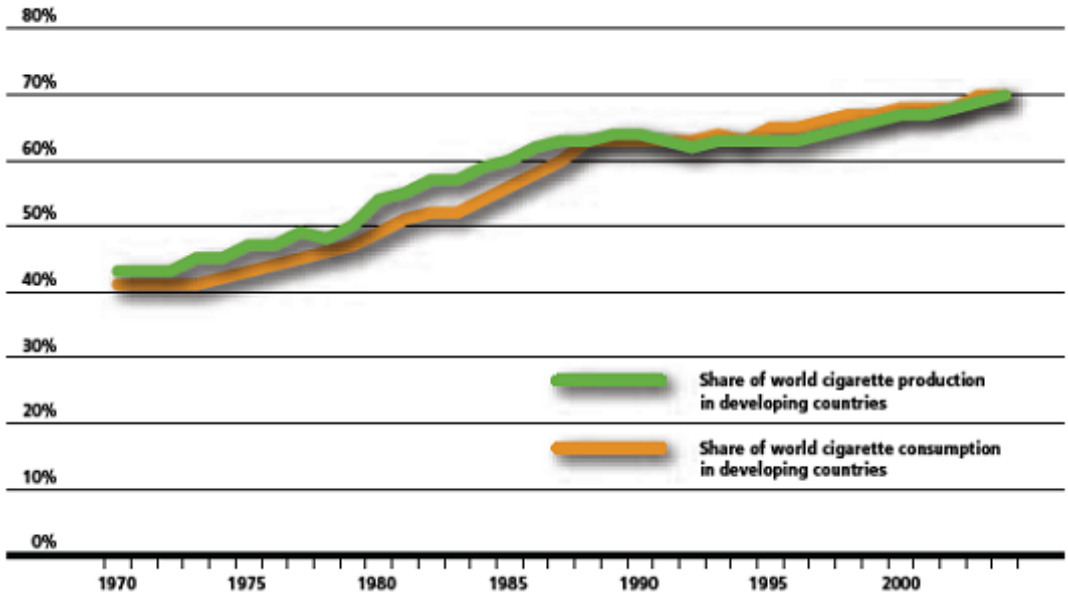
Source: U.S. Department of Agriculture 1996 [51]

Figure 1.6: Age-standardized prevalence of daily tobacco smoking in adults aged 15+ years, by WHO Region and World Bank income group, comparable estimates, 2008



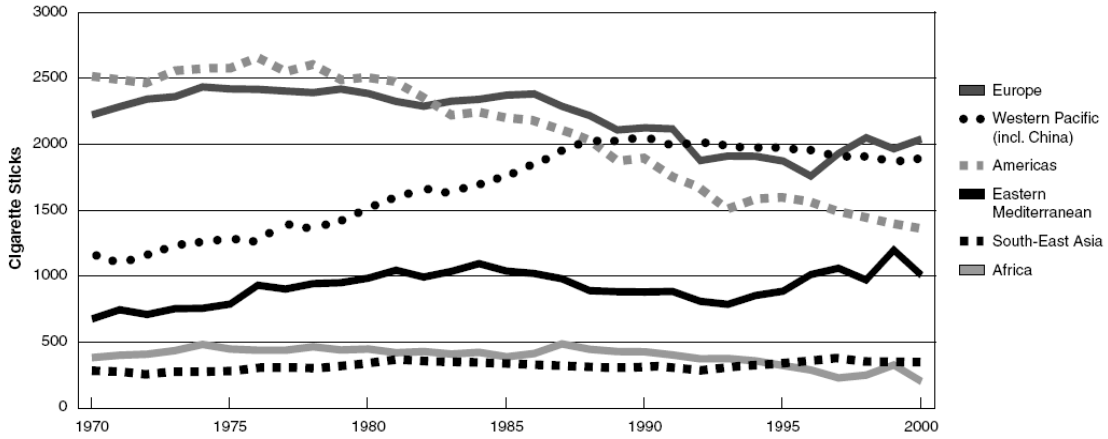
Note: AFR: African Region, AMR: Region of the Americas, EMR: Eastern Mediterranean Region, EUR: European Region, SEAR: South-East Asia Region, WPR: Western Pacific Region. World Bank income groups are created dividing all Member States into 4 income groups based on 2004 Gross National Income per capita: low, lower middle, upper middle, and high. Source: WHO Global status report on non-communicable diseases 2010 [7].

Figure 1.7: Share of cigarette production and consumption in developing countries.



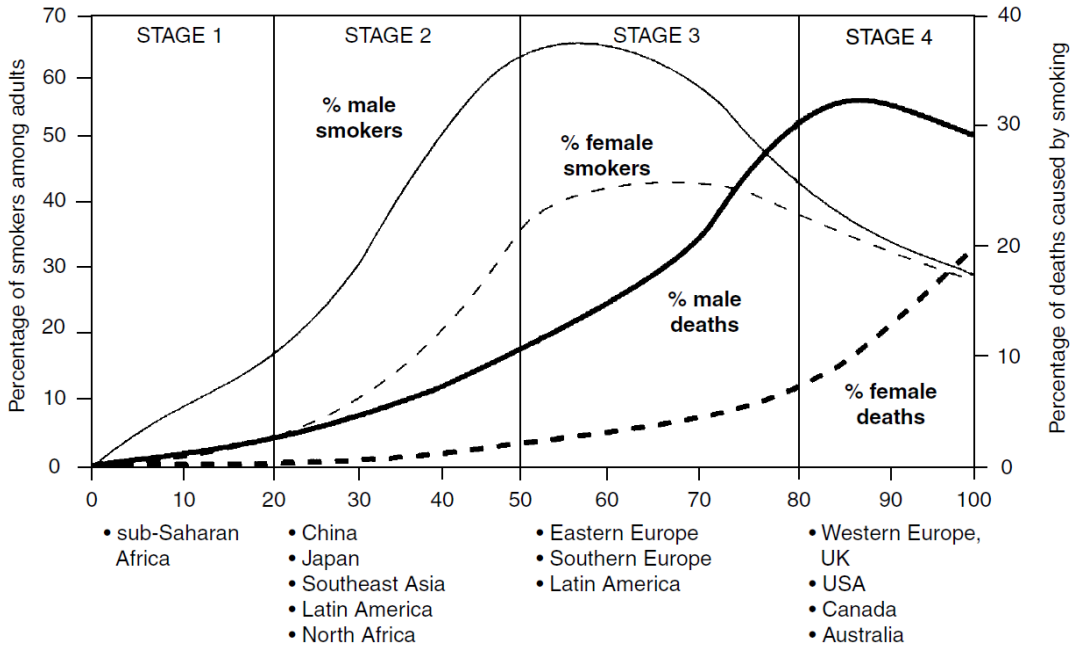
Source: Based on data from Food and Agriculture Organization FAOSTAT, United Nations Commodity Trade Statistics Database, United Nations Common Database, United States Department of Agriculture Economic Research Service, World Health Organization Statistical Information System, and ERC Group Plc.'s world Cigarettes Report 2005, extracted from WHO report on tobacco epidemic 2008 [132].

Figure 1.8: Average per capita cigarette consumption in persons aged ≥15 years by WHO region



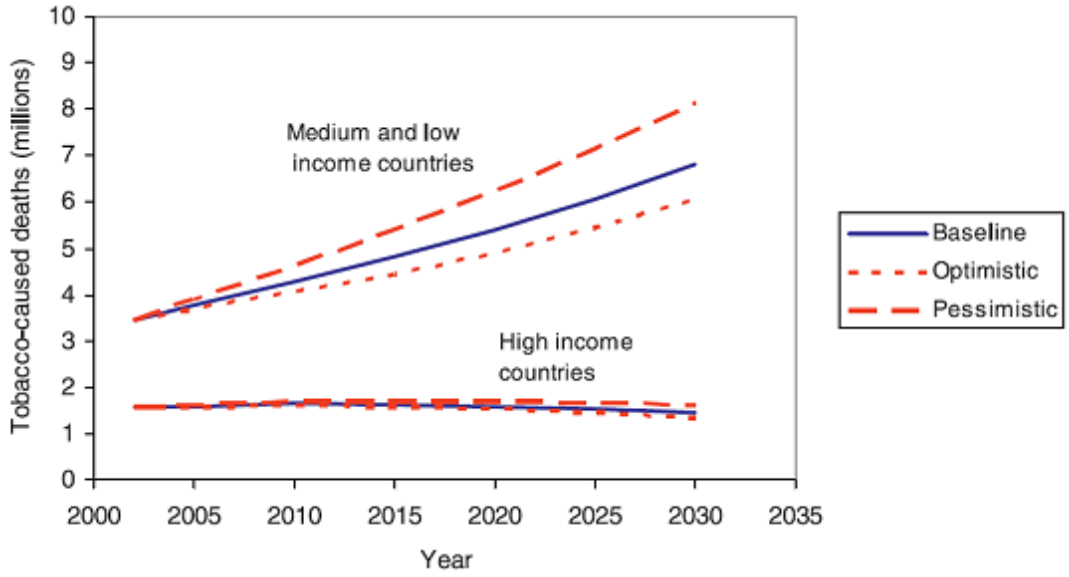
Source: United Nations Statistics Division. 2003. Commodity Trade Statistics Database [51]

Figure 1.9: Four stages of the tobacco epidemic



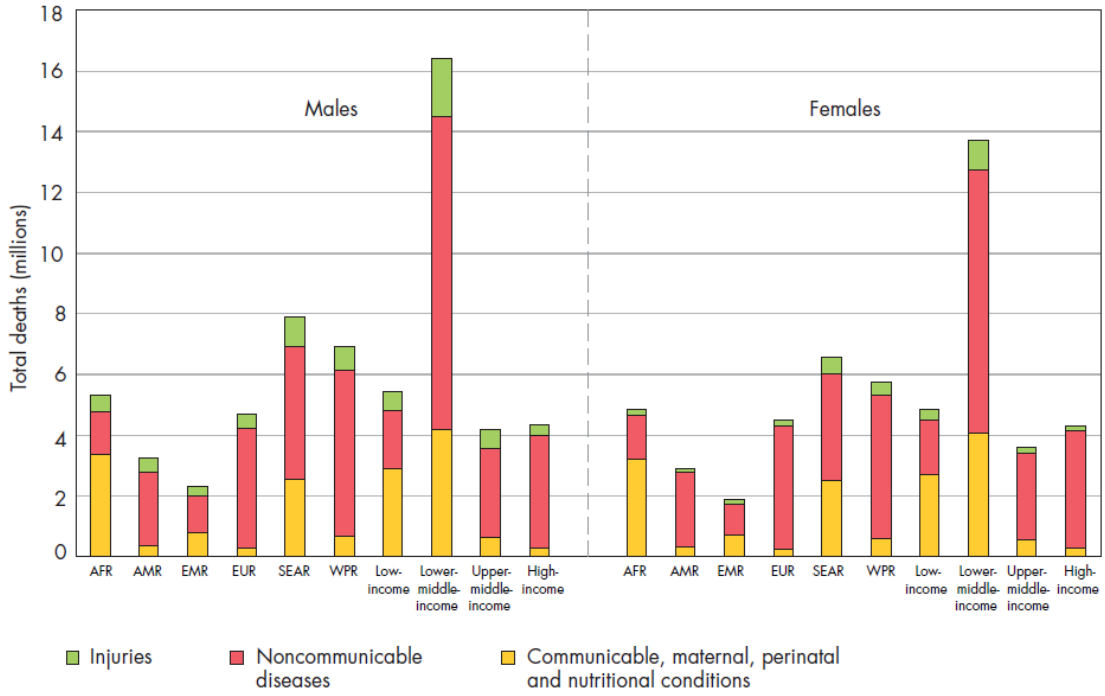
Note: x-axis indicates the number of years since smoking began. Source: Thun, 2012 [133].

Figure 1.10: Projected number of tobacco-related deaths for high and middle plus low income countries, 2002-2030



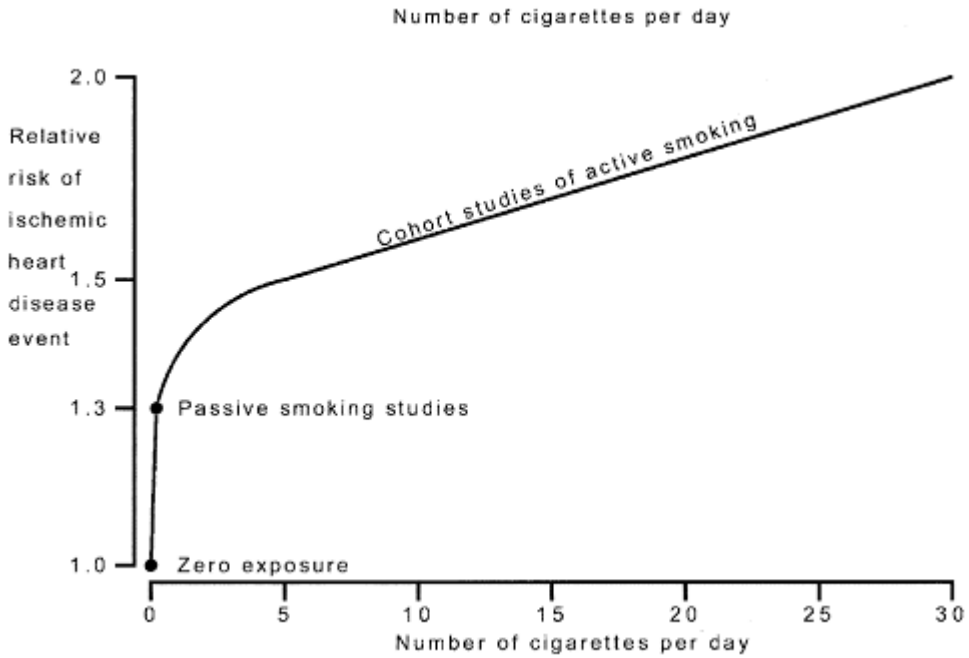
Source: Mathers 2006 Plos Medicine [62].

Figure 1.11: Total deaths by broad cause group, by WHO Region, World Bank income group and by sex, 2008



Note: AFR: African Region, AMR: Region of the Americas, EMR: Eastern Mediterranean Region, EUR: European Region, SEAR: South-East Asia Region, WPR: Western Pacific Region. World Bank income groups are created dividing all Member States into 4 income groups based on 2004 Gross National Income per capita: low, lower middle, upper middle, and high. Source: WHO Global status report on non-communicable diseases 2010 [104].

Figure 1.12: Model of the dose-response relationship between smoking and the risk of ischemic heart disease



Note: Summary of evidence from a meta-analysis of 5 large cohort studies of active smoking combined with the summary estimate from the studies of environmental tobacco smoke exposure (taken to be equivalent to actively smoking 0.2 cigarettes per day). Source: Law & Wald 2003 [134]

.

Table 1.1: Literature review of epidemiological studies on the association between tobacco use and CVD risk

Ref. no	Study acronym	Study full name	Study design	No. people recruited	Country	Baseline survey	Follow-up period	Number of CVD events
107	3CS	3 Copenhagen study (CCHS + Glostrup Population Study	PC	13,897 individuals	Denmark	1976-78 (CCHS) 1964 (GPS)	7-16 years	-
49, 71, 72, 103	ARIC	Atherosclerosis Risk in Community Study	PC	15,792 individuals aged 45-64 years	USA	1987-89	13 years	932 CHD events
96, 99	ASPC	Asia Pacific Studies Collaboration	Meta-analysis of 40 PCs	500,000 Asians and 100,000 Australians	Asia	1961-99	7 years	4183 fatal & non fatal MI; 5930 fatal & non fatal strokes
76	BDS	British Doctors Study	PC	35,000 Male	UK	1951	50 years	in 2001, 7628 CHD deaths, 3307 stroke deaths
129	BRHS	British Regional Heart Study	PC	7,735 individuals	UK	1978	20 years	1,766 fatal or non fatal CVD events at 15 years follow-up
135	CCHS	Copenhagen City Heart Study	PC	6,505 women and 5,644 men	Denmark	1976-78	21 years	1348 fatal and non fatal MI
136	CPS-1	Cancer Prevention Study 1	PC	~1 million individuals	USA	1960	6 years	-
105	CPS-2	Cancer Prevention Study 2	PC	~1 million individuals	USA	1982	6 years	14,585 CHD & 3,539 strokes deaths
82, 114, 137	EAS	Edinburgh Artery Study	PC	1,592 individuals aged 55-74 years	UK	1988	5 years	141 fatal & non fatal CHD
81	Finmark	Finmark Study	PC	11,843 individuals aged 35-52 years	Finland	1977	12 years	498 fatal & non fatal MI
111	Framingham	Framingham Heart Study	PC	4,255 individuals age 36-68 years	USA	1960	26 years	459 strokes
	GOTO	Goteborg study	PC	6879 men aged 47-55 years	Sweden	1970-80s	12 years	2277 CHD event
122	INTERHEART	INTERHEART	CC	12,461 MI cases and 14,637 controls	World	2000-2	-	12460 MI cases
101	Iowa Study	Iowa Study	PC	41,836 women	USA	1986	13 years	757 CHD deaths
87	KPS	Kaiser Permanente Medical Care Program	PC	60,838 individuals aged ≥35 years	USA	1979	8 years	-
112	MHS	Munster Heart Study (also called PROCAM)	PC	20,696 men and 10,212 women aged 40-65 years	Germany	1978-95	8 years	39 CHD events
138	MRFIT	Multiple Risk Factors Intervention Trial	Clinical trial	12,866 men	USA	1980s	10 years	-
74	MSA	Multiethnic Study of Atherosclerosis	CS	6,384 individuals aged 45-84 years	USA	2000-2	-	-
102	NHANES-III	Third National Health and Examination Survey	CS	33,994 individuals	USA	1988-1994	-	-
85	NHS	Nurses' Health Study	PC	100,000 female nurses	USA	1980	24 years	1385 CHD & 734 strokes deaths
121	SCS	Seven Countries Study	PC	12,763 men aged 40-89 years	Europe, USA, Japan	1957-1964	25 years	1827 CHD & 797 stroke deaths
70	USVS	US Veteran Study	PC	300,000 men	USA	1954	26 years	16,586 CVD deaths
78	Zutphen	Zutphen study	PC	1373 men	Netherlands	1960	40 years	-
	-	Hammon & Horn study	PC	188,000 white men age 50-69 years old	USA	1952	6 years	5,297 CVD deaths
	Norway	Oslo city and 3 counties in Norway	PC	23,521 men and 19,201 women aged 35-49 years	Norway	1970	32 years	2253 CHD deaths
116	ISIS	International Study of Infarct Survival	CC	14,000 MI cases and 32,000 controls	UK	1990s	-	-

Ref. no: Reference number; CC: case-control study. PC: prospective cohort study; CS: cross-sectional study; CT: clinical trial.

References

(1) www.who.int/fctc/en. WHO Framework Convention on Tobacco Control. 24-8-2012.

(2) World Health Organization. *Tobacco: deadly in any form or disguise*. 2006.

(3) Baker F, Ainsworth SR, Dye JT et al. Health risks associated with cigar smoking. *Stat Med* 2002 June 15;21(11):1539-58.

(4) US code of federal regulation. 2011 April 24;Available at: URL: http://cfr.vlex.com/vid/40-11-meaning-terms-19672629.

(5) Advocacy forum for tobacco control. 2011 May 1;Available at: URL: http://www.aftcindia.org/tob_pro_india.htm .

(6) Ambrose JA, Barua RS. The pathophysiology of cigarette smoking and cardiovascular disease: an update. *J Am Coll Cardiol* 2004 May;%19;43(10):1731-7.

(7) World Health Organization. WHO report on the global tobacco epidemic; the MPOWER package. 2008.

(8) Gupta PC, Hamner JE. *Control of tobacco-related cancers and other diseases. International Symposium*. 1992.

(9) Third International Conference on Smokeless Tobacco; Fact sheet. 2002.

(10) The British Medical Association. *Illustrated Medical Dictionary*. 2nd edition ed. 2007.

(11) Aaronson PI, Ward JPT. *The cardiovascular system at a glance*. 3rd edition ed. 2007.

(12) Libby, Bonow, Mann, Zipes. *Braunwald's heart disease: a textbook of cardiovascular disease*. 8th edition Saunders Elsevier ed. 2009.

(13) Licastro F, Candore G, Lio D et al. Innate immunity and inflammation in ageing: a key for understanding age-related diseases. *Immun Ageing* 2005 May 18;2:8.:8.

(14) Benowitz NL. Cigarette smoking and cardiovascular disease: pathophysiology and implications for treatment. *Prog Cardiovasc Dis* 2003 July;46(1):91-111.

(15) Craig WY, Palomaki GE, Haddow JE. Cigarette smoking and serum lipid and lipoprotein concentrations: an analysis of published data. *BMJ* 1989 March 25;298(6676):784-8.

(16) Harats D, Ben-Naim M, Dabach Y, Hollander G, Stein O, Stein Y. Cigarette smoking renders LDL susceptible to peroxidative modification and enhanced metabolism by macrophages. *Atherosclerosis* 1989 October;79(2-3):245-52.

(17) Libby P, Ridker PM, Maseri A. Inflammation and atherosclerosis. *Circulation* 2002 March 5;105(9):1135-43.

(18) Wannamethee SG, Lowe GD, Shaper AG, Rumley A, Lennon L, Whincup PH. Associations between cigarette smoking, pipe/cigar smoking, and smoking cessation, and haemostatic and inflammatory markers for cardiovascular disease. *Eur Heart J* 2005 September;26(17):1765-73.

(19) Lehr HA, Frei B, Arfors KE. Vitamin C prevents cigarette smoke-induced leukocyte aggregation and adhesion to endothelium in vivo. *Proc Natl Acad Sci U S A* 1994 August 2;91(16):7688-92.

(20) Adams MR, Jessup W, Celermajer DS. Cigarette smoking is associated with increased human monocyte adhesion to endothelial cells: reversibility with oral L-arginine but not vitamin C. *J Am Coll Cardiol* 1997 March 1;29(3):491-7.

(21) Metz L, Waters DD. Implications of cigarette smoking for the management of patients with acute coronary syndromes. *Prog Cardiovasc Dis* 2003 July;46(1):1-9.

(22) Burke A, Fitzgerald GA. Oxidative stress and smoking-induced vascular injury. *Prog Cardiovasc Dis* 2003 July;46(1):79-90.

(23) Matetzky S, Tani S, Kangavari S et al. Smoking increases tissue factor expression in atherosclerotic plaques: implications for plaque thrombogenicity. *Circulation* 2000 August 8;102(6):602-4.

(24) Xie XT, Liu Q, Wu J, Wakui M. Impact of cigarette smoking in type 2 diabetes development. *Acta Pharmacol Sin* 2009 June;30(6):784-7.

(25) Rahman MM, Laher I. Structural and functional alteration of blood vessels caused by cigarette smoking: an overview of molecular mechanisms. *Curr Vasc Pharmacol* 2007 October;5(4):276-92.

(26) Heeschen C, Jang JJ, Weis M et al. Nicotine stimulates angiogenesis and promotes tumor growth and atherosclerosis. *Nat Med* 2001 July;7(7):833-9.

(27) Henningfield JE, Stapleton JM, Benowitz NL, Grayson RF, London ED. Higher levels of nicotine in arterial than in venous blood after cigarette smoking. *Drug Alcohol Depend* 1993 June;33(1):23-9.

(28) Arcavi L, Jacob P, III, Hellerstein M, Benowitz NL. Divergent tolerance to metabolic and cardiovascular effects of nicotine in smokers with low and high levels of cigarette consumption. *Clin Pharmacol Ther* 1994 July;56(1):55-64.

(29) Jolma CD, Samson RA, Klewer SE, Donnerstein RL, Goldberg SJ. Acute cardiac effects of nicotine in healthy young adults. *Echocardiography* 2002 August;19(6):443-8.

(30) Teter CJ, Asfaw B, Ni L, Lutz M, Domino EF, Guthrie SK. Comparative effects of tobacco smoking and nasal nicotine. *Eur J Clin Pharmacol* 2002 August;58(5):309-14.

(31) Symons JD, Stebbins CL. Hemodynamic and regional blood flow responses to nicotine at rest and during exercise. *Med Sci Sports Exerc* 1996 April;28(4):457-67.

(32) Sabha M, Tanus-Santos JE, Toledo JC, Cittadino M, Rocha JC, Moreno H, Jr. Transdermal nicotine mimics the smoking-induced endothelial dysfunction. *Clin Pharmacol Ther* 2000 August;68(2):167-74.

(33) Csonka E, Somogyi A, Augustin J, Haberbosch W, Schettler G, Jellinek H. The effect of nicotine on cultured cells of vascular origin. *Virchows Arch A Pathol Anat Histopathol* 1985;407(4):441-7.

(34) Krupski WC, Olive GC, Weber CA, Rapp JH. Comparative effects of hypertension and nicotine on injury-induced myointimal thickening. *Surgery* 1987 August;102(2):409-15.

(35) Aicher A, Heeschen C, Mohaupt M, Cooke JP, Zeiher AM, Dimmeler S. Nicotine strongly activates dendritic cell-mediated adaptive immunity: potential role for progression of atherosclerotic lesions. *Circulation* 2003 February 4;107(4):604-11.

(36) Axelsson T, Jansson PA, Smith U, Eliasson B. Nicotine infusion acutely impairs insulin sensitivity in type 2 diabetic patients but not in healthy subjects. *J Intern Med* 2001 June;249(6):539-44.

(37) Eliasson B, Taskinen MR, Smith U. Long-term use of nicotine gum is associated with hyperinsulinemia and insulin resistance. *Circulation* 1996 September 1;94(5):878-81.

(38) Louie D. The effects of cigarette smoking on cardiopulmonary function and exercise tolerance in teenagers. *Can Respir J* 2001 July;8(4):289-91.

(39) Sheps DS, Herbst MC, Hinderliter AL et al. Production of arrhythmias by elevated carboxyhemoglobin in patients with coronary artery disease. *Ann Intern Med* 1990 September 1;113(5):343-51.

(40) DeBias DA, Banerjee CM, Birkhead NC, Greene CH, Scott SD, Harrer WV. Effects of carbon monoxide inhalation on ventricular fibrillation. *Arch Environ Health* 1976 January;31(1):42-6.

(41) Yanbaeva DG, Dentener MA, Creutzberg EC, Wesseling G, Wouters EF. Systemic effects of smoking. *Chest* 2007 May;131(5):1557-66.

(42) Newby DE, McLeod AL, Uren NG et al. Impaired coronary tissue plasminogen activator release is associated with coronary atherosclerosis and cigarette smoking: direct link between endothelial dysfunction and atherothrombosis. *Circulation* 2001 April 17;103(15):1936-41.

(43) Bauer V, Sotnikova R. Nitric oxide--the endothelium-derived relaxing factor and its role in endothelial functions. *Gen Physiol Biophys* 2010 December;29(4):319-40.

(44) Penn A, Snyder C. Arteriosclerotic plaque development is 'promoted' by polynuclear aromatic hydrocarbons. *Carcinogenesis* 1988 December;9(12):2185-9.

(45) Penn A, Snyder CA. Butadiene inhalation accelerates arteriosclerotic plaque development in cockerels. *Toxicology* 1996 October 28;113(1-3):351-4.

(46) World Health Organization. *The Tobacco Atlas*. 2002.

(47) CNN's brief history of tobacco. *http://edition cnn com/US/9705/tobacco/history/index.html* Accessed on 2011 April 28.

(48) White C. Research on smoking and lung cancer: a landmark in the history of chronic disease epidemiology. *The Yale Journal of Biology and Medicine* 1990;(63):29-46.

(49) Doll R, HILL AB. The mortality of doctors in relation to their smoking habits: a preliminary report. *BMJ* 1954;328(7455):1529-33.

(50) Pierce JP, Messer K, White MM, Cowling DW, Thomas DP. Prevalence of heavy smoking in California and the United States, 1965-2007. *JAMA* 2011 March 16;305(11):1106-12.

(51) World Health Organization. Global status report on noncommunicable diseases 2010. 2011 Apr.

(52) Allender S, Scarborough P, Peto V et al. European cardiovascular disease statistics: 2008 edition. 2008.

(53) O'Connor RJ, McNeill A, Borland R et al. Smokers' beliefs about the relative safety of other tobacco products: findings from the ITC collaboration. *Nicotine Tob Res* 2007 October;9(10):1033-42.

(54) Delnevo CD. Smokers' choice: what explains the steady growth of cigar use in the U.S.? *Public Health Rep* 2006 March;121(2):116-9.

(55) Tomar SL. The prevalence of smokeless tobacco use remained low and unchanged among US men in 2000-2005 while smoking declined. *J Evid Based Dent Pract* 2010 September;10(3):183-5.

(56) Nishtar S. Prevention of coronary heart disease in south Asia. *Lancet* 2002 September 28;360(9338):1015-8.

(57) IRIN. Pakistan: anti-smoking law have little impact. *IRIN Humanitarian news and analysis* 2004 July 14.

(58) Simpson D. India: gutka plastic pack ban. *Tobacco Control* 2011;20:89-92.

(59) Liu Y, Chen L. New medical data and leadership on tobacco control in China. *Lancet* 2011 March 23.

(60) Giovino GA, Mirza SA, Samet JM et al. Tobacco use in 3 billion individuals from 16 countries: an analysis of nationally representative cross-sectional household surveys. *Lancet* 2012 August 18;380(9842):668-79.

(61) Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. *Lancet* 2003 September 13;362(9387):847-52.

(62) Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med* 2006 November;3(11):e442.

(63) Murray CJ, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet* 1997 May 17;349(9063):1436-42.

(64) Katanoda K, Marugame T, Saika K et al. Population attributable fraction of mortality associated with tobacco smoking in Japan: a pooled analysis of three large-scale cohort studies. *J Epidemiol* 2008;18(6):251-64.

(65) Liu BQ, Peto R, Chen ZM et al. Emerging tobacco hazards in China: 1. Retrospective proportional mortality study of one million deaths. *BMJ* 1998 November 21;317(7170):1411-22.

(66) Ezzati M, Henley SJ, Thun MJ, Lopez AD. Role of smoking in global and regional cardiovascular mortality. *Circulation* 2005 July 26;112(4):489-97.

(67) National Institute on drug abuse. NIDA InfoFacts: Cigarettes and other tobacco products. [http://www.nida.nih.gov/infofacts/tobacco.html#References], Accessed on 01/06/2010.

(68) Best CM, Sun K, de PS, Bloem MW, Stallkamp G, Semba RD. Parental tobacco use is associated with increased risk of child malnutrition in Bangladesh. *Nutrition* 2007 October;23(10):731-8.

(69) Hu TW, Mao Z, Liu Y, de BJ, Ong M. Smoking, standard of living, and poverty in China. *Tob Control* 2005 August;14(4):247-50.

(70) Hammond EC, Horn D. Smoking and death rates: report on forty-four months of follow-up of 187,783 men. 2. Death rates by cause. *J Am Med Assoc* 1958 March 15;166(11):1294-308.

(71) Doll R, Hill AB. Mortality in relation to smoking: ten years' observations of British Doctors. *Br Med J* 1964 May 30;1(5395):1399-410.

(72) Doll R, Peto R. Mortality in relation to smoking: 20 years' observations on male British doctors. *Br Med J* 1976 December 25;2(6051):1525-36.

(73) U.S.Department of Health and Human Services. *How tobacco smoke causes disease: the biology and behavioral basis of smoking-attributable disease: a report of the surgeon general*. 2010.

(74) Kenfield SA, Stampfer MJ, Rosner BA, Colditz GA. Smoking and smoking cessation in relation to mortality in women. *JAMA* 2008 May 7;299(17):2037-47.

(75) Teo KK, Ounpuu S, Hawken S et al. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet* 2006 August;368(9536):647-58.

(76) Prescott E, Scharling H, Osler M, Schnohr P. Importance of light smoking and inhalation habits on risk of myocardial infarction and all cause mortality. A 22 year follow up of 12 149 men and women in The Copenhagen City Heart Study. *J Epidemiol Community Health* 2002 September;56(9):702-6.

(77) Willett WC, Green A, Stampfer MJ et al. Relative and absolute excess risks of coronary heart disease among women who smoke cigarettes. *N Engl J Med* 1987 November;317(21):1303-9.

(78) Bjartveit K, Tverdal A. Health consequences of smoking 1-4 cigarettes per day. *Tob Control* 2005 October;14(5):315-20.

(79) Gallo V, Neasham D, Airoldi L et al. Second-hand smoke, cotinine levels, and risk of circulatory mortality in a large cohort study of never-smokers. *Epidemiology* 2010 March;21(2):207-14.

(80) Whincup PH, Gilg JA, Emberson JR et al. Passive smoking and risk of coronary heart disease and stroke: prospective study with cotinine measurement. *BMJ* 2004 July 24;329(7459):200-5.

(81) Teo KK, Ounpuu S, Hawken S et al. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet* 2006 August;368(9536):647-58.

(82) Wolf PA, D'Agostino RB, Kannel WB, Bonita R, Belanger AJ. Cigarette smoking as a risk factor for stroke. The Framingham Study. *JAMA* 1988 February;259(7):1025-9.

(83) Neaton JD, Kuller LH, Wentworth D, Borhani NO. Total and cardiovascular mortality in relation to cigarette smoking, serum cholesterol concentration, and diastolic blood pressure among black and white males followed up for five years. *Am Heart J* 1984 September;108(3 Pt 2):759-69.

(84) Bhat VM, Cole JW, Sorkin JD et al. Dose-response relationship between cigarette smoking and risk of ischemic stroke in young women. *Stroke* 2008 September;39(9):2439-43.

(85) U.S. National Cancer Institutes of Health. *Smoking and tobacco control: eighth monograph*. Feb. 1997. Available online at wwww.cancer.gov.

(86) Caraballo RS, Giovino GA, Pechacek TF et al. Racial and ethnic differences in serum cotinine levels of cigarette smokers: Third National Health and Nutrition Examination Survey, 1988-1991. *JAMA* 1998 July 8;280(2):135-9.

(87) Kuller LH, Ockene JK, Meilahn E, Wentworth DN, Svendsen KH, Neaton JD. Cigarette smoking and mortality. MRFIT Research Group. *Prev Med* 1991 September;20(5):638-54.

(88) Tverdal A. Calculation of risk for the development of acute myocardial infarction in the normal population based on long-term follow-up studies: smokers compared with non-smokers. *J Cardiovasc Risk* 1999 October;6(5):287-91.

(89) Prescott E, Bjerg AM, Andersen PK, Lange P, Vestbo J. Gender difference in smoking effects on lung function and risk of hospitalization for COPD: results from a Danish longitudinal population study. *Eur Respir J* 1997 April;10(4):822-7.

(90) Thun MJ, Apicella LF, Henley SJ. Smoking vs other risk factors as the cause of smoking-attributable deaths: confounding in the courtroom. *JAMA* 2000 August 9;284(6):706-12.

(91) Dept.of Health and Human Services. The health consequences of smoking: a report of the Surgeon General. 2004.

(92) Woodward M, Lam TH, Barzi F et al. Smoking, quitting, and the risk of cardiovascular disease among women and men in the Asia-Pacific region. *Int J Epidemiol* 2005 October;34(5):1036-45.

(93) Anand SS, Islam S, Rosengren A et al. Risk factors for myocardial infarction in women and men: insights from the INTERHEART study. *Eur Heart J* 2008 April;29(7):932-40.

(94) Thun MJ, Day-Lally C, Myers D et al. Trends in tobacco smoking and mortality from cigarette use in Cancer Prevention Studies I (1959 through 1965) and II (1982 through 1988). Changes in Cigarette-Related Disease Risks and Their Implication for Prevention and Control.Smoking and Tobacco Control Monograph No.8.Rockville (MD) [*a*], 305-382. 1997. Shopland DR, Burns DM, Garfinkel L, Samet JM.

(95) Godtfredsen NS, Osler M, Vestbo J, Andersen I, Prescott E. Smoking reduction, smoking cessation, and incidence of fatal and non-fatal myocardial infarction in Denmark 1976-1998: a pooled cohort study. *J Epidemiol Community Health* 2003 June;57(6):412-6.

(96) Cook DG, Shaper AG, Pocock SJ, Kussick SJ. Giving up smoking and the risk of heart attacks. A report from The British Regional Heart Study. *Lancet* 1986 December 13;2(8520):1376-80.

(97) Bakhru A, Erlinger TP. Smoking cessation and cardiovascular disease risk factors: results from the Third National Health and Nutrition Examination Survey. *PLoS Med* 2005 June;2(6):e160.

(98) Wannamethee SG, Shaper AG, Walker M, Ebrahim S. Lifestyle and 15-year survival free of heart attack, stroke, and diabetes in middle-aged British men. *Arch Intern Med* 1998 December 7;158(22):2433-40.

(99) Shaper AG, Pocock SJ, Walker M, Phillips AN, Whitehead TP, Macfarlane PW. Risk factors for ischaemic heart disease: the prospective phase of the British Regional Heart Study. *J Epidemiol Community Health* 1985 September;39(3):197-209.

(100) Price JF, Mowbray PI, Lee AJ, Rumley A, Lowe GD, Fowkes FG. Relationship between smoking and cardiovascular risk factors in the development of peripheral arterial disease and coronary artery disease: Edinburgh Artery Study. *Eur Heart J* 1999 March;20(5):344-53.

(101) Cullen P, Schulte H, Assmann G. Smoking, lipoproteins and coronary heart disease risk. Data from the Munster Heart Study (PROCAM). *Eur Heart J* 1998 November;19(11):1632-41.

(102) Jacobs DR, Jr., Adachi H, Mulder I et al. Cigarette smoking and mortality risk: twenty-five-year follow-up of the Seven Countries Study. *Arch Intern Med* 1999 April 12;159(7):733-40.

(103) Doll R, Gray R, Hafner B, Peto R. Mortality in relation to smoking: 22 years' observations on female British doctors. *Br Med J* 1980 April 5;280(6219):967-71.

(104) Prescott E, Hippe M, Schnohr P, Hein HO, Vestbo J. Smoking and risk of myocardial infarction in women and men: longitudinal population study. *BMJ* 1998 April 4;316(7137):1043-7.

(105) Njolstad I, Arnesen E, Lund-Larsen PG. Smoking, serum lipids, blood pressure, and sex differences in myocardial infarction. A 12-year follow-up of the Finnmark Study. *Circulation* 1996 February 1;93(3):450-6.

(106) Hozawa A, Folsom AR, Sharrett AR, Payne TJ, Chambless LE. Does the impact of smoking on coronary heart disease differ by low-density lipoprotein cholesterol level?: the Atherosclerosis Risk in Communities (ARIC) Study. *Circ J* 2006 September;70(9):1105-10.

(107) Nakamura K, Barzi F, Huxley R et al. Does cigarette smoking exacerbate the effect of total cholesterol and high-density lipoprotein cholesterol on the risk of cardiovascular diseases? *Heart* 2009 June;95(11):909-16.

(108) Karim R, Buchanan TA, Hodis HN, Li Y, Mack WJ. The association of smoking and subclinical atherosclerosis in Type 2 diabetes: modification by duration of diabetes. *Diabet Med* 2005 January;22(1):81-7.

(109) Richey SA, Coady SA, Folsom AR, Couper DJ, Heiss G. Smoking and diabetes differ in their associations with subclinical atherosclerosis and coronary heart disease-the ARIC Study. *Atherosclerosis* 2004 January;172(1):143-9.

(110) Xu WH, Zhang XL, Gao YT et al. Joint effect of cigarette smoking and alcohol consumption on mortality. *Prev Med* 2007 October;45(4):313-9.

(111) Ebbert JO, Janney CA, Sellers TA, Folsom AR, Cerhan JR. The association of alcohol consumption with coronary heart disease mortality and cancer incidence varies by smoking history. *J Gen Intern Med* 2005 January;20(1):14-20.

(112) Sharrett AR, Ding J, Criqui MH et al. Smoking, diabetes, and blood cholesterol differ in their associations with subclinical atherosclerosis: the Multiethnic Study of Atherosclerosis (MESA). *Atherosclerosis* 2006 June;186(2):441-7.

(113) Stratton K, Shetty P, Wallace R, Bondurant S. Clearing the smoke: the science base for tobacco harm reduction--executive summary. *Tob Control* 2001 June;10(2):189-95.

(114) Castelli WP, Garrison RJ, Dawber TR, McNamara PM, Feinleib M, Kannel WB. The filter cigarette and coronary heart disease: the Framingham study. *Lancet* 1981 July 18;2(8238):109-13.

(115) Baldassarre D, Castelnuovo S, Frigerio B et al. Effects of timing and extent of smoking, type of cigarettes, and concomitant risk factors on the association between smoking and subclinical atherosclerosis. *Stroke* 2009 June;40(6):1991-8.

(116) Gajalakshmi V, Peto R, Kanaka TS, Jha P. Smoking and mortality from tuberculosis and other diseases in India: retrospective study of 43000 adult male deaths and 35000 controls. *Lancet* 2003 August 16;362(9383):507-15.

(117) Gallus S, Randi G, Negri E, Tavani A, La VC. Tar yield and risk of acute myocardial infarction: pooled analysis from three case-control studies. *Eur J Cardiovasc Prev Rehabil* 2007 April;14(2):299-303.

(118) Pais P, Pogue J, Gerstein H et al. Risk factors for acute myocardial infarction in Indians: a case-control study. *Lancet* 1996 August 10;348(9024):358-63.

(119) Pais P, Fay MP, Yusuf S. Increased risk of acute myocardial infarction associated with beedi and cigarette smoking in Indians: final report on tobacco risks from a case-control study. *Indian Heart J* 2001 November;53(6):731-5.

(120) Shaper AG, Wannamethee SG, Walker M. Pipe and cigar smoking and major cardiovascular events, cancer incidence and all-cause mortality in middle-aged British men. *Int J Epidemiol* 2003 October;32(5):802-8.

(121) Streppel MT, Boshuizen HC, Ocke MC, Kok FJ, Kromhout D. Mortality and life expectancy in relation to long-term cigarette, cigar and pipe smoking: the Zutphen Study. *Tob Control* 2007 April;16(2):107-13.

(122) Iribarren C, Tekawa IS, Sidney S, Friedman GD. Effect of cigar smoking on the risk of cardiovascular disease, chronic obstructive pulmonary disease, and cancer in men. *N Engl J Med* 1999 June 10;340(23):1773-80.

(123) Jacobs EJ, Thun MJ, Apicella LF. Cigar smoking and death from coronary heart disease in a prospective study of US men. *Arch Intern Med* 1999 November 8;159(20):2413-8.

(124) Henley SJ, Thun MJ, Chao A, Calle EE. Association between exclusive pipe smoking and mortality from cancer and other diseases. *J Natl Cancer Inst* 2004 June 2;96(11):853-61.

(125) Eliasson M, Asplund K, Evrin PE, Lundblad D. Relationship of cigarette smoking and snuff dipping to plasma fibrinogen, fibrinolytic variables and serum insulin. The Northern Sweden MONICA Study. *Atherosclerosis* 1995 February;113(1):41-53.

(126) Wallenfeldt K, Hulthe J, Bokemark L, Wikstrand J, Fagerberg B. Carotid and femoral atherosclerosis, cardiovascular risk factors and C-reactive protein in relation to

smokeless tobacco use or smoking in 58-year-old men. *J Intern Med* 2001 December;250(6):492-501.

(127) Stegmayr B, Johansson I, Huhtasaari F, Moser U, Asplund K. Use of smokeless tobacco and cigarettes--effects on plasma levels of antioxidant vitamins. *Int J Vitam Nutr Res* 1993;63(3):195-200.

(128) Rahman MA, Zaman MM. Smoking and smokeless tobacco consumption: possible risk factors for coronary heart disease among young patients attending a tertiary care cardiac hospital in Bangladesh. *Public Health* 2008 December;122(12):1331-8.

(129) Boffetta P, Straif K. Use of smokeless tobacco and risk of myocardial infarction and stroke: systematic review with meta-analysis. *BMJ* 2009 August 18;339.

(130) Boersma E, Mercado N, Poldermans D, Gardien M, Vos J, Simoons ML. Acute myocardial infarction. *Lancet* 2003 March 8;361(9360):847-58.

(131) Benowitz NL, Porchet H, Sheiner L, Jacob P, III. Nicotine absorption and cardiovascular effects with smokeless tobacco use: comparison with cigarettes and nicotine gum. *Clin Pharmacol Ther* 1988 July;44(1):23-8.

(132) Shafey O, Dolwick S, Guindon EG, World Health Organization, American Cancer Society, International Union Against Cancer. *Tobacco control country profile*. 2nd ed. 2003.

(133) Thun M, Peto R, Boreham J, Lopez AD. Stages of the cigarette epidemic on entering its second century. *Tob Control* 2012 March;21(2):96-101.

(134) Law MR, Wald NJ. Environmental tobacco smoke and ischemic heart disease. *Prog Cardiovasc Dis* 2003 July;46(1):31-8.

(135) Steenland K, Thun M, Lally C, Heath C, Jr. Environmental tobacco smoke and coronary heart disease in the American Cancer Society CPS-II cohort. *Circulation* 1996 August 15;94(4):622-8.

(136) Price JF, Mowbray PI, Lee AJ, Rumley A, Lowe GD, Fowkes FG. Relationship between smoking and cardiovascular risk factors in the development of peripheral arterial disease and coronary artery disease: Edinburgh Artery Study. *Eur Heart J* 1999 March;20(5):344-53.

(137) Dawber TR, Kannel WB, REVOTSKIE N, STOKES J, III, KAGAN A, GORDON T. Some factors associated with the development of coronary heart disease: six years' follow-up experience in the Framingham study. *Am J Public Health Nations Health* 1959 October;49:1349-56.:1349-56.

(138) Seeman T, Merkin SS, Crimmins E, Koretz B, Charette S, Karlamangla A. Education, income and ethnic differences in cumulative biological risk profiles in a national sample of US adults: NHANES III (1988-1994). *Soc Sci Med* 2008 January;66(1):72-87.

Section A: Smoking and the risk of cardiovascular diseases in developed countries

Chapter 2: Description of the Emerging Risk Factors Collaboration

Summary

The aim of this Chapter is to present the dataset used to investigate the effect of smoking on CVD in developed populations. The Emerging Risk Factors Collaboration (ERFC) was set up with the aim of better characterizing associations between lipids and inflammatory markers with the risk of CVD. Compared to literature-based meta-analyses, ERFC collected individual participants' data, enabling meta-analyses of study specific estimates with consistent adjustment for confounders and tests for interaction across studies. By April 2011, ERFC had enrolled 135 prospective epidemiological studies with information on lipids and inflammatory markers as well as lifestyle, medical and demographic characteristics, and follow-up to first major cardiovascular event or main cause of death. After exclusion of studies conducted in developing countries or with insufficient information on smoking status, a subset of 929,335 individuals from 114 studies was selected for the analyses of cigarette smoking presented in **Chapter 4**. A subset of 20 studies has collected information on cigar use and a subset of 22 studies on pipe use presented in **Chapter 5**. During an average follow-up of 14.2 years; participants experienced 40,218 incident coronary artery disease outcomes, 17,445 strokes, 9,788 lung cancer deaths, and a total of 128,137 deaths from all causes. The ERFC will allow more detailed analyses of the risk of CVD in developed populations in relation to smoking cigarettes, pipes or cigars.

2.1 Background of ERFC

The aim of the Emerging Risk Factors Collaboration is to *"characterize more precisely and in greater detail than had previously been possible the shape and strength of associations of several lipids and inflammatory markers with incident CHD and other cardiovascular outcomes, under a wide range of circumstances"* (http://www.phpc.cam.ac.uk/ceu/research/erfc/) [2]. For this purpose, the ERFC established a central database with ongoing recruitment of suitable studies (**List 2.1**). By September 2012, the ERFC included more than 2 million participants from over 135 prospective studies with population-based samples. Subsets of participants had information on lipid and inflammatory markers, lifestyle and behavioural factors as well as major CHD and cerebrovascular events, and cause-specific mortality. The ERFC main objectives have been fulfilled with individual participants' meta-analyses on the relationship with CVD of triglycerides [3], major lipids and apolipoproteins [4], lipoprotein(a) [5], and C-reactive protein [6]. To help harvest the information made available by the collaboration, the ERFC has been enlarging its scope of analysis to lifestyle factors, such as tobacco, obesity [7], and diabetes [8,9]. Data requests were sent to all the ERFC investigators in December 2010 (**List 2.2**) requiring more detailed information on a range of lifestyle and other risk factors, including smoking amount, duration, pack-years, type, age starting and age stopping smoking. As a result, the ERFC represents one of the largest collections of information on smoking habits and other CVD risk factors, with follow-up of participants for major vascular and non-vascular events, similar in scale to the Cancer Prevention Studies set up in the U.S.A in 1960s and 1980, which included over 1 million participants [10]. Its uniqueness resides in the length of follow-up of prospective studies included in the Collaboration, in the depth of information gathered on smoking and other cardiovascular risk factors, such as lipids and inflammatory markers, and in the way individual prospective cohort studies enrolled random samples from the general population.

2.2 Methods

2.2.1 Inclusion of studies

As the initial aim of the ERFC was to investigate lipids and inflammatory markers, the criteria for inviting studies to join the ERFC were (1) data available from baseline measurements on at least one circulating lipid or inflammatory marker, (2) at least 1 year of follow-up, (3) participants not selected on the basis of having previous cardiovascular diseases; and (4) information on cause-specific mortality and/or major cardiovascular morbidity collected during follow-up [2]. To enable proper adjustment

for confounding factors and investigation of effect modifiers, additional data were sought on lifestyle, demographic and biochemical risk factors (**List 2.1**). Studies were identified through meta-analyses publications, literature search of databases, scanning of reference lists and correspondence with authors of relevant reports.

Data obtained from each participating study were checked for consistency and harmonized to a standard format by the coordinating centre. For biochemical factors expressed in different units of measurement, conversions to a standard unit were operated [11]. Information was stored on the different coding systems and on the assay methods used by individual studies to measure biochemical factors. For categorical variables, the number of categories was defined in order to maximize information whilst taking into account the need for harmonization of differential coding systems operated by individual studies.

2.2.2 Smoking information

Information relative to smoking was self-reported. Some studies provided smoking status already coded as “current smoker”, “ex-smoker” and “never smoker” whilst some studies provided answers to questions such as “when did you last smoke”, “have you smoked at least once over the past week / month / year preceding the interview?”. The data managing team of the ERFC, communicating with study investigators, created a harmonized variable of smoking status categorized as “current”, “ex-smoker” and “never smoker”.

Smoking amount was recorded for cigarettes, pipes and cigars separately when available. To enable comparisons across different types of smoking and also to be able to quantify total exposure to tobacco, self reported amounts of cigars, pipes and cigarettes were converted into a unique scale using the approximate amount of tobacco contained in each form of tobacco. One cigar was thought to contain on average 4 grams of tobacco and correspond to approximately 3 cigarettes. One pipe was converted into 1.6 cigarettes and 1 cigarillo was converted into 2 cigarettes (personal correspondence between data management teams of the ERFC and Prospective Study Collaboration [12]). Smoking pack-years was defined as the number of packs of cigarettes or cigarettes equivalent per day – 1 pack corresponding to 20 cigarettes equivalent per day – multiplied by the number of years smoked.

Regarding type of smoking, individuals reporting the use of small or large cigar, or of cigarillo, were coded as current cigar smokers. Some studies provided information

on only pipe, only cigar or only cigarette smoking. In order to maximize information available, individuals with missing information on cigar or on pipe were considered as non pipe or non cigar smokers in the main analyses. This is a reasonable assumption because currently pipe and cigar smokers represent below 5% of the population in most populations from developed countries [13,14]. Subsidiary analyses in which only studies providing information on all three types: pipe, cigar and cigarettes were also conducted to assess robustness of the results.

2.2.3 Follow-up and outcomes

For each individual, data were sought on each of the following outcomes and on their dates of occurrence: fatal and non-fatal MI events and/or fatal and non-fatal stroke events; and cause-specific mortality. In registering fatal outcomes, all contributing studies used coding from the International Classification of Diseases to at least 3 digits or study-specific classification systems, and ascertainment was based on death certificates (**Table 2.1**). Attribution of death referred to the primary cause provided or, in its absence, the underlying cause provided.

The study was approved by the Cambridgeshire Ethics Review Committee and analyzed independently from its funders.

2.3 Summary of dataset

2.3.1 Characteristics of studies

Out of the 114 studies, 5 were nested case-control studies, 10 were clinical trials and the rest were cohort studies (**Table 2.2**). Studies included spanned 4 decades, with the earliest cohort starting recruitment in 1960 and the most recent having initiated enrolment in 2001. Most individuals were recruited in the 1970s (32%), 1980s (37%) or 1990's (24%) while only 8% were enrolled in 1960s and less than 1% in 2000 and after. Average follow-up time was 14.2 years (inter-quartile range: 3.3 years to 30.4 years). Length of follow-up varied widely between studies, with the longest follow-up exceeding 35 years (KARELIA) and the shortest follow-up below 4 years (MOGERAUG3). Overall, 32% of individuals had less than 10 years of follow-up, 41% between 10 and 20 years, 21% between 20 and 30 years and 6% were followed up for 30 years or over. Average age at baseline was 49.5 years old (Standard Deviation: 12.3 years) and 51% of the data were men.

2.3.2 Baseline smoking characteristics

Smoking information was available in April 2011 for 117 studies. Out of these, 3 studies were excluded because they were set up in a developing country (China, Turkey and the Caribbean), leaving 114 studies, which included 929,335 individuals (**List 2.3**). In this data, 34% declared being current smokers (316,688 individuals), 23% ex-smokers (214,017 individuals) and 43% never smokers (398,630 individuals) (**Figure 2.1**).

Subsets of the 114 studies shared further information on smoking: 43 studies provided data on number of pack-years (114,662 individuals with non-missing and non-null values), 20 studies on cigar (131,816 individuals) and 22 studies on pipe smoking (132,060 individuals), 44 studies on starting age (119,907 individuals), and 54 studies on number of years since quitting smoking (80,318 individuals) (**Table 2.3**).

The distribution of amount smoked showed evidence of digit preferences for multiple of 5 which may be due to the fact that cigarettes are usually sold in packs of 20 (and therefore individuals are more likely to answer that they smoke 1 pack or half a pack of cigarettes per day, and it could also be that the question was asked in this way in the original study questionnaire). Indeed, 47% of smokers answered that they were currently smoking a multiple of 5 (5, 10, 15, 20 or 25) cigarettes per day.

Most individuals had started smoking during teenage-hood or early adulthood: more than 50% of current and ex-smokers had initiated smoking at age 18 years old or younger (IQR: 16-22) and 75% had begun before age 22 years old. Quitting age was more evenly spread, with a median age of 40 years old and an inter-quartile range of 17 years old (31 to 48 years old); and mean 42.9 years old (SD: 11.2). Those who had stopped were on average 39 years old (SD: 12) and had stopped for 9 years (IQR: 3 to 18) (**Table 2.2**). Mean pack-years smoked was 21 (SD: 6.8) in current smokers and 22 (SD: 19.4) in ex-smokers (**Table 2.4**).

Cigarette was the most popular form of smoking: 34% of individuals declared being current cigarette smokers at baseline (**Table 2.5**). Cigar was the second most common form of smoking: 8% of individuals who were asked a question on cigar smoking declared being current cigar smokers at baseline. The least common form of smoking was pipe: amongst those asked the question, only 5% reported being current pipe smokers at baseline. Pipe and cigar smokers reported on average a

lower number of pack-years, with medians respectively equal to 4 (inter-quartile range: 2-10) and 5 (2-9), versus 20 (10-32) for cigarette smoking.

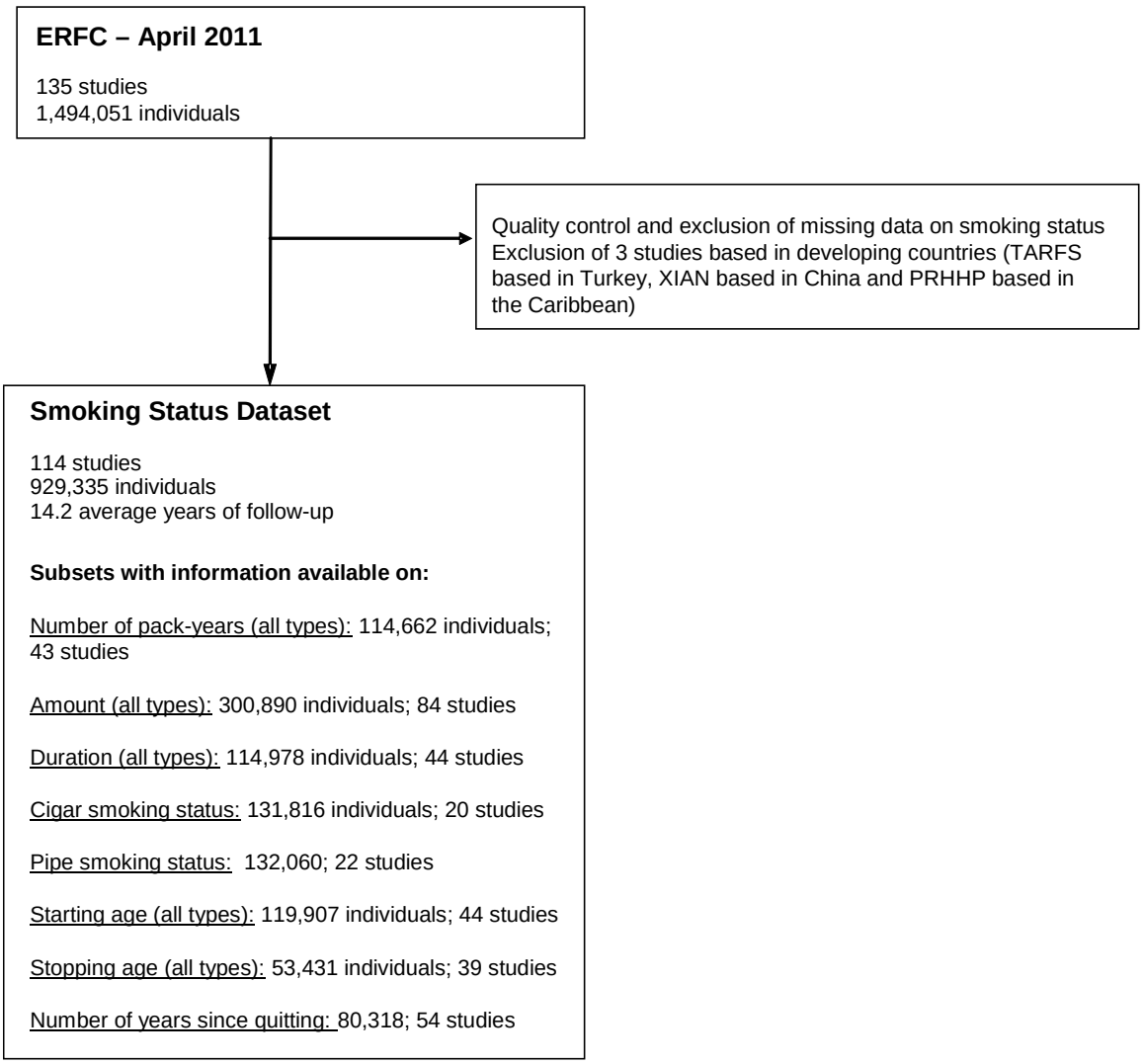
2.3.3 Outcomes

A total of 77 out of the 114 studies contributing to this analysis also involved medical records, autopsy findings and other supplementary sources to help classify deaths (**Table 2.6**). In total, 72 studies used standard definitions of myocardial infarction based on World Health Organization criteria. Out of 114 studies, 111 reported on ischemic stroke as a stroke subtype and 66 reported diagnosis of strokes on the basis of typical clinical features and brain imaging. During an average follow-up of 14.2 years; there were 40,218 incident coronary artery disease outcomes (16,390 non-fatal myocardial infarction and 23,828 coronary deaths); 17,445 strokes (5,023 ischaemic stroke and 2,014 haemorrhagic stroke and the rest was unclassified or information was not provided); 2,108 deaths from heart failures; 1,298 sudden deaths; 919 deaths from cardiac dysrhytmia; 704 deaths from pulmonary embolism; 1331 deaths from aortic aneurysm and 385 deaths from peripheral vascular disease (**Table 2.7**). There were also 9,788 lung cancer deaths, and, in total, 128,137 deaths from all causes.

2.4 Use of datasets in later Chapters

The dataset described above is used in **Chapter 3**, **4** and **5**. In **Chapter 3**, the dataset including all studies with smoking status (114 studies and 929,335 individuals) is further restricted to studies with information on a specific correlate (blood pressure, lipids, inflammatory markers). To investigate shapes of relationship between smoking and other factors, subsets with information on pack-years (43 studies, 114,662 individuals) and with information on years since quitting smoking (54 studies, 80,318 individuals) are restricted to further subsets with information on specific correlates in **Chapter 3**. **Chapter 4** uses studies which provide information on smoking status and on the following subsets: amount, duration, starting age, stopping age and number of years since quitting. **Chapter 5** uses studies which provide information on cigarette, cigar and pipe smoking status. A more detailed description of the datasets can be found in each Chapter.

Figure 2.1: Flow chart of study selection



ERFC: Emerging Risk Factor Collaboration

Table 2.1: Definition of endpoints recorded in The ERFC using the International Classification of Diseases version 10

Endpoint	ICD-10 codes
All cardiovascular*	G45, I01, I03-I82, I87, I95-I99, F01, Q20-2Q28, R96
Coronary heart disease (CHD)*	I21-I25
Myocardial infarction	I21, I22
All cerebrovascular*	F01, I60-I69
Ischemic stroke*	I63
Hemorrhagic stroke*	I61
Subarachnoid stroke*	I60
Unclassified stroke†*	I64
Other vascular deaths	Remainder of cardiovascular disease (fatal)
Cardiac dysrhythmia	I47-I49
Hypertensive disease	I10-I15
Pulmonary embolism	I26
Ill-defined descriptions and complications of the dearth disease	I51
Sudden death	R96
Aortic aneurysm	I71
Heart failure	I50
Peripheral vascular disease	I73-I74, I77-I78
Other	Remainder of vascular
All cancer	C00-C97, D00-D48
Oral	C00-C14
Colorectum	C18-C21
Oesophagus	C15
Stomach	C16
Liver	C22
Pancreas	C25
Lung	C34
Prostate	C61
Ovary	C56
Bladder	C67
Hematological	C81-C96
Endocrine & nervous	C69-C75
Melanoma	C43
Connective tissue	C40-C42, C45-C49
Breast (female)	C50
Other/unspecified	Remainder of cancer/ unspecified to ERFC
All non-cancer, non-vascular	A00-A99, B00-B99, D50-D99,E00-E99, F00, F02-F99, G00-G44, G46-G99, H00-H99, I00, I02, I83-I86, I88-I89, J00-J99, K00-K99, L00-L99, M00-M99, N00-N99, O00-O99, P00-P99, Q00-Q18, Q30-Q99, S00-S99, T00-T99, U04, V00-V99, W00-W99, X00-X99, Y00-Y99, Z00-Z99
All external cause	S00-S99, T00-T98, U04, V01-V99, W00-W99, X00-X99, Y00-Y98, Z00-Z99
Falls	W00-W19
Intentional self-harm	X60-X84
Infections	A00-A99, B00-B14, B20-B99
Diabetes mellitus	E10-E14
Mental disorders	F04-F99
Alzheimer's disease and related conditions	F00, F02, F03, G30-G32
Liver disease	B15-B19, K70-K77
Respiratory system disease	J00-J99
Pneumonia	J12-J18
COPD and related conditions	J40-J47
Digestive system disease (except liver)	K00-K69, K78-K93
Renal disease	N00-N19
Other/unspecified	Remainder of non-cancer, non-vascular/ unspecified to ERFC
Deaths of unknown cause or ill-defined cause	R00-R96, R97-R99 and non-vascular deaths defined according to study-specific read-codes for mortality, and not standard ICD codes.
All-cause mortality	A00-Y89

* includes both fatal and non-fatal events. † Unclassified stroke was defined by the ICD codes stated, or as strokes not specified as ischaemic or hemorrhagic by study-specific codes. NB: Corresponding ICD-6, 7, 8 or 9 codes were used for studies that recorded outcomes using earlier ICD versions

Table 2.2: Characteristics of individual studies with complete information on at least smoking status, age and sex

						Smoking baseline characteristics					
Study name / study design	Country	Median follow-up (5th & 95th percentiles)	No with smoking status	Age at survey (yrs) mean (sd)	Male sex N (%)	Current all types smoking N (%)	Current pipe or cigar smokers, %	Smoking pack years[$], Median (IQR)	Smoking age start, median (IQR)	Smoking age stopped, mean (sd)	Smoking number years stopped, median (IQR)
Nested case-control studies											
EPICNOR	UK	7.5 (3.4 to 9.3)	3334	65 (8)	2110 (63)	352 (11)	-	-	-	-	-
FIA	Sweden	4.0 (0.3 to 9.5)	2684	54 (8)	2161 (81)	633 (24)	-	-	-	-	-
FLETCHER	New Zealand	5.7 (4.9 to 6.4)	706	56 (15)	537 (76)	125 (18)	-	18 (10 to 29)	19 (16 to 22)	-	-
GLOSTRUP	Denmark	4.5 (0.5 to 12.5)	390	50 (9)	286 (73)	237 (61)	-	-	-	-	-
WHIHABPS	USA	6.8 (1.2 to 9.3)	1560	68 (6)	0 (0)	91 (6)	-	-	-	-	-
SUBTOTAL		*6.2 (1.0 to 9.3)*	*8674*	*61 (11)*	*5094 (59)*	*1438 (17)*	-	-	-	-	-
Clinical trials											
AFTCAPS	USA	5.1 (4.1 to 6.6)	652	57 (7)	540 (83)	652 (100)	-	-	-	-	-
ALLHAT	USA/Canada/ Puerto Rico/ US Virgin Islands	4.4 (0.4 to 6.7)	28684	66 (8)	14031 (49)	7037 (25)	-	-	-	-	-
LEADER	UK	4.2 (0.6 to 6.1)	471	66 (9)	471 (100)	382 (81)	9	-	-	-	10 (2 to 21)
MRFIT	USA	6.9 (4.4 to 7.8)	12833	47 (6)	12833 (100)	8164 (64)	-	-	-	-	-
PREVEND	Netherlands	7.6 (4.7 to 8.2)	4612	49 (11)	2130 (46)	2381 (52)	-	-	-	-	-
PROSPER	Scotland/Ireland/ Netherland	3.2 (1.1 to 3.8)	3253	75 (3)	1351 (42)	1083 (33)	-	-	-	-	-
TPT	UK	7.5 (2.8 to 10.5)	7190	55 (7)	7190 (100)	7190 (100)	-	-	-	-	-
USPHS2	USA	10.9 (4.9 to 11.5)	10707	64 (8)	10707 (100)	507 (5)	-	-	-	-	-
WHS	USA	10.2 (8.4 to 10.8)	27456	55 (7)	0 (0)	3185 (12)	-	-	-	-	-
WOSCOPS	UK	4.8 (2.2 to 5.8)	2717	55 (5)	2717 (100)	2717 (100)	-	-	-	-	-
SUBTOTAL		*7.2 (1.8 to 11.0)*	*98575*	*59 (10)*	*51970 (53)*	*33298 (34)*	-	-	-	-	-
Cohort studies											
ARIC	USA	14.0 (5.0 to 15.7)	14584	54 (6)	6294 (43)	3745 (26)	-	24 (11 to 38)	19 (16 to 22)	40 (11)	14 (6 to 23)
ATENA	Italy	6.7 (5.2 to 8.1)	4740	50 (7)	0 (0)	1900 (40)	-	16 (8 to 27)	18 (16 to 24)	41 (10)	8 (3 to 12)
ATS_SAR	Italy	8.7 (5.7 to 9.1)	3778	46 (8)	1963 (52)	1201 (32)	-	19 (10 to 30)	20 (16 to 24)	43 (8)	7 (3 to 10)
ATTICA	Greek	5.0 (0.0 to 5.0)	2113	52 (10)	1076 (51)	897 (42)	-	25 (11 to 40)	22 (19 to 29)	-	-
AUSDIAB	Australia	5.0 (4.9 to 8.5)	8260	53 (12)	3619 (44)	1387 (17)	2	-	-	-	16 (9 to 25)
BHS	Australia	26.6 (7.2 to 33.2)	5931	45 (16)	2795 (47)	1857 (31)	5	-	-	-	-
BRHS	UK	24.5 (4.7 to 25.4)	6795	50 (6)	6795 (100)	2797 (41)	-	-	-	-	-
BRUN	Italy	20.2 (3.9 to 20.5)	817	58 (11)	398 (49)	198 (24)	-	-	-	-	-
BUPA	UK	23.7 (11.1 to 26.8)	14517	47 (8)	14517 (100)	8232 (57)	-	-	-	-	6 (1 to 10)
BWHHS	UK	7.3 (3.1 to 8.4)	2789	68 (5)	0 (0)	321 (12)	0	17 (8 to 31)	18 (17 to 22)	46 (13)	22 (13 to 33)
CAPS	UK	13.0 (4.0 to 13.0)	2129	52 (5)	2129 (100)	1180 (55)	27	-	-	-	4 (4 to 5)
CASTEL	Italy	11.1 (2.0 to 14.0)	330	72 (4)	245 (74)	330 (100)	-	20 (10 to 30)	21 (17 to 31)		
CHA	USA	32.0 (11.6 to 35.6)	34250	41 (13)	19894 (58)	14243 (42)	-	-	-	-	-
CHARL	USA	24.0 (3.4 to 40.0)	2028	50 (11)	934 (46)	1114 (55)	-	-	-	-	-
CHS1	USA	12.1 (2.0 to 12.9)	3778	72 (5)	1437 (38)	452 (12)	-	-	18 (16 to 21)	51 (14)	20 (10 to 31)
CHS2	USA	9.1 (1.9 to 9.5)	462	72 (5)	173 (37)	76 (16)	-	-	-	-	-
COPEN	Denmark	13.2 (2.6 to 14.9)	6166	57 (14)	2525 (41)	4000 (65)	0	-	-	-	-
DISCO	Italy	5.5 (5.5 to 9.5)	1825	50 (11)	813 (45)	512 (28)		19 (8 to 34)	18 (15 to 22)	47 (12)	7 (3 to 10)
DRECE	Spain	16.4 (15.5 to 16.6)	2748	41 (11)	1338 (49)	1098 (40)		5 (2 to 7)	22 (15 to 30)	37 (11)	7 (3 to 14)
DUBBO	Australia	14.1 (2.0 to 15.0)	1523	68 (7)	537 (35)	323 (21)	0				15 (6 to 30)
EAS	Scotland	15.2 (2.8 to 15.8)	1036	64 (6)	515 (50)	243 (23)	5	21 (10 to 33)	20 (17 to 25)		
EMOFRI	Italy	6.8 (6.5 to 7.2)	360	55 (6)	176 (49)	92 (26)		19 (9 to 32)	20 (18 to 23)	41 (10)	15 (6 to 20)
ESTHER	Germany	5.0 (0.0 to 6.1)	1286	59 (6)	622 (48)	1286 (100)	-	-	-	-	-
FINE_FIN	Finland	5.3 (0.5 to 10.0)	40	75 (4)	40 (100)	40 (100)	-	-	-	-	-
FINE_IT	Italy	9.9 (1.9 to 21.4)	459	72 (4)	459 (100)	123 (27)		-	-	-	10 (5 to 10)
FINRISK92	Finland	16.9 (7.9 to 16.9)	5737	44 (11)	2647 (46)	1761 (31)	0	-	-	-	15 (9 to 21)
FINRISK97	Finland	11.8 (6.7 to 11.9)	6306	49 (12)	3059 (49)	1879 (30)	0	-	-	-	15 (9 to 24)

FRAMOFF	USA	5.0 (2.0 to 6.9)	335	57 (9)	147 (44)	335 (100)	-	-	-	-	-
FUNAGATA	Japan	7.3 (5.3 to 10.2)	1125	53 (12)	504 (45)	331 (29)	-	-	-	-	-
GOH	Israel	35.0 (12.9 to 36.0)	3874	43 (8)	1957 (51)	1693 (44)	-	-	-	-	-
GOTO13	Sweden	21.0 (4.0 to 30.5)	404	54 (0)	404 (100)	404 (100)	-	-	-	-	-
GOTO33	Sweden	12.8 (5.7 to 13.1)	720	51 (0)	720 (100)	264 (37)	-	-	-	-	-
GOTO43	Sweden	11.0 (7.9 to 11.7)	776	50 (0)	776 (100)	240 (31)	-	-	-	-	-
GOTOW	Sweden	32.2 (10.6 to 32.7)	1425	47 (6)	0 (0)	582 (41)	-	10 (5 to 16)	21 (18 to 26)	38 (9)	10 (3 to 17)
GREPCO	Italy	7.9 (7.7 to 8.4)	779	44 (8)	0 (0)	319 (41)	-	8 (4 to 17)	25 (20 to 30.5)	-	-
GRIPS	Germany	9.8 (4.9 to 10.0)	5320	48 (5)	5320 (100)	2179 (41)	-	26 (19 to 35)	18 (17 to 20	-	-
GUBBIO	Italy	8.4 (5.8 to 9.4)	3408	55 (13)	1515 (44)	1161 (34)	-	-	-	-	-
HBS	Finland	20.5 (5.0 to 20.5)	1302	60 (4)	1302 (100)	227 (17)	-	-	-	-	-
HELSINAG	Finland	9.3 (2.2 to 11.0)	385	79 (4)	100 (26)	35 (9)	-	-	-	-	-
HISAYAMA	Japan	14.0 (3.2 to 14.0)	2576	59 (12)	1088 (42)	349 (14)	-	32 (18 to 47)	20 (20 to 23)	57 (11)	1 (0 to 1)
HONOL	USA	6.3 (1.5 to 7.6)	2436	78 (4)	2436 (100)	184 (8)	-	-	-	-	-
HOORN	Netherlands	8.8 (3.2 to 9.9)	2228	61 (7)	983 (44)	706 (32)	-	-	-	-	-
HPFS1	USA	20.2 (6.2 to 21.9)	45857	54 (10)	45857 (100)	4562 (10)	-	-	-	-	-
IKNS	Japan	11.1 (5.1 to 18.6)	8038	58 (10)	3299 (41)	1911 (24)	-	32 (19 to 45)	20 (20 to 21)	-	-
ISRAEL	Israel	23.3 (7.9 to 23.9)	7698	49 (7)	7698 (100)	4111 (53)	5	-	-	-	-
KARELIA	Finland	36.7 (6.5 to 36.9)	9735	41 (10)	4638 (48)	3271 (34)	0	-	-	-	1 (1 to 1)
KIHD	Finland	19.0 (1.8 to 23.9)	632	52 (6)	632 (100)	632 (100)		-	-	-	
LASA	Netherlands	9.9 (1.5 to 10.5)	1236	69 (8)	549 (44)	285 (23)	21	-	-	-	20 (10 to 31)
MALMO	Sweden	18.2 (8.2 to 22.6)	31675	45 (7)	21913 (69)	14413 (46)	-	-	-	-	-
MATISS83	Italy	18.7 (6.8 to 19.5)	2552	51 (10)	1198 (47)	759 (30)	-	22 (11 to 37)	18 (15 to 20)	44 (12)	8 (3 to 15)
MATISS87	Italy	15.6 (7.2 to 16.2)	2019	52 (9)	895 (44)	443 (22)	-	23 (12 to 36)	18 (15 to 20)	42 (12)	10 (3 to 18)
MATISS93	Italy	8.3 (7.1 to 9.3)	1214	49 (9)	587 (48)	325 (27)	-	22 (10 to 36)	18 (15 to 20)	41 (11)	7 (2 to 13)
MCVDRFP	Netherlands	16.7 (10.5 to 18.9)	23522	42 (10)	10946 (47)	10329 (44)	4	14 (6 to 25)	17 (15 to 20)	33 (10)	8 (3 to 15)
MESA	USA	4.8 (2.0 to 5.2)	3576	62 (10)	2092 (59)	1017 (28)	3	17 (6 to 34)	18 (15 to 20)	42 (14)	21 (11 to 31)
MICOL	Italy	5.9 (4.5 to 7.1)	18858	51 (10)	10650 (56)	6164 (33)	-	18 (9 to 30)	21 (18 to 28)	-	-
MOGERAUG1	Germany	13.0 (3.6 to 13.4)	871	54 (6)	871 (100)	294 (34)	-	-	-	-	-
MOGERAUG2	Germany	7.9 (2.3 to 8.4)	3962	53 (12)	1949 (49)	969 (24)	-	-	-	-	-
MOGERAUG3	Germany	3.0 (1.8 to 3.6)	3373	55 (10)	1664 (49)	747 (22)	-	-	-	-	-
MONFRI86	Italy	16.7 (7.7 to 16.9)	1407	49 (9)	690 (49)	447 (32)	-	17 (8 to 26)	19 (16 to 21)	40 (11)	7 (2 to 14)
MONFRI89	Italy	13.6 (7.5 to 13.7)	1341	49 (8)	666 (50)	380 (28)	-	17 (8 to 29)	18 (16 to 20)	39 (10)	8 (3 to 13)
MONFRI94	Italy	8.5 (8.0 to 8.8)	1293	49 (8)	629 (49)	349 (27)	-	15 (7 to 27)	19 (16 to 20)	38 (10)	10 (5 to 16)
MONICA	Italy	6.5 (2.1 to 10.5)	3639	49 (9)	1825 (50)	1242 (34)	-	18 (7 to 32)	18 (15 to 22)	47 (10)	7 (3 to 10)
MORGEN	Netherlands	10.8 (8.3 to 13.0)	18246	46 (9)	8307 (46)	6826 (37)	6	15 (7 to 26)	17 (15 to 18)	34 (10)	12 (5 to 19)
MOSWEGOT	Sweden	13.9 (7.6 to 19.6)	4158	47 (11)	1973 (47)	1222 (29)	-	-	-	-	-
MRCOLD	UK	8.7 (1.2 to 11.7)	10137	80 (4)	3825 (38)	1222 (12)	-	12 (5 to 27)	-	50 (16)	28 (16 to 43)
NCS1	Norway	16.1 (14.4 to 16.7)	23996	42 (4)	11815 (49)	10467 (44)	20	-	-	40 (4)	5 (3 to 5)
NCS2	Norway	17.3 (15.6 to 17.8)	12743	42 (4)	6491 (51)	5070 (40)	8	-	-	40 (5)	5 (3 to 5)
NCS3	Norway	18.2 (12.2 to 18.8)	9658	42 (4)	4994 (52)	5485 (57)	5	-	-	40 (5)	3 (3 to 5)
NFR	Italy	10.2 (6.1 to 11.2)	3075	55 (5)	3075 (100)	1564 (51)	-	25 (14 to 38)	20 (17 to 24)	51 (6)	7 (3 to 10)
NHANESI	USA	19.0 (3.7 to 21.0)	6629	51 (15)	2456 (37)	1402 (21)	5	22 (11 to 39)	18 (16 to 21)	38 (14)	8 (4 to 15)
NHANESIII	USA	14.3 (3.5 to 17.7)	7357	51 (17)	4339 (59)	3979 (54)	3	8 (2 to 19)	20 (17 to 29)	41 (16)	13 (5 to 23)
NHS1	USA	28.6 (12.0 to 30.3)	117887	43 (7)	0 (0)	38884 (33)	-	-	-	-	-
NPHSI	UK	14.6 (4.4 to 18.6)	1392	52 (7)	1392 (100)	704 (51)	54	-	-	-	9 (4 to 17)
NPHSII	UK	8.3 (3.5 to 10.4)	1941	57 (3)	1941 (100)	1101 (57)	30	-	-	-	-
NSHS	Canada	9.7 (4.0 to 10.0)	1168	53 (15)	614 (53)	411 (35)	-	24 (11 to 37)	17 (15 to 20)	51 (13)	1 (1 to 1)
OB43	Italy	7.5 (5.1 to 9.1)	3531	47 (8)	1707 (48)	1264 (36)	-	16 (7 to 30)	21 (17 to 27)	43 (9)	7 (3 to 10)
OSAKA	Japan	10.2 (3.9 to 18.8)	12380	52 (10)	8415 (68)	4795 (39)	-	29 (17 to 42)	20 (20 to 20)	-	-
OSLO	Norway	29.5 (10.7 to 30.5)	16832	44 (6)	16832 (100)	9600 (57)	40	-	-	41 (6)	5 (3 to 5)
OYABE	Japan	10.4 (4.2 to 10.6)	973	57 (12)	873 (90)	972 (100)	-	-	-	-	-
PARIS1	France	22.9 (7.0 to 26.1)	5959	47 (2)	5959 (100)	4771 (80)	12	-	-	-	6 (2 to 15)
PRIME	France / NI	5.2 (5.0 to 7.3)	9486	55 (3)	9486 (100)	2544 (27)	5	-	-	-	14 (6 to 20)
PROCAM	Germany	9.8 (3.8 to 18.9)	12790	44 (10)	10333 (81)	7878 (62)	5	18 (10 to 28)	18 (17 to 21)	35 (9)	10 (4 to 15)
QUEBEC	Canada	26.4 (4.2 to 26.9)	1798	46 (7)	1798 (100)	1429 (79)	22	37 (25 to 50)	16 (15 to 19)	38 (10)	6 (2 to 12)
RANCHO	USA	14.2 (2.0 to 18.1)	1819	68 (11)	753 (41)	250 (14)	-	20 (8 to 40)	22 (18 to 30)	48 (13)	19 (11 to 28)
REYK	Iceland	24.7 (6.3 to 37.1)	16760	52 (9)	8028 (48)	7944 (47)	20	20 (12 to 31)	22 (17 to 22)	38 (10)	14 (6 to 22)

RF2	Italy	13.7 (11.3 to 14.1)	5392	44 (9)	2536 (47)	1979 (37)	-	15 (6 to 28)	20 (16 to 25)	41 (9)	7 (3 to 10)
ROTT	Netherlands	12.0 (3.3 to 14.3)	2279	68 (9)	550 (24)	892 (39)	-	-	-	-	-
SHHEC	UK	10.0 (6.3 to 10.0)	13402	49 (8)	6533 (49)	5962 (44)	10	27 (18 to 38)	17 (15 to 19)	38 (11)	11 (4 to 19)
SHS	USA	12.4 (2.0 to 14.3)	4132	56 (8)	1613 (39)	1371 (33)	-	-	-	-	-
SPEED	UK	16.7 (3.3 to 18.2)	2114	55 (4)	2114 (100)	1011 (48)	18	-	-	50 (5)	10 (7 to 10)
TOYAMA	Japan	12.7 (7.8 to 12.8)	4524	46 (7)	2908 (64)	1752 (39)	-	-	-	-	-
TROMSØ	Norway	18.9 (11.4 to 19.3)	13665	39 (10)	6646 (49)	6610 (48)		8 (4 to 15)	18 (16 to 22)	-	10 (6 to 15)
ULSAM	Sweden	26.5 (6.0 to 37.6)	1629	50 (2)	1629 (100)	1167 (72)	71	13 (7 to 23)	20 (18 to 21)	38 (9)	8 (3 to 10)
VHMPP	Austria	13.1 (2.2 to 16.7)	120612	48 (14)	55101 (46)	22370 (19)	-	-	-	-	-
VITA	Italy	3.3 (1.7 to 5.3)	2353	50 (8)	1205 (51)	2353 (100)	0	-	-	-	25 (14 to 36)
WHITEI	UK	8.2 (2.0 to 8.4)	3993	76 (5)	3993 (100)	554 (14)	-	-	-	-	-
WHITEII	UK	12.4 (4.9 to 14.1)	10174	45 (6)	6787 (67)	2624 (26)	16	21 (12 to 31)	18 (16 to 20)	-	-
ZARAGOZA	Spain	5.1 (4.1 to 5.1)	2601	60 (12)	1101 (42)	478 (18)	-	-	-	-	2 (1 to 3)
ZUTE	Netherlands	7.0 (0.9 to 10.1)	123	75 (4)	123 (100)	99 (80)	-	30 (16 to 44)	17 (15 to 19)	73 (4)	-
SUBTOTAL		*15.7 (4.1 to 31.3)*	*822086*	*48 (12)*	*419745 (51)*	*281952 (34)*	-	-	-	-	-
TOTAL		**14.2 (3.3 to 30.4)**	**929,335**	**50 (12)**	**476809 (51)**	**316688 (34)**	**10**	**18 (8 to 31)**	**18 (16 to 22)**	**39 (12)**	**9 (3 to 18)**

No: Number of individuals; -: Not available; $: smoking pack-years in current and ex-smokers (excluding never smokers); IQR: Inter-quartile range; SD: standard deviation from the mean. Means and SD were computed as weighted average of study-specific means and SD.

Table 2.3: Studies providing information on smoking type

Cohort abbreviation	Type of smoking			
	Cigarettes	Pipe or cigars	Pipes only	Cigars only
AFTCAPS	o	•	•	•
ALLHAT	•	•	•	•
ARIC	o	•	•	•
ATENA	o	•	•	•
ATS_SAR	o	•	•	•
ATTICA	o	•	•	•
AUSDIAB	o	o	o	o
BHS	o	o	•	•
BRHS	o	•	•	•
BRUN	•	•	•	•
BUPA	o	o	o	o
BWHHS	o	•	•	•
CAPS	o	o	o	o
CASTEL	o	•	•	•
CHA	o	•	•	•
CHARL	•	•	•	•
CHS1	o	•	•	•
CHS2	o	•	•	•
COPEN	•	•	•	•
DISCO	o	•	•	•
DRECE	o	•	•	•
DUBBO	•	•	•	•
EAS	o	o	•	•
EMOFRI	o	•	•	•
EPICNOR	o	•	•	•
ESTHER	o	•	•	•
FIA	•	•	•	•
FINE_FIN	o	•	•	•
FINE_IT	•	•	•	•
FINRISK92	•	•	•	•
FINRISK97	•	•	•	•
FLETCHER	o	•	•	•
FRAMOFF	o	•	•	•
FUNAGATA	•	•	•	•
GLOSTRUP	•	•	•	•
GOH	•	•	•	•
GOTO13	o	o	•	•
GOTO33	•	•	•	•
GOTO43	•	•	•	•
GOTOW	o	•	•	•
GREPCO	o	•	•	•
GRIPS	o	•	•	•
GUBBIO	o	•	•	•
HBS	•	•	•	•
HELSINAG	•	•	•	•
HISAYAMA	o	•	•	•
HONOL	•	•	•	•
HOORN	o	•	•	•
HPFS1	o	•	•	•
IKNS	o	•	•	•
ISRAEL	o	o	•	•
KARELIA	•	•	•	•
KIHD	•	•	•	•
LASA	•	o	o	o
LEADER	o	o	o	o
MALMO	o	•	•	•
MATISS83	o	•	•	•
MATISS87	o	•	•	•
MATISS93	o	•	•	•
MCVDRFP	o	o	o	o
MESA	o	o	o	o
MICOL	o	•	•	•
MOGERAUG1	o	•	•	•
MOGERAUG2	o	•	•	•
MOGERAUG3	o	•	•	•
MONFRI86	o	•	•	•
MONFRI89	o	•	•	•
MONFRI94	o	•	•	•
MONICA	o	•	•	•
MORGEN	o	o	•	•
MOSWEGOT	o	•	•	•
MRCOLD	o	•	•	•
MRFIT	o	•	•	•
NCS1	o	o	o	o
NCS2	o	o	o	o
NCS3	o	o	o	o
NFR	o	•	•	•
NHANESI	o	o	o	o
NHANESIII	o	o	o	o
NHS1	o	•	•	•
NPHSI	o	o	o	o
NPHSII	•	o	o	•
NSHS	o	•	•	•
OB43	o	•	•	•
OSAKA	o	•	•	•
OSLO	o	o	o	o
OYABE	o	•	•	•
PARIS1	o	o	o	o
PREVEND	o	•	•	•
PRIME	o	o	o	o
PROCAM	o	o	•	•
PROSPER	•	•	•	•
QUEBEC	o	o	o	o
RANCHO	o	•	•	•
REYK	o	o	o	o
RF2	o	•	•	•

ROTT	•	•	•	•
SHHEC	º	º	º	º
SHS	º	•	•	•
SPEED	º	º	•	•
TOYAMA	•	•	•	•
TPT	•	•	•	•
TROMSØ	º	º	º	•
ULSAM	º	º	º	º
USPHS2	•	•	•	•
VHMPP	•	•	•	•
VITA	•	•	•	•
WHIHABPS	•	•	•	•
WHITEI	º	º	•	•
WHITEII	º	º	•	•
WHS	º	•	•	•
WOSCOPS	•	•	•	•
ZARAGOZA	•	•	•	•
ZUTE	º	º	•	•

•: Information not collected by study investigators or collected by study investigators but not shared with THE ERFC. º: Information provided

Table 2.4: Smoking intensity and demographics at baseline

			Smoking status	
Characteristics	***No of studies***	***No of individuals***	**Mean (SD) in current smokers**	**Mean (SD) in ex-smokers**
All types of smoking combined				
Smoking pack-years	*43*	*114,662*	21 (6.8)	22 (19.4)
Duration of smoking (yrs)	*44*	*114,978*	30.2 (10)	21.8 (10.8)
Amount smoked (cpd)	*84*	*300,890*	16.5 (14.2)	17.6 (11.6)
Age at smoking initiation (yrs)	*44*	*119,907*	21 (6.8)	19.7 (5.4)
Age at smoking cessation (yrs)	*39*	*53,431*	-	42.9 (11.2)
Time since quitting smoking (yrs)	*54*	*80,318*	-	11.2 (9.5)

No: Number of individuals excluding null values (for example never smokers will have null value for amount smoked); cpd: cigarettes equivalent per day. No: Number of individuals; -: Not available; $: smoking pack-years in current and ex-smokers (excluding never smokers); IQR: Inter-quartile range; SD: standard deviation from the mean. Means and SD were computed as weighted average of study-specific means and SD.

Table 2.5: Smoking status at baseline

	No studies	*No subjects*	Smoking status, %			
			Current smoker	Ex-smoker	Never smoker	Other *
All types of smoking combined	*114*	*929,335*	34%	23%	43%	0%
Cigarette smoking	*83*	*689,104*	34%	25%	37%	4%
Cigar or pipe smoking	*33*	*178,566*	11%	4%	45%	40%
Cigar smoking only	*20*	*131,816*	8%	3%	43%	46%
Pipe smoking only	*22*	*132,060*	5%	3%	43%	49%

No: Number. *: Other includes individuals who were classified as “non current” (meaning ex or never) or “users” (meaning ex or current) without the possibility to classify them further as either current, ex or never. A total of 19 studies, including 110,456 individuals who were current or ex smokers, had information on the three smoking types. Some studies only provided information for “cigar or pipe smoking” without further precision regarding the type of smoking.

Table 2.6: Endpoints characteristics

Study		Definition of incident endpoint						Classification of incident endpoints						
	Death	Non-fatal MI			Non-fatal stroke			MI			Stroke			
		Clinical feature	ECG	Cardiac enzymes	Clinical features	CT/MRI imaging		Definite	Probable	Silent	Ischemic	Hemorrhagic	SAH	Unclassified
AFTCAPS	**	√	√	√	√	√		√	-	√NC	√	√	√	√
ALLHAT	**	√	√	√	√	√		√	o	o	√ NC	√ NC	√ NC	√
ARIC	**	√	√	√	√	√		√	√	√	√	√	√	√
ATENA†	**	√	√	√	√	√		√	√	√	√	√	√	√
ATS_SAR	*	NA	NA	NA	NA	NA		√	√ NC	o	√	√	√	√
ATTICA	**	NA	NA	NA	NA	NA		NS	NS	NS	o	o	o	o
AUSDIAB	*	NS	NS	NS	NS	NS		NS	NS	NS	√	√	√	√
BHS	*	NA	NA	NA	NA	NA		√	o	o	√	√	√	√
BRHS	*	√	√	√	NS	NS		√	o	o	√	√	√	√
BRUN	**	√	√	√	√	√		√	o	o	√	√	o	o
BUPA	*	NA	NA	NA	NA	NA		√	o	o	√	√	√	√
BWHHS	**	√	√	√	√	√		√	o	o	√	√	√	√
CAPS	**	√	√	o	√	√		√	√	o	√	√	√	√
CASTEL	**	NA	NA	NA	NA	NA		√	o	o	√ NC	√ NC	√ NC	√
CHA	*	NA	NA	NA	NA	NA		NS	NS	NS	NS	NS	NS	NS
CHARL	**	√	√	o	√	o		√	√	o	√	√	√	√
CHS 1	**	√	√	√	√	√		√	√ NC	√	√	√	o	√
CHS 2	**	√	√	√	√	√		√	√ NC	√	√	√	o	√
COPEN	**	√	√	√	√	√		√	o	o	√	√	√	√
DISCO‡	*	NA	NA	NA	NA	NA		√	√ NC	o	√	√	√	√
DRECE	*	NA	NA	NA	NA	NA		√	o	o	NS	NS	NS	√
DUBBO	**	√	√	√	√	√		√	NS	o	√	√	√	√
EAS	**	√	√	√	√	√		√	√	√	√	√	√	√
EMOFRI†	**	√	√	√	√	√		√	√	√	√	√	√	√
EPIC-	*	√	√	√	√	√		√	o	o	√ NC	√ NC	√ NC	√
ESTHER	*	NS	NS	NS	NS	NS		NS	NS	NS	NS	NS	NS	NS
FIA	**	√	√	√	NA	NA		√	o	o	√ NC	√ NC	√ NC	√
FINE_FIN	**	√	√	√	√	NA		√	NS	√	√	√	√	√
FINE_IT	**	√	√	√	√	NA		√	NS	√	√	√	√	√
FINRISK9	**	√	√	√	√	√		√	o	o	√	√	√	√
FINRISK9	**	√	√	√	√	√		√	o	o	√	√	√	√
FLETCHE	*	√	√	√	√	√		√	o	o	√ NC	√ NC	√ NC	√
FRAMOF	**	√	√	√	√	√		√	o	√	√	√	√	√
FUNAGA	NS	NS	NS	NS	NS	NS		NS	NS	NS	NS	NS	NS	NS
GLOSTR	**	√	√	√	NA	NA		o	o	o	√ NC	√ NC	√ NC	√
GOH	**	NA	NA	NA	NA	NA		√	√ NC	o	√	√	√	√
GOTO13	**	√	√	√	√	√		√	o	o	√ NC	√ NC	√ NC	√
GOTO33	**	√	√	√	√	√		√	o	o	√ NC	√ NC	√ NC	√
GOTO43	**	√	√	√	√	√		√	o	o	√	√	√	√
GOTOW	*	√	√	√	√	√		√	o	o	NS	NS	NS	√
GREPCO‡	*	NA	NA	NA	NA	NA		√	√ NC	o	√	√	√	√
GRIPS	**	√	√	√	√	√		√	√	o	√	√	√	√
GUBBIO‡	*	NA	NA	NA	NA	NA		√	√ NC	o	√	√	√	√
HBS	**	NA	NA	NA	NA	NA		√	o	NS	√ NC	√ NC	√ NC	√
HELSINA	**	NA	NA	NA	NA	NA		√	o	o	√	√	√	√
HISAYAM	**	√	√	√	√	√		√	√	√	√	√	√	√
HONOL	**	√	√	√	√	√		√	√ NC	√	√	√	√	√
HOORN	*	√	√	√	√	o		√	o	o	√	√	√	√
HPFS	**	NA	NA	NA	NA	NA		√	√ NC	√	√	√	√	√
IKNS	**	√	√	√	√	√		√	√	o	√	√	√	√
ISRAEL	**	NA	NA	NA	NA	NA		√ NC	o	o	√ NC	√ NC	√ NC	√
KARELIA	**	√	√	√	√	√		√	o	o	√	√	√	√
KIHD	**	√	√	√	√	√		√	√ NC	o	√	√	√	√
LASA	*	NS	NS	NS	√	√		√	√	o	o	o	o	√
LEADER	**	√	√	√	√	√		√	o	o	√	√	√	√
MALMO	**	√	√	√	√	√		√	o	o	√	√	√	√
MATISS8	**	√	√	√	√	√		√	√	√	√	√	√	√
MATISS8	**	√	√	√	√	√		√	√	√	√	√	√	√
MATISS9	**	√	√	√	√	√		√	√	√	√	√	√	√
MCVDRF	NS	NA	NA	NA	NA	NA		NS	NS	NS	√	√	√	√
MESA	**	√	√	?	√	√		√	√	√	√	√	√	√

MICOL‡	*	NA	NA	NA	NA	NA	√	√ NC	o	√	√	√	√
MOGERA	**	√	√	√	NA	NA	√	√ NC	o	√	√	√	√
MOGERA	**	√	√	√	NA	NA	√	√ NC	o	√	√	√	√
MOGERA	**	√	√	√	NA	NA	√	√ NC	o	√	√	√	√
MONFRI8	**	√	√	√	√	√	√	√	√	√	√	√	√
MONFRI8	**	√	√	√	√	√	√	√	√	√	√	√	√
MONFRI9	**	√	√	√	√	√	√	√	√	√	√	√	√
MONICA‡	*	NA	NA	NA	NA	NA	√	√ NC	o	√	√	√	√
MORGEN	*	NA	NA	NA	NA	NA	NS	NS	NS	√	√	√	√
MOSWEG	**	√	√	√	√	√	√	o	o	√	√	√	√
MRCOLD	*	NA	NA	NA	NA`	NA	NA	NA	NA	√	√	√	√
MRFIT	**	√	√	√	√	√	√	o	√	√	√	√	√
NCS1	*	NA	NA	NA	NA	NA	√	o	o	√	√	√	√
NCS2	*	NA	NA	NA	NA	NA	√	o	o	√	√	√	√
NCS3	*	NA	NA	NA	NA	NA	√	o	o	√	√	√	√
NFR‡	*	NA	NA	NA	NA	NA	√	√ NC	o	√	√	√	√
NHANES1	*	√	√	√	√	√	√	o	o	√	√	√	√
NHANES3	*	NA	NA	NA	NA	NA	√ NC	o	o	√ NC	√ NC	√ NC	√
NHS	**	NA	NA	NA	NA	NA	√	√ NC	o	√	√	√	√
NPHSI	**	√	√	√	√	√	√	√ NC	√	√ NC	√ NC	√ NC	√
NPHSII	**	√	√	√	√	√	√	√ NC	√	√	√	√	√
NSHS	**	√	√	√	√	√	√	o	o	√	√	√	√
OB43‡	*	NA	NA	NA	NA	NA	√	√ NC	o	√	√	√	√
OSAKA	**	√	√	√	√	√	√	√ NC	o	√	√	√	√
OSLO	*	NA	NA	NA	NA	NA	√	o	o	√	√	√	√
OYABE	**	NA	NA	NA	√	√	√	o	o	√	√	√	√
PARIS1	**	NA	NA	NA	NA	NA	√ NC	o	o	o	o	o	√
PREVEN	**	√	√	√	√	√	NS	NS	NS	√	√	√	√
PRIME	**	√	√	√	√	√	√	o	o	√	√	√	√
PROCAM	**	√	√	√	√	√	√	√NC	√	√	√	o	√
PROSPE	**	√	√	√	√	√	√	√	o	o	o	o	√
QUEBEC	**	√	√	√	√	√	√	o	√	o	o	o	√
RANCHO	*	√	√	√	√	√	√	o	o	√	√	√	√
REYK	**	√	√	√	√	√	√	√	o	√	√	√	√
RF2‡	*	NA	NA	NA	NA	NA	√	√ NC	o	√	√	√	√
ROTT	**	√	√	√	NA	NA	√	√	o	√	√	√	√
SHHEC	**	√	√	√	√	√	√	√	o	√	√	√	√
SHS	**	√	√	√	√	√	√	√	o	√	√	√	√
SPEED	**	√	√	√	√	√	√	o	√	√	√	√	√
TOYAMA	**	√	√	√	√	√	NS	NS	NS	√	√	√	√
TPT	**	√	√	√	√	√	√	√ NC	NS	√	o	o	√
TROMSO	*	√	√	√	√	√	√	√	√	√	√	√	√
ULSAM	**	√	√	√	√	√	√	o	o	√	√	√	√
USPHS2	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
VHMPP	*	NA	NA	NA	NA	NA	√	o	o	√	√	√	√
VITA	**	√	√	√	√	√	√	√NC	o	√	√	√	√
WHIHABP	**	NC	NC	NC	√	√	√	o	o	√	√ NC	√ NC	√
WHITE I	*	NA	NA	NA	NA	NA	NS	NS	NS	√	√	√	√
WHITE II	*	√	√	√	√	√	√	o	o	√	√	√	√
WHS	**	√	√	√	√	√	√	o	o	√	√	√	√
WOSCOP	**	√	√	√	√	√	√	√	√	o	o	o	√
ZARAGO	**	√	√	√	NS	NS	√	√NC	o	√	o	o	√
ZUTE	**	√	√	√	√	√	√	o	o	√	√	√	√

–: Not recorded; +: Self-report only; ++: Self-report supplemented by objective criteria (e.g., Electrocardiogram, Physical examination). * Death certificate only; ** Death certificate supplemented by medical record. SAH: Subarachnoid haemorrhage; NS: Not stated. NC = reportedly measured but data not contributed to the ERFC; NA = not applicable, where cohorts contributed data on fatal endpoints only. 0: Feature not included in criteria; √: Feature included in criteria.

Table 2.7: Summary of endpoints for each study with information on smoking status, age and sex

	Cardiovascular outcomes															
Cohort abbreviation	**All cardiovascular**	**CHD death and non fatal MI**	**Myocardial Infarction (non fatal)**	**All CHD (fatal)**	**All cerebrovascular (fatal and non fatal)**	**Ischaemic stroke (fatal and non fatal)**	**Haemorrhagic stroke (fatal and non fatal)**	**Subarachnoid haemorrhage (fatal)**	**Heart failure (fatal)**	**Sudden death (fatal)**	**Cardiac dysrhytmia (fatal)**	**Pulmonary embolism (fatal)**	**Aortic aneurysm (fatal)**	**Peripheral vascular disease (fatal)**	**Lung cancer (fatal)**	**All causes of death**
AFTCAPS	23	18	17	1	2	2	0	0	0	2	0	0	0	0	0	8
ALLHAT	1666	1124	1119	5	542	0	0	0	0	0	0	0	0	0	0	11
ARIC	1637	872	672	200	562	453	56	33	13	0	21	12	13	7	228	1507
ATENA	30	18	17	1	4	1	2	1	0	0	2	0	1	0	4	40
ATS_SAR	32	19	0	19	8	0	1	1	0	0	0	0	1	0	13	113
ATTICA	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50
AUSDIAB	127	77	41	36	38	12	2	4	2	0	2	1	3	2	21	269
BHS	914	509	0	509	217	22	21	4	37	0	11	9	37	9	84	1932
BRHS	1854	1213	674	539	514	7	13	10	11	0	7	12	50	8	246	1954
BRUN	151	66	28	38	63	43	19	0	9	0	0	7	5	0	0	240
BUPA	1073	723	0	723	176	22	28	8	24	0	2	17	69	2	230	2495
BWHHS	198	91	78	13	91	0	1	0	1	0	0	3	4	2	17	217
CAPS	289	249	138	111	18	3	3	1	1	0	0	7	3	1	46	340
CASTEL	73	16	3	13	15	0	0	0	28	10	0	4	0	0	0	175
CHA	4820	3000	0	3000	786	116	154	39	174	3	119	62	88	32	940	11613
CHARL	967	549	256	293	257	29	34	5	13	0	32	10	11	5	88	1205
CHS1	1063	569	366	203	442	345	62	0	0	0	0	0	0	0	95	1045
CHS2	107	53	31	22	48	39	5	0	0	0	0	0	0	0	15	89
COPEN	1013	382	348	34	430	265	53	14	46	7	8	20	12	0	122	1279
DISCO	10	8	0	8	2	0	1	0	0	0	0	0	0	0	3	26
DRECE	29	15	0	15	6	0	1	0	2	0	0	1	0	0	13	132
DUBBO	386	196	153	43	144	57	14	2	19	0	1	4	3	0	18	338
EAS	169	82	41	41	68	0	3	2	3	0	1	1	4	0	33	284
EMOFRI	8	2	2	0	3	2	1	0	0	0	3	0	0	0	0	9
EPICNOR	503	484	257	227	0	0	0	0	0	0	0	0	0	0	0	353
ESTHER	52	18	16	2	31	1	0	0	0	0	0	0	2	0	7	28
FIA	584	584	448	136	0	0	0	0	0	0	0	0	0	0	0	136
FINE_FIN	10	5	2	3	5	1	0	0	0	0	0	0	0	0	9	24
FINE_IT	208	66	18	48	103	4	5	0	19	2	1	0	1	1	17	328
FINRISK92	323	162	123	39	136	85	39	3	6	2	1	6	1	0	18	268
FINRISK97	219	103	75	28	91	64	16	1	9	0	0	9	1	0	10	190
FLETCHER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FRAMOFF	23	18	17	1	5	5	0	0	0	0	0	0	0	0	0	22
FUNAGATA	49	13	6	7	33	17	6	1	0	0	0	0	1	1	4	50
GLOSTRUP	76	76	58	18	0	0	0	0	0	0	0	0	0	0	0	18
GOH	437	201	0	201	99	2	8	2	14	12	50	6	2	3	29	1422
GOTO13	208	131	130	1	59	0	1	0	9	2	0	3	1	0	17	143
GOTO33	44	27	14	13	8	0	0	0	0	0	2	2	0	0	7	81
GOTO43	48	29	28	1	17	12	1	2	0	0	0	0	1	0	3	25
GOTOW	369	146	94	52	178	2	0	0	1	0	2	0	1	1	14	407
GREPCO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5
GRIPS	409	272	272	0	95	0	0	0	8	15	0	12	3	0	28	181
GUBBIO	107	69	0	69	29	11	2	1	0	0	0	1	0	1	12	239
HBS	135	91	0	91	27	0	0	0	0	0	0	0	0	0	0	417
HELSINAG	96	41	0	41	32	17	2	0	1	0	2	7	3	2	4	203
HISAYAMA	356	77	67	10	220	148	49	21	2	0	1	2	12	2	30	387
HONOL	298	151	110	41	124	12	39	2	0	0	3	1	7	1	38	493
HOORN	170	73	60	13	53	3	4	0	11	13	11	3	4	0	19	211
HPFS1	4259	2492	0	2492	709	97	127	36	198	341	55	41	141	11	745	12290
IKNS	495	84	37	47	344	158	71	25	57	0	2	1	3	0	18	757
ISRAEL	987	723	0	723	264	0	0	0	0	0	0	0	0	0	0	2531
KARELIA	2970	1854	1303	551	882	64	44	34	80	3	13	35	17	2	142	2357
KIHD	219	154	151	3	48	34	10	2	2	0	0	0	2	1	30	162
LASA	34	22	22	0	12	0	0	0	0	0	0	0	0	0	0	333
LEADER	93	53	20	33	33	24	2	0	2	0	1	1	1	2	16	93
MALMO	2413	2043	1229	814	143	36	49	21	18	1	6	17	46	3	334	3270
MATISS83	335	83	47	36	99	26	10	3	53	0	71	1	0	0	12	410
MATISS87	166	43	21	22	55	9	7	2	26	1	34	0	0	0	7	200
MATISS93	31	14	11	3	7	1	2	1	5	0	4	0	0	0	0	29
MCVDRFP	456	196	0	196	97	15	31	14	27	8	19	8	16	4	247	1775
MESA	105	59	47	12	41	30	9	1	0	0	0	0	0	0	0	102
MICOL	147	102	0	102	33	7	3	0	0	0	0	3	3	2	75	507
MOGERAUG1	108	79	47	32	5	0	2	0	10	2	0	5	2	1	8	126
MOGERAUG2	129	104	63	41	7	1	1	1	8	1	0	4	0	0	16	199
MOGERAUG3	36	18	11	7	5	2	1	0	3	0	0	2	0	0	2	54
MONFRI86	107	28	20	8	26	14	5	2	4	1	43	2	1	0	5	166
MONFRI89	82	28	22	6	20	10	5	0	6	0	23	2	1	0	1	100
MONFRI94	39	10	10	0	17	6	7	1	0	1	9	0	1	0	0	40
MONICA	38	28	0	28	8	0	1	0	0	0	0	0	0	0	10	100
MORGEN	149	77	0	77	24	3	10	7	5	4	4	3	6	0	80	586
MOSWEGOT	307	155	116	39	132	75	19	22	1	0	2	7	2	1	15	234
MRCOLD	2632	1146	0	1146	842	53	61	13	170	0	61	47	93	47	219	6301

MRFIT	901	772	588	184	80	5	4	8	0	0	8	6	5	0	62	484
NCS1	545	372	0	372	67	9	17	26	8	43	5	2	12	0	75	1430
NCS2	270	187	0	187	25	2	5	10	1	20	5	1	3	1	43	778
NCS3	446	274	0	274	81	8	21	20	3	38	5	0	5	0	60	952
NFR	124	90	0	90	27	2	9	1	0	0	0	1	4	0	39	330
NHANESI	1221	649	237	412	343	89	27	10	24	1	35	9	11	5	89	1691
NHANESIII	756	423	0	423	131	0	0	0	38	0	0	0	14	0	226	2064
NHS1	5198	2267	0	2267	1327	23	104	648	365	342	1	100	150	155	2194	23241
NPHSI	200	157	89	68	24	0	0	0	0	0	0	0	0	0	0	215
NPHSII	211	140	126	14	46	27	5	3	0	13	0	2	5	0	17	135
NSHS	58	13	0	13	38	1	1	1	3	0	4	0	0	0	0	23
OB43	23	14	0	14	8	1	1	1	0	0	0	0	1	0	5	76
OSAKA	261	42	26	16	144	57	27	16	62	1	1	0	4	0	10	628
OSLO	2575	1585	0	1585	372	55	78	30	60	116	35	14	155	5	503	5816
OYABE	67	13	0	13	45	33	6	5	5	0	0	0	0	0	13	134
PARIS1	436	174	0	174	91	2	29	3	3	24	21	0	7	4	119	1899
PREVEND	127	87	72	15	17	0	9	6	2	0	3	0	8	1	24	183
PRIME	205	144	127	17	42	33	6	0	0	16	0	0	0	0	29	183
PROCAM	594	405	304	101	79	59	14	0	7	76	2	7	8	0	92	753
PROSPER	396	267	202	65	115	0	0	0	0	0	0	0	0	0	0	243
QUEBEC	423	297	260	37	93	0	0	0	4	29	0	0	0	0	0	414
RANCHO	543	246	243	3	197	0	1	0	10	0	9	1	5	5	36	486
REYK	4531	3245	2028	1217	768	183	162	45	82	12	45	78	71	6	532	6672
RF2	90	64	0	64	18	2	7	0	0	0	0	2	1	0	27	320
ROTT	313	100	85	15	79	28	7	1	41	30	1	0	16	2	56	718
SHHEC	674	453	321	132	183	55	21	21	3	2	2	2	6	1	122	757
SHS	782	449	303	146	214	8	10	0	15	4	24	6	2	4	39	1155
SPEED	354	253	98	155	77	66	2	1	0	0	1	5	9	0	69	475
TOYAMA	92	34	33	1	51	24	17	10	4	0	0	0	0	0	6	83
TPT	737	555	367	188	145	82	16	8	6	0	2	5	17	0	146	675
TROMSØ	1001	604	550	54	365	278	40	32	1	9	2	0	7	1	80	505
ULSAM	770	458	336	122	229	134	40	16	15	0	4	6	18	2	65	737
USPHS2	643	310	282	28	259	217	40	0	0	38	0	0	0	0	0	791
VHMPP	3281	1683	0	1683	783	81	122	24	184	1	61	45	57	34	460	6929
VITA	26	18	15	3	5	5	0	0	0	0	2	0	1	0	4	23
WHIHABPS	768	48	42	6	719	719	0	0	0	0	0	0	0	0	0	38
WHITEI	470	218	0	218	138	19	14	4	20	0	12	6	40	4	62	1228
WHITEII	349	317	255	62	10	2	2	2	1	0	0	4	3	1	15	328
WHS	603	236	228	8	286	239	26	19	0	52	0	0	0	0	0	625
WOSCOPS	259	213	174	39	39	0	0	0	0	0	0	0	0	0	0	111
ZARAGOZA	80	43	30	13	37	7	0	0	0	0	0	0	0	0	0	20
ZUTE	41	16	13	3	14	1	1	0	3	0	0	1	7	0	4	65
Total	***69174***	***40218***	***16390***	***23828***	***17445***	***5023***	***2014***	***1313***	***2108***	***1298***	***919***	***704***	***1331***	***385***	***9788***	***128137***

Note: Some studies did not provide information regarding non-fatal myocardial infarction or non fatal stroke. Some studies did not provide information on type of cerebrovascular event (ischaemic versus haemorrhagic stroke).

List 2.1: List of core variables sought in the Emerging Risk Factors Collaboration

From baseline examination

- Date of baseline survey
- Unique (but anonymous) participant identifier
- Date of birth (or age at baseline) and sex
- Unique identifier for case-control matched sets for studies in which controls are 'individually matched' to cases

Clinical and biochemical measurements made at baseline examination

- Ethnicity
- Smoking and alcohol use (current / ex / never; amount / duration etc.)
- Use of cardiovascular medications (current and past use, in as much detail as possible, including anti-hypertensive drugs, 'statins', fibrates) and other medications (e.g. hypoglycemic agents, hormone replacement therapy) – also, treatment allocation made in randomized controlled trials
- Use of postmenopausal hormone therapy or oral contraceptives
- Prior history of coronary heart disease (in particular myocardial infarction and angina), stroke, transient ischemic attack (TIA), peripheral vascular disease (PVD) and diabetes
- Systolic and diastolic blood pressure
- Weight, height, waist and hip circumference
- Physical activity and socio-economic status
- Total, high- and low-density lipoprotein cholesterol (including particle size and numbers, where available); triglycerides; lipoprotein (a); apolipoprotein-AI and -B (including information about fasting status at the time blood samples were taken); lipoprotein-associated phospholipase A_2 mass and activity levels
- Inflammatory markers (including C-reactive protein, fibrinogen, albumin, interleukin-6 and the leucocyte count)
- Creatinine, uric acid
- Haemostatic factors (including von-Willebrand factor, fibrin D-dimer)
- Metabolic factors (including fasting glucose, post load glucose, glycosylated haemoglobin and insulin)

From re-survey examinations

- The unique (but anonymous) participant identifier used for baseline visit
- Date of the visit (or, if not available, age at visit)
- Data on baseline items that were collected at repeat surveys (particularly established risk factors and other biochemical markers)

Non-fatal events during follow-up

- Myocardial infarction and date of MI
- Stroke (including subtype if available: e.g. ischaemic / haemorrhagic) and date of stroke
- Other subsidiary cardiovascular outcomes: e.g. angina, PVD, coronary artery bypass grafting (CABG), percutaneous transluminal coronary angioplasty (PTCA), congestive heart failure
- Dates of censoring for end of follow-up for non-fatal events

Fatal events during follow-up

- Date last known to be alive (if not recorded as dead)
- Date of death (or, if not available, age at death)
- Underlying cause of death (preferably coded according to some specified version of the three-digit International Classification of Diseases (ICD); but if a three-digit ICD code is not available then whatever code the study already uses)
- Date of censoring for end of follow-up for fatal cases

Source: Extract from the Protocol of ERFC [2].

List 2.2: Data request sent to all the ERFC investigators in December 2010

ERFC

Emerging Risk Factors Collaboration

Name of study: _____

Name of contact: _____

Information on behavioural, social, and hereditary characteristics

Some studies sharing data in the ERFC have already provided the Coordinating Centre with all available data (see below for list of relevant variables). If this applies to your study, please tick the box immediately below and return this form to the Coordinating Centre.

☐ Our study has previously provided all available data on a cohort-wide basis on the variables listed below.

If, however, your study has not provided complete data on **a cohort-wide basis** on the variables listed below, please complete and return the remainder of this form.

Many thanks.

Variables	Available at the "baseline" survey		Available at any subsequent resurvey?		Willing to share in the ERFC?	
More detailed information sought:						
Smoking: eg, amount, duration, cessation date	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
Alcohol consumption: eg, amount, type, frequency, pattern	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
Physical activity: eg, type, duration, frequency	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
Socioeconomic characteristics: eg, income, education, occupation, marital status, number of children	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
Medication use: eg, statins, antihypertensives, aspirin, HRT	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
New information sought:						
Family history of CVD: eg, family members, age at onset, definition of disease	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
Known coronary genetic loci: eg, 9p21, 1p13, 6q25, 1q22, 1q41, 6p24, 19p13, 1p32, 10q11, 3q22	☐ Yes	☐ No OR	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
History of chronic disease: eg, rheumatoid arthritis, COPD, kidney disease, disability	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
Physiological measurement: eg, ECG, heart rate, spirometry	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
Mental & general health: eg, depression, anxiety, sleep, self-rated health	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
Non-fatal incidence of:						
Heart failure	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no
Atrial fibrillation/flutter	☐ Yes	☐ No	☐ Yes	☐ No	☐ Probably yes	☐ Probably no

Please reply to Mat Walker by **31 January 2011,** by email (erfc@phpc.cam.ac.uk), or by fax (+44-[0]1223-741339), or by post to: ERFC Coordinating Centre, University of Cambridge, Strangeways Laboratory, Wort's Causeway, Cambridge, CB1 8RN, UK

Source: Personal communication from Dr M. Walker, data manager of ERFC.

List 2.3: List of study acronyms in the ERFC

AFTCAPS, Air Force/Texas Coronary Atherosclerosis Prevention Study; **ALLHAT**, Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial; **AMORIS**, Apolipoprotein Related Mortality Risk Study; **ARIC**, Atherosclerosis Risk in Communities Study; **ATENA**, cohort of Progetto CUORE; **ATS_SAR**, cohort of Risk Factors and Life Expectancy Pooling Project; **ATTICA**, ATTICA Study; **AUSDIAB**, Australian Diabetes, Obesity and Lifestyle Study; **BHS**, Busselton Health Study; **BRHS**, British Regional Heart Study; **BRUN**, Bruneck Study; **BUPA**, BUPA Study; **BWHHS**, British Women's Heart and Health Study; **CaPS**, Caerphilly Prospective Study; **CASTEL**, Cardiovascular Study in the Elderly; **CHA**, Chicago Heart Association Study; **CHARL**, Charleston Heart Study; **CHS-1**, original cohort of the Cardiovascular Health Study; **CHS-2**, supplemental African-American cohort of the Cardiovascular Health Study; **COPEN**, Copenhagen City Heart Study; **DISCO**, cohort of Risk Factors and Life Expectancy Pooling Project; **CUORE**, Progetto CUORE; **DRECE**, Diet and Risk of Cardiovascular Disease in Spain; **DUBBO**, Dubbo Study of the Elderly; **EAS**, Edinburgh Artery Study; **EMOFRI**, part of CUORE; **EPESEBOS**, The Established Populations for the Epidemiologic Study of the Elderly Studies, Boston; **EPESEIOW**, The Established Populations for the Epidemiologic Study of the Elderly Studies, Iowa; **EPESENCA**, The Established Populations for the Epidemiologic Study of the Elderly Studies, North Carolina; **EPESENHA**, The Established Populations for the Epidemiologic Study of the Elderly Studies, New Haven; **EPICNOR**, European Prospective Investigation of Cancer Norfolk Study; **ESTHER**, Epidemiologische Studie zu Chancen der Verhutung und optimierten Therapie chronischer Erkrankungen in der alteren Bevolkerung; **FIA**, First Myocardial Infarction in Northern Sweden; **FINE-FIN**, Finland, Italy and Netherlands Elderly Study - Finland cohort; **FINE-IT**, Finland, Italy and Netherlands Elderly Study – Italian cohort; **FLETCHER**, Fletcher Challenge Blood Study; **FINRISK-92**, Finrisk Cohort 1992; **FINRISK-97**, Finrisk Cohort 1997; **FRAMOFF**, Framingham Offspring Study; **FUNAGATA**, The Funagata Study; **GLOSTRUP**, Research Centre for Prevention and Health; **GOH**, The Glucose Intolerance, Obesity and Hypertension Study; **GOTO13**, Goteborg Study 1913; **GOTO33**, Göteborg 1933 Study; **GOTO43**, Göteborg 1943 Study; **GOTOW**, Population Study of Women in Gothenburg, Sweden; **GREPCO**, cohort of Risk Factors and Life Expectancy Pooling Project; **GRIPS**, Göttingen Risk Incidence and Prevalence Study; **GUBBIO**, cohort of Risk Factors and Life Expectancy Pooling Project; HBS, Helsinki Businessmen Study; **HELSINAG**, Helsinki Aging Study; **HISAYAMA**, Hisayama Study; **HONOL**, Honolulu Heart Program; **HOORN**, Hoorn 275 Study; **HPFS**, Health Professionals Follow-up Study; **IKNS**, Ikawa, Kyowa, and Noichi Study; **ISRAEL**, Israeli Ischaemic Heart Disease Study; **KARELIA**, North Karelia Project; **KIHD**, Kuopio Ischaemic Heart Disease Study; **LASA**, Longitudinal Aging Study Amsterdam; **LEADER**, Lower Extremity Arterial Disease Event Reduction Trial; **MALMO**, Malmö Study; **MATISS-83**, cohort of Progetto CUORE; **MATISS-87**, cohort of Progetto CUORE; **MATISS-93**, cohort of Progetto CUORE; **MCVDRFP**, Monitoring of CVD Risk Factors Project; **MESA**, Multi-Ethnic Study of Atherosclerosis; **MICOL**, cohort of Risk

Factors and Life Expectancy Pooling Project; **MOGERAUG1**, MONICA/KORA Augsburg Surveys S1; **MOGERAUG2**, MONICA/KORA Augsburg Surveys S2; **MOGERAUG3**, MONICA/KORA Augsburg Surveys S3; **MONFRI-86,** cohort of Progetto CUORE; **MONFRI-89**, cohort of Progetto CUORE; **MONFRI-94**, cohort of Progetto CUORE; **MONICA**, cohort of Risk Factors and Life Expectancy Pooling Project; **MORGEN**, Monitoring Project on Chronic Disease Risk Factors; **MOSWEGOT**, MONICA Göteborg Study; **MRCOLD**, MRC Study of Older People; **MRFIT**, Multiple Risk Factor Intervention Trial 1; **NCS 1, 2 and 3**, Norwegian Counties Studies; **NFR**, cohort of Risk Factors and Life Expectancy Pooling Project; **NHANES I**, First National Health and Nutrition Examination Survey; **NHANES III,** Third National Health and Nutrition Examination Survey; **NHS**, Nurses' Health Study; **NPHSI**, Northwick Park Heart Study I; **NPHSII**, Northwick Park Heart Study II; **NSHS**, Nova Scotia Health Survey; **OB43**, cohort of Risk Factors and Life Expectancy Pooling Project; **OSAKA**, Osaka Study; **OSLO**, Oslo Study; **OYABE**, Oyabe study; **PARIS1**, Paris Prospective Study I; **PREVEND,** Prevention of Renal and Vascular End Stage Disease Study; **PRHHP**, Puerto Rico Heart Health Program; **PRIME**, Prospective Epidemiological Study of Myocardial Infarction; **PROCAM**, Prospective Cardiovascular Münster Study; **PROSPER**, Prospective Study of Pravastatin in the Elderly at Risk; **QUEBEC**, Quebec Cardiovascular Study; **RANCHO**, Rancho Bernardo Study; **REYK**, Reykjavik Study; **RF2**, cohort of Risk Factors and Life Expectancy Pooling Project; **RIFLE**, Risk Factors and Life Expectancy Pooling Project; **ROTT**, The Rotterdam Study; **SHHEC**, Scottish Heart Health Extended Cohort; **SHS**, Strong Heart Study; **SPEED**, Speedwell Study; **TARFS**, Turkish Adult Risk Factor Study; **TOYAMA**, Toyama; **TROMSØ**, Tromsø Study; **ULSAM**, Uppsala Longitudinal Study of Adult Men; **USPHS**, U.S. Physicians Health Study; **USPHS2**, U.S. Physicians Health Study II; **VHMPP**, Vorarlberg Health Monitoring and Promotion Programme; **VITA**, Vicenza Thrombophilia and Athrosclerosis Project; **WHIHABPS**, Women's Health Initiative (Hormones and Biomarkers Predicting Stroke in Women); **WHITE I**, Whitehall I Study;**WHITE II**, Whitehall II Study; **WHS,** Womens Health Study; **WOSCOPS**, West of Scotland Coronary Prevention Study; **ZARAGOZA**, Zaragosa study; **ZUTE**, Zutphen Elderly Study

References

1. Strong K, Mathers C, Leeder S, Beaglehole R. Preventing chronic diseases: how many lives can we save? *Lancet.* 2005;366:1578-1582.

2. Danesh J, Erqou S, Walker M et al. The Emerging Risk Factors Collaboration: analysis of individual data on lipid, inflammatory and other markers in over 1.1 million participants in 104 prospective studies of cardiovascular diseases. *Eur J Epidemiol.* 2007;22:839-869.

3. Sarwar N, Sandhu MS, Ricketts SL et al. Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies. *Lancet.* 2010;375:1634-1639.

4. Di AE, Sarwar N, Perry P et al. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA.* 2009;302:1993-2000.

5. Erqou S, Kaptoge S, Perry PL et al. Lipoprotein(a) concentration and the risk of coronary heart disease, stroke, and nonvascular mortality. *JAMA.* 2009;302:412-423.

6. Kaptoge S, Di AE, Lowe G et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *Lancet.* 2010;375:132-140.

7. Wormser D, Kaptoge S, Di AE et al. Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. *Lancet.* 2011;377:1085-1095.

8. Seshasai SR, Kaptoge S, Thompson A et al. Diabetes mellitus, fasting glucose, and risk of cause-specific death. *N Engl J Med.* 2011;364:829-841.

9. Sarwar N, Gao P, Seshasai SR et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. *Lancet.* 2010;375:2215-2222.

10. Thun, M. J., Day-Lally, C., Myers, DG, Calle, EE, Flanders, WD, Zhu, B-P, Namboodiri, MM, and Heath, CW Jr. Trends in tobacco smoking and mortality from cigarette use in Cancer Prevention Studies I (1959 through 1965) and II (1982 through 1988). Changes in Cigarette-Related Disease Risks and Their Implication for Prevention and Control.Smoking and Tobacco Control Monograph No.8.Rockville (MD) [*a*], 305-382. 1997. Shopland DR, Burns DM, Garfinkel L, Samet JM.

11. Iverson C, Christiansen S, Flanagin A, et al. *American Medical Association Manual of Style: A Guide for Authors and Editors*. 10 ed. New York, NY: 2007.

12. Whitlock G, Lewington S, Sherliker P et al. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet.* 2009;373:1083-1096.

13. O'Connor RJ, McNeill A, Borland R et al. Smokers' beliefs about the relative safety of other tobacco products: findings from the ITC collaboration. *Nicotine Tob Res.* 2007;9:1033-1042.

14. Delnevo CD. Smokers' choice: what explains the steady growth of cigar use in the U.S.? *Public Health Rep.* 2006;121:116-119.

Chapter 3: Distribution and cross-sectional correlates of smoking in developed countries

Summary

A smoker is exposed to the inhaled smoke as well as to the ingested tobacco when smoking a cigarette, pipe or cigar. As the tar and gas phases of tobacco consist of thousands of chemical compounds, it has proven difficult to pinpoint the causal element which promotes CVD. A better understanding of correlates of tobacco may shed light on how smoking induces heart diseases. Another reason for investigating correlates of tobacco is appropriate adjustment of the association of tobacco with CVD for potential confounders, enhancing quality of the estimates of risk and the chances that an observed relationship is causal.

In this Chapter, I first summarize the distribution of smoking variables, and then investigate correlations between smoking and conventional CVD risk factors such as blood pressure, lipid levels, inflammatory markers and adiposity. Up to 929,335 individuals from 114 studies were included in these analyses. A majority of current and ex-smokers had initiated smoking at age 18 years old or younger and 75% had begun before age 22 years old. BMI was slightly lower in current (25.4 kg/m^2, SD: 3.8) compared to ex (26.4 kg/m^2, SD: 4.0) and never smokers (26.3 kg/m^2, SD: 4.2). Current smokers were more likely to be male and white, to declare being current alcohol drinkers and to be employed in skilled rather than manual jobs. Current smokers, especially women, had higher baseline levels of total cholesterol, non-HDL-C, $\log_e$ triglycerides, and ApoB, while they had lower levels of HDL-C and ApoA1 compared to never smokers. Smoking was most strongly correlated to markers of inflammation. There was a positive gradient from never to ex-smokers and to current smokers for CRP, fibrinogen, leukocyte counts and interleukin-6. In current smokers, per SD increase in number pack-years, CRP was raised by 0.18 mg/l (95% CI: 0.11 to 0.24) and fibrinogen by 0.22 µmol/l (0.16 to 0.28).

3.1 Background

Vasomotor dysfunction, inflammation and modification of lipids are all responsible for the development of atherosclerotic plaques in the arteries, later leading to thrombosis and clinical CVD. Smoking has been shown to accelerate atherosclerosis and promote thrombosis in the coronary arteries, the aorta, the carotid and cerebral arteries and the large arteries in the peripheral circulation [1] (see **Section 1.4**). Studies have shown a relationship between smoking and hyperlipidaemia [3], blood pressure [4], inflammation [5,6], oxidative stress and platelet activation [7], and a role for nicotine, carbone monoxide and oxidant chemicals has been suggested. However, these studies have been characterized by small sample sizes and inconsistent adjustment. Studies have, as a result, sometimes found inconsistent results, and the strength of correlations between smoking and other known cardiovascular risk factors remains uncertain. Published studies have also not been able to assess how adjustment for potential correlates modifies the association between smoking and CVD risk. In this context, ERFC represents a unique opportunity to adjust associations for a large range of questionnaire-based and blood-based factors significantly correlated with smoking.

In this context, there is an interest in better quantifying the relationship between smoking and a range of lifestyle and biochemical risk factors of CHD using information contained in the ERFC. Smoking status was available in 929,335 individuals from 114 studies characterized by random sampling at baseline from the general population of the country involved. Subsets had information on anthropometric measurements, blood pressure, lipid levels, and inflammatory markers.

3.2 Methods

3.2.1 Dataset

Details of study selection, data collection and harmonization within the ERFC have been described in **Chapter 1**. In total, 929,335 individuals included provided information on age, sex, smoking status and body mass index (BMI). Smaller subsets had information on other conventional risk factors: 154,674 on waist to hip ratio (WHR), 754,077 on systolic blood pressure, 738,215 on total cholesterol (on 382,079

on HDL-C), 107,879 on C-reactive protein, 10,974 on leukocyte counts and 180,379 on fibrinogen levels.

3.2.2 Statistical methods

Positively skewed variables such as Lp(a), triglycerides and C-reactive protein were log-transformed, apart from smoking pack-years which was kept untransformed for ease of interpretation. The proportion of current and ex-smokers at baseline was drawn by cohort, and ordered by decade of baseline enrollment of the study. Ninety five percent confidence intervals for the frequencies were derived assuming a binomial distribution $\mathrm{B}(p)$, with *p* the probability of being a current or ex-smoker in a specific study. Graphically, this translated into the width of the confidence interval being inversely proportional to the number of individuals in the study of a specific sex. For continuous variables, box-plots were represented for each study and grouped by decade of enrollment. Mean of smoking amount reported at baseline was drawn by sex and by tenths of smoking duration, using a linear mixed model adjusted for study and age 50 years old (see below). Unadjusted means of all continuous variables were calculated within studies and pooled across studies by random effect meta-analysis; and the pooled variance was obtained as an average of study-specific variances weighted by study size. Categorical variables were summarized as raw counts and proportions.

Statistical methods used in the ERFC for the assessment of cross-sectional correlates using linear mixed models follow the example of methods used in the Fibrinogen Study Collaboration [8]. First, I estimated the effect of 1 standard deviation increase in smoking pack-years (equal to approximately 17 pack-years) on other characteristics. Secondly, I graphically assessed the strength and shape of correlations by plotting predicted means of continuous characteristics based on linear mixed model regressions (i) over smoking pack-years, in current smokers only, divided into tenths and (ii) over smoking status categorized as current, ex and never. In comparison to a standard regression, a so-called “mixed” regression allows the effect of smoking, and additionally of selected covariates, to truly vary across studies. Studies included in the ERFC dataset recruited participants from all over the World and over a time scale of 50 years. Therefore, unexplained heterogeneity between studies may reasonably be assumed to be real heterogeneity. A mixed model relaxes the assumption of a fixed true estimate, allowing estimation of true heterogeneous

effects of smoking on correlates across studies, even after adjustment for available confounder and effect modifiers. A mixed model can be written as:

$$Y_{si} = \alpha_s + (\beta + u_s)E_{si} + \lambda X_{si} + \varepsilon_{si}$$

Where *s=1…S* indicates the study. *i=1…n_s* individuals belonging to study *s,* Y_{si} the risk factor levels, E_{si} the exposure of interest, X_{si} other covariates of adjustment. α represents the constant fixed effect of the cohort and λ is a vector of coefficients for covariates of adjustment. β is the parameter of interest, being the change in risk factor per unit increase in exposure, adjusted for covariates X_{si}. The random noise around the predicted value of the risk factor level is represented by a random variable ε and the heterogeneous effect of the exposure of interest across study is represented by another random variable u. I assume a normal distribution of these two random variables : $u_s \sim N(0, \sigma_u^2), \varepsilon_{si} \sim N(0, \sigma_r^2)$.

All mixed models included a random effect of *smoking* (and optionally *age, age^2 and sex*) at the study level, while the main effect of study was modeled as a separate fixed effect. In addition to study, models were also adjusted for at least *age and sex (*and when stated in the legends for *age^2, age x sex* and *age* and *sex* interactions with *smoking*). For each fitted mixed model, predicted means and their 95% confidence intervals were plotted, separately for men and women, and for age fixed to 50 years old. An inverse-variance weighted polynomial curve was superimposed on the graph for men and women separately, to investigate whether the association was consistent with a linear or a quadratic shape. For categorical variables, instead of the effect of 1 SD increase in pack-years, the difference in the number of pack-years compared to the reference category was reported. This was obtained by regressing pack-years over categories of the variable using a linear mixed model, and adjusting at least for age and sex.

I also reported correlation coefficients between smoking pack-years in current smokers and continuous characteristics, as they provide useful summaries when the shape of association is approximately linear. To obtain these correlation coefficients, I followed the steps listed below:

1. "Partial" Pearson correlation coefficients r_s adjusted for baseline age and sex were computed for each study *k=1…s*.

2. These coefficients were z-transformed to fulfill the assumption of normality [9] using Fisher's formula:

$$z_s = 0.5 \times \log_e((1 + r_s)/(1 - r_s))$$

3. Z-transformed coefficients were pooled across studies by random-effects meta-analysis [10]

4. The pooled estimate was back-transformed using the same formula as before to obtain the overall correlation coefficient:

$$r_{pooled} = (\exp(2 \times z_{pooled}) - 1))/(\exp(2 \times z_{pooled}) + 1)$$

All analyses were performed using STATA statistical software, release 11 (StataCorp LP, College Station, Texas, USA).

3.3 Results

3.3.1 Baseline characteristics

There was heterogeneity between studies and between sexes regarding the prevalence of current and ex-smokers, ranging for current users from 5% for USPHS2 to above 80% for ZUTE (**Figures 3.1** & **3.2**). Studies which began to enroll in the 1960s and 1970s had a tendency to report higher prevalence of current smokers and lower prevalence of ex-smokers than studies started in the 1980s, 1990s and 2000s, for both men and women (**Figure 3.3** & **3.4**). Substantial heterogeneity was also observed at a study level for other variables of smoking, namely number of pack-years in current smokers, and years since quitting smoking in ex-smokers, which were not explained by the decade in which the study was started (**Figure 3.5** & **3.6**). Within each study, women consistently reported lower prevalence of current smokers and, amongst current users, lower levels of pack-years than men.

After adjustment for age, baseline amount and duration were positively and approximately linearly correlated in both men and women, with wider confidence intervals in women due to their overall lower prevalence of smoking compared to men (**Figure 3.7**). For every 10 additional years of smoking, amount was raised by approximately 2.3 cigarettes per day for both men and women.

3.3.2 Smoking association with demographic characteristics

The proportion of current smokers was 41% amongst men and 27% amongst women (**Table 3.1**). The proportion of ex-smokers was 28% amongst men whilst it was 18% amongst women. Mean age was 53.3 years old (SD: 8.5) in current smokers, 55.0 years old in ex-smokers (SD: 8.6) and 54.3 years old in never smokers (SD: 10.2). One SD increase in age (corresponding to approximately 9 years) translated into an average of 4.5 additional smoking pack-years in current smokers at baseline. Whites were 5% more likely to report being current smokers than non-Whites (38% versus 33%)

3.3.3 Smoking association with lifestyle factors

Alcohol use and smoking were positively correlated. Current alcohol drinkers were slightly more likely than non-current alcohol drinkers (38% versus 32%) to be also current smokers (**Table 3.1**). By contrast, being a diabetic had a negative effect on the likelihood of being a smoker. Non-diabetics were 13% more likely than diabetics (36% versus 23%) to declare being current smokers and 10% less likely to have stopped smoking (26% versus 36%). Smoking was also associated with education. The percentage of current smokers was highest amongst individuals reaching secondary education (41%), whilst the percentage of never smokers was highest amongst individuals reporting no schooling. Regarding occupation, the percentage of smoker was highest amongst office workers (65% current smokers), whilst it was lowest amongst individuals not working (19%), a heterogeneous category of older individuals which was comprised of housewives, unemployed individuals and retired individuals, with a mean age of 57.4 (SD: 11.5). Leaving aside individuals not working, the percentage of never smokers was highest amongst individuals working in services (52%). Whites reported an average 9.22 (95% confidence interval of 6.36 to 12.07) additional pack-years than non-Whites at baseline. Number of smoking pack-years was not significantly different depending on alcohol status, history of diabetes, level of education and occupation.

3.3.4 Smoking association with adiposity and blood pressure

BMI was slightly lower in current (25.4 kg/m^2, SD: 3.8) compared to ex (26.4 kg/m^2, SD: 4.0) and never smokers (26.3 kg/m^2, SD: 4.2) (**Table 3.2** & **Figure 3.8**). In current smokers, correlation with pack-years was low and the Pearson correlation

coefficient was equal to 0.02 (**Figure 3.9**). In contrast, WHR was not different amongst the smoking groups and not associated with number of pack-years in current smokers (difference in WHR per 1-SD increase in pack-years was 0.01). Women current smokers had relatively lower SBP and DBP levels than ex and never smokers, but there were no significant trends with number of pack-years (partial correlations equal to 0). Associations of blood pressure with number of years since stopping to smoke in ex-smokers were also null (**Figure 3.10**).

3.3.5 Smoking association with lipids

Current smokers, especially women, had higher baseline levels of total cholesterol, non-HDL-C, $\log_e$ triglycerides, ApoB; while they had lower levels of HDL-C and ApoA1 compared to never smokers (**Table 3.2**). Levels of lipids and lipoproteins were not significantly different for ex-smokers compared to those of never smokers, except for total cholesterol, where they were slightly raised (**Figure 3.8**). When focusing on current smokers and looking at the relationships with number of pack-years smoking, there were slight positive relationships for total cholesterol, non-HDL-C, $\log_e$ triglycerides, ApoB; and slightly negative relationships for HDL-C and ApoA1. Correlation coefficients with pack-years smoking were all less than 0.10 with these variables. In ex-smokers, log triglycerides decreased linearly and HDL-C increased slightly with number of years since stopping (**Figure 3.9**).

3.3.6 Smoking association with inflammation

Smoking was most strongly correlated to markers of inflammation. There was a positive gradient from never to ex-smokers and to current smokers for CRP, fibrinogen, leukocyte counts and interleukin-6. In current smokers, per SD increased in pack-years, CRP was raised by 0.18 mg/l (95% CI: 0.11 to 0.24). The corresponding increase for fibrinogen was 0.22 µmol/l (0.16 to 0.28) (**Table 3.2**). For all inflammatory markers (apart from albumin men ex-smokers), levels in ex-smokers were not significantly different from levels in never smokers. The shape of association for these markers was curvilinear, with an apparent tailoring of the association for $\log_e$ leukocyte counts and $\log_e$ interleukin-6 above 40 pack-years for both men and women. Overall, partial correlations were 0.14 for CRP, 0.11 for fibrinogen, 0.17 for $\log_e$ leukocyte counts, 0.13 for $\log_e$ interleukin-6 and -0.04 for albumin (**Figure 3.9**). In ex-smokers $\log_e$ C-reactive protein and $\log_e$ leukocyte

counts were negatively associated with the number of years since stopping smoking, while associations with other markers were not significant (**Figure 3.10**).

3.4 Discussion

Smoking status, defined as being a current, ex- or never smoker, was available in 929,335 individuals from 114 studies, around a half of participants being women. In that respect, the ERFC represents an effort similar in scale to Cancer Prevention Studies conducted in the USA in the 1960's and 1980's, which both included more than 1 million participants but did not have information on biochemical risk factors relevant to CVD such as blood pressure, lipid levels and inflammatory biomarkers [11]. The present analysis is the most comprehensive and detailed analysis on the cross-sectional correlates of smoking status and dose-duration (as measured by pack-years) to date, with adequate power to assess shapes of association separately for each gender.

3.4.1 Baseline characteristics

The prevalence of active smokers was 34% in the overall data and was higher than current levels in the developed world which comprise between 20% and 25% [12]. There was evidence of a gender gap, with prevalence of current smoking in men much higher than in women (41% versus 27%). A majority of studies included in the ERFC recruited individuals in the 1960's and 1970's and these percentages are a reflection of a time period when smoking was more prevalent, especially amongst men, in developed countries [13]. The drop in smoking prevalence, and the narrowing in the gender gap observed since the 1960's, was observable when plotting sex-specific prevalence within studies by starting decade of recruitment.

A positive and approximately linear association was found between baseline amount and baseline duration of smoking. For every 10 additional years of smoking, smoking amount was raised by approximately 2.3 cigarettes per day for both men and women. This cross-sectional observation shows that continuing smokers are likely to increase their smoking intake over time. In my dataset, most smokers had started in teenagehood or early adulthood: median age was 18 years old and inter-quartile range 16 to 22 years old. Studies have shown that adolescents initiating smoking are particularly vulnerable to becoming dependent on nicotine and rapidly increase their intake over time [14]. The mechanisms by which nicotine establishes dependence in smokers are thought to involve adaptative changes of nicotinic acetylcholine receptors, which are widely distributed through the central nervous system, [15]. The

activation of these receptors by nicotine increases the release of neurotransmitters and produces effects on a large number of physiological processes such as locomotion, anxiety, learning and memory.

3.4.2 Smoking and lifestyle factors

In the ERFC, current smokers were over-represented amongst individuals reaching secondary education and working in offices, and never smokers were over-represented amongst individuals without schooling, not working or in manual jobs. These are surprising findings which contrast with previous reports showing an over-representation of smokers in deprived socio-economic groups characterized by a poor diet, low levels of physical activity and high prevalence of alcohol drinking [16,17], as a result of higher exposure to public smoking [18] and lower education about smoking hazards [19]. One reason for this could be the over-representation of unqualified women in never smokers as well as older individuals having reached retirement age. Alcohol drinking and smoking were correlated habits. In that respect, sociological studies have shown that smoking tends to become socially acceptable and normative in student parties where alcohol is also consumed, whereas it is stigmatized the rest of the time [20]. Diabetics were less likely to smoke and more likely to have stopped smoking than non-diabetics. This may reflect a greater awareness of the health effects of smoking amongst diabetics in the developed world.

3.4.3 Smoking and adiposity

It has been previously reported that smokers have decreased BMI but increased central adiposity [21-23]. In the ERFC, never smokers were leaner than current smokers, and WHR was not significantly associated with the number of pack-years amongst current smokers. These paradoxical effects are thought to result from increased resting metabolic rate, increased sympathetic nervous system activity and thermogenesis in smokers [23]. Never and ex-smokers had similar BMI levels, which contrasts with available epidemiological evidence showing that smoking cessation is accompanied by weight gain, a stronger appetite and an increase in adipose tissue lipoprotein lipase activity [23].

3.4.4 Smoking and blood pressure

Smoking is thought to affect heart rate and blood pressure [24] Nicotine uptake has been shown to immediately activate the sympathetic nervous system [25]. In both animal and human models, studies have shown that active and passive cigarette smoke exposures are associated with a decrease in vasodilatory function [2]. In humans, cigarette smoke exposure impaired endothelium-dependent vasodilatation in macrovascular beds such as coronary and brachial arteries and in microvascular beds. Cigarette smoke has also been shown to alter nitric oxide biosynthesis, a vasoregulatory molecule which helps regulate inflammation, leucocyte adhesion, platelet activation and thrombosis. In the ERFC, current, never and ex-smokers had similar distributions of SBP and DBP, with slightly lower levels in never smokers. This surprising lack of correlation between smoking and blood pressure could be due to smoking having short-term rather than chronic effects on blood pressure. One consequence is that most of the smoking effect on CVD is unlikely to be mediated through modification of blood pressure levels.

3.4.5 Smoking and lipids

Smoking could promote atherosclerosis, in part, through modifying the lipid profile of smokers [26]. In the ERFC, associations with cholesterol and lipoprotein levels were significant but modest and correlations were below 0.10. There were positive correlations between smoking and $\log_e$ triglycerides, total cholesterol, non-HDL-C and Apolipoprotein B; and negative correlations with HDL-C and Apolipoprotein A. Several mechanisms have been proposed to explain these correlations. They include lipid oxidation, changes in composition of lipoproteins, alteration in plasma- and lipoprotein- associated lipid transfer enzymes, changes in metabolism of fatty acids, effects of levels of postprandial lipids and changes in cholesterol fluxes, particularly reverse cholesterol transport [27]. Specifically, higher levels of triglycerides and lower levels of HDL-C have been suggested to be related to insulin resistance [2]. Smoking has also been shown to increase oxidative modification of LDL and decrease plasma activity of paraoxanase, an enzyme that protects against LDL oxidation [2].

3.4.6 Smoking and inflammation

In the ERFC, current smoking was positively associated with several markers of inflammation and most strongly with CRP and fibrinogen; and there was a negative trend with number of years since stopping. Correlations between smoking and chronic low grade inflammation, such as leukocyte counts, C-reactive protein,

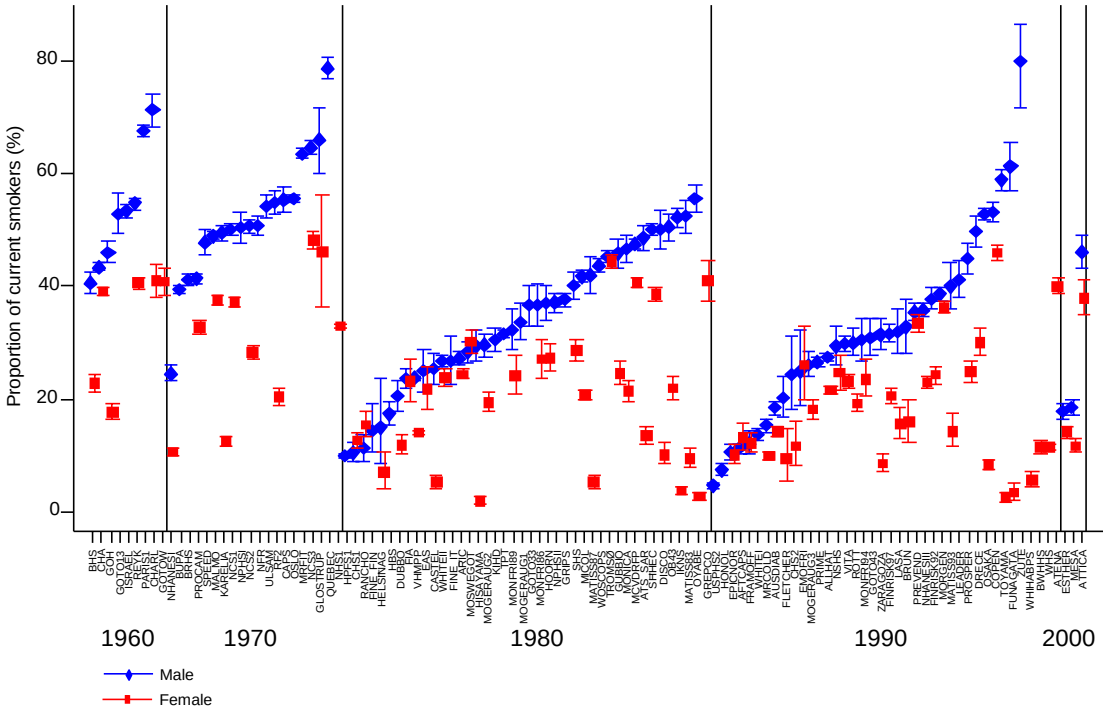
interleukin-6 and fibrinogen, have already been reported in the literature [5-7,28]. Markers of inflammation are known to be associated with an increased CVD risk [29,30], but whether they are causally related to CVD remains subject to debate. Local recruitment of leukocytes on the surface of endothelial cells is an early event in atherosclerosis [2]. Elevation of various pro-inflammatory cytokines increases leukocyte-endothelial cell interaction leading to leukocyte recruitment; and soluble VCAM-1, ICAM-1 and E-selecting levels have been shown to be elevated in smokers [2]. Cigarette smoking has also been found to activate pro-atherogenic molecules leading to alteration in cell-cell interactions [2]. However, the association between CRP and vascular disease has been shown to attenuate considerably after adjustment for non-inflammatory risk factors [5,28] and a Mendelian randomization study has reported a lack of concordance of the associations of CRP genotypes and CRP concentrations with CVD risk [31]. The association between fibrinogen and CVD has also been shown to be partially altered by progressive adjustment [29]. These observations translate into uncertainties regarding the relevance of inflammation as mediating the association between smoking and CVD and therefore the need to consider them as potential confounders.

3.4.7 Limitations

Despite its strengths, this analysis contains several limitations. Prevalence of smoking was highly variable across studies and only part of this heterogeneity could be attributed to study characteristics such as decade of baseline survey. Limited information was available on the criteria used by each study to define categories of “current”, “ex” and “never” smokers (such as “smoking at least once over the past day/month/year or the interview”), which may have accounted for part of the heterogeneity. Efforts to retrieve this information proved unsuccessful because this information was rarely published by studies and had not been requested by ERFC investigators. Evidence of digit preference in the reporting of amount and duration may partly be attributed to misreporting or difficulty in recalling the exact number of years since starting smoking and precisely estimating the amount smoked each day. In that respect, distortion in levels of biochemical risk factors and misreporting of smoking status were minimized by excluding individuals with pre-existing CVD. More accurate information on smoking amount, for example measurement of biomarkers such as cotinine levels [32], was not available. Despite efforts in harmonizing coding across studies, measurement error may have been introduced when recoding level of education and occupation which were assessed differently across study designs and

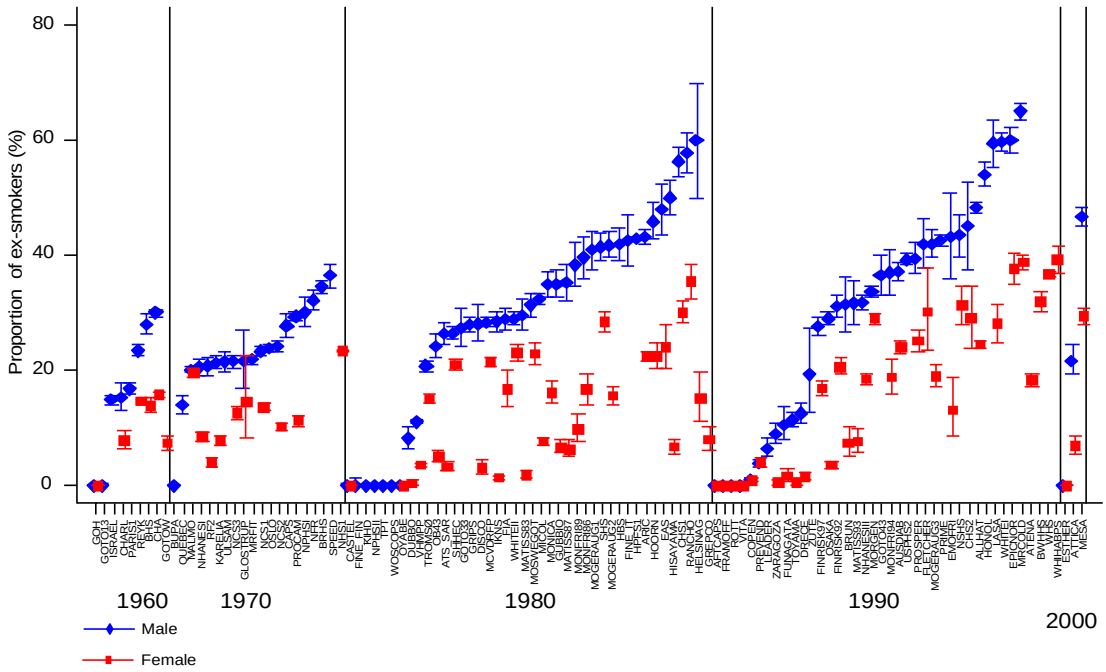
categorization may not be transferable across countries. The assessment of correlates was done using cross-sectional information, and investigation of these correlations using prospective data would be needed to confirm that they are causal. Despite the large scale of the dataset, information for some biomarkers, especially markers of inflammation, was only available in relatively small subsets of the data.

Figure 3.1: Proportion of current smokers in each cohort, by sex and by decade of baseline survey



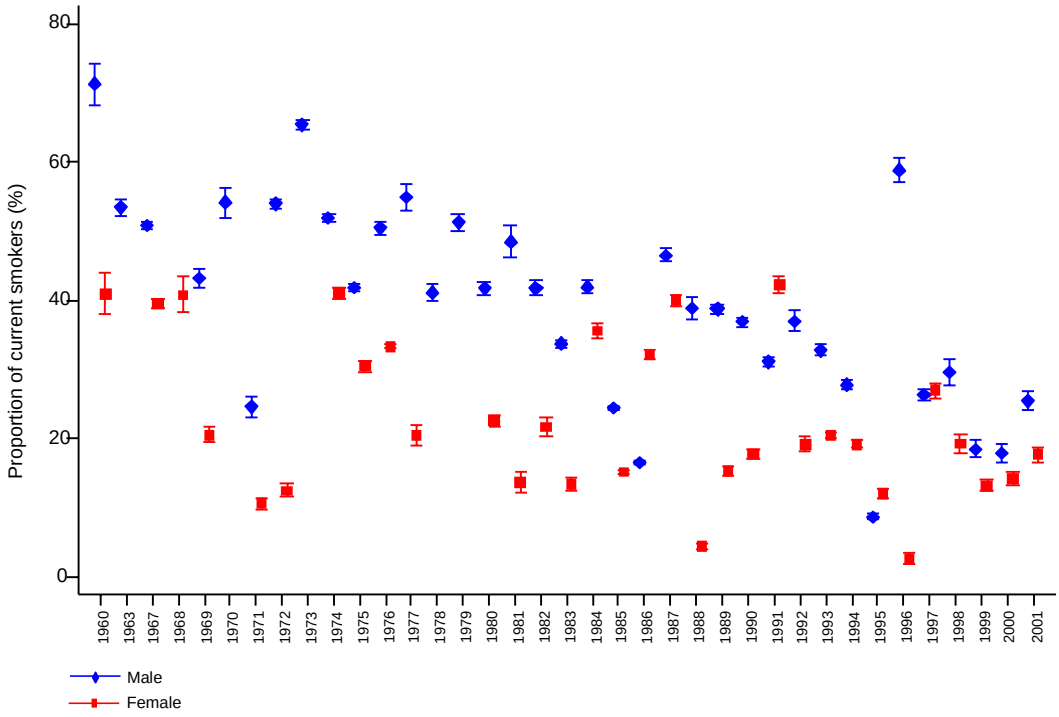
Cohorts are grouped by decade of baseline survey. Within each decade, cohorts are ordered by increasing proportion of current smokers. 95% Confidence intervals are derived assuming a binomial distribution and width is inversely proportional to the number of individuals in the study of a specific sex. A total of 114 cohorts contributed to this graph.

Figure 3.2: Proportion of ex-smokers in each cohort, by sex and by decade of baseline survey



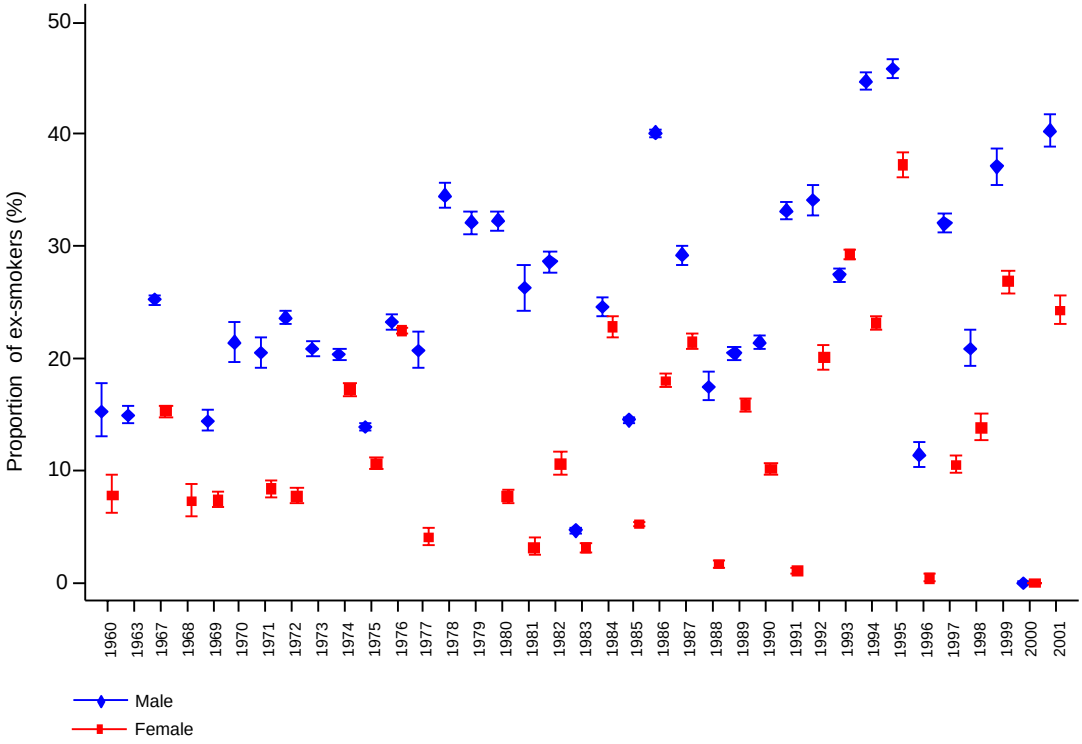
Cohorts were grouped by decade of baseline survey. Within each decade, cohorts are ordered by increasing proportion of ex-smokers. 95% Confidence intervals are derived assuming a binomial distribution and width is inversely proportional to the number of individuals in the study of a specific sex. A total of 114 cohorts were included in this graph.

Figure 3.3: Proportion of current smokers by baseline year of study and by sex



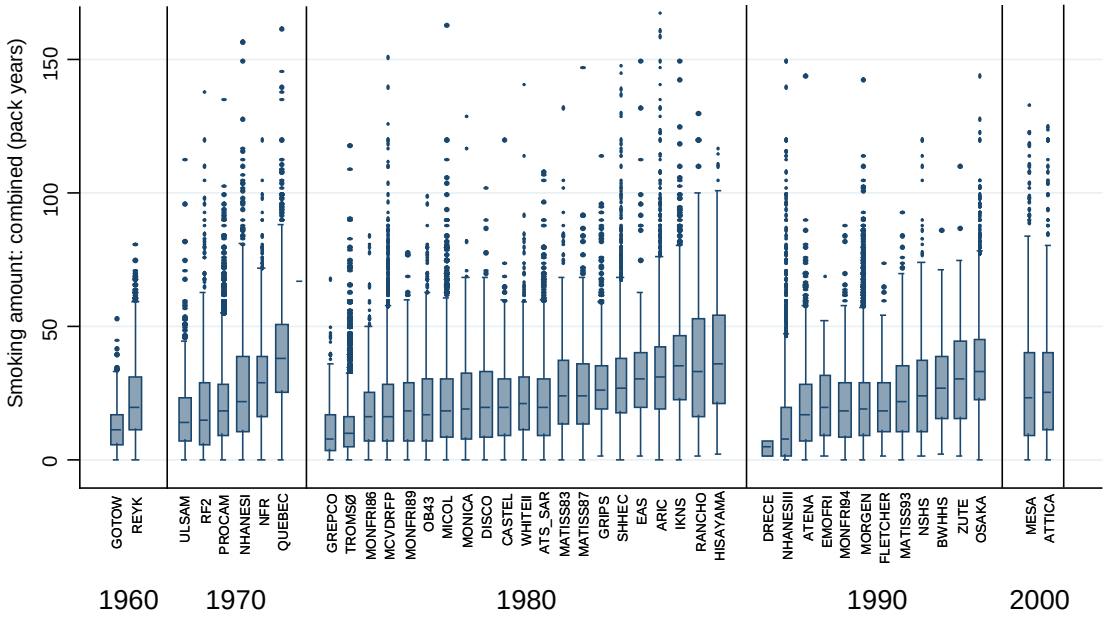
95% Confidence intervals are derived assuming a binomial distribution and width is inversely proportional to the number of current smokers recruited in a specific year all studies combined. A total of 114 cohorts were included in this graph.

Figure 3.4: Proportion of ex-smokers by baseline year of study and by sex



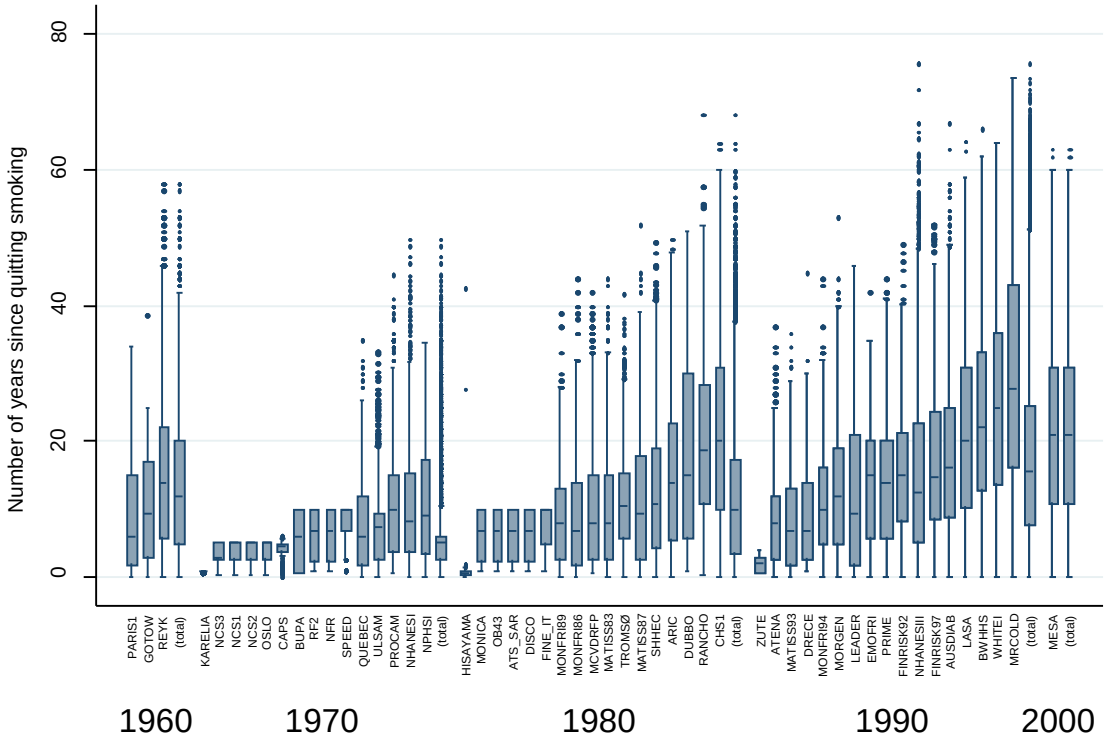
95% Confidence intervals are derived assuming a binomial distribution and width is inversely proportional to the number of ex smokers recruited in a specific year all studies combined. A total of 114 cohorts were included in this graph.

Figure 3.5: Box plots of pack-years in current smokers by studies and decade of baseline survey



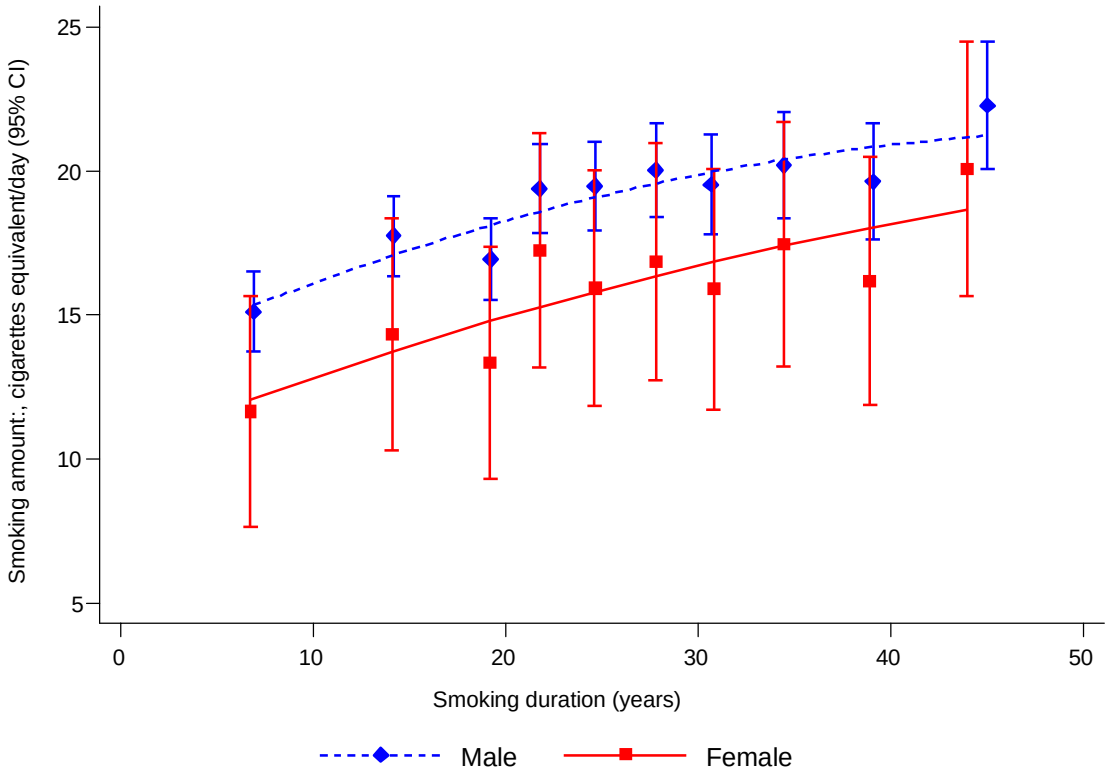
In total, 43 studies provided information on the number of pack-years. Cohorts are ordered by decade of baseline survey and by increasing median of pack-years. A total of 44 studies including 82.922 individuals were included in this graph.

Figure 3.6: Box plots of number of years since stopping smoking by studies and decade of baseline survey



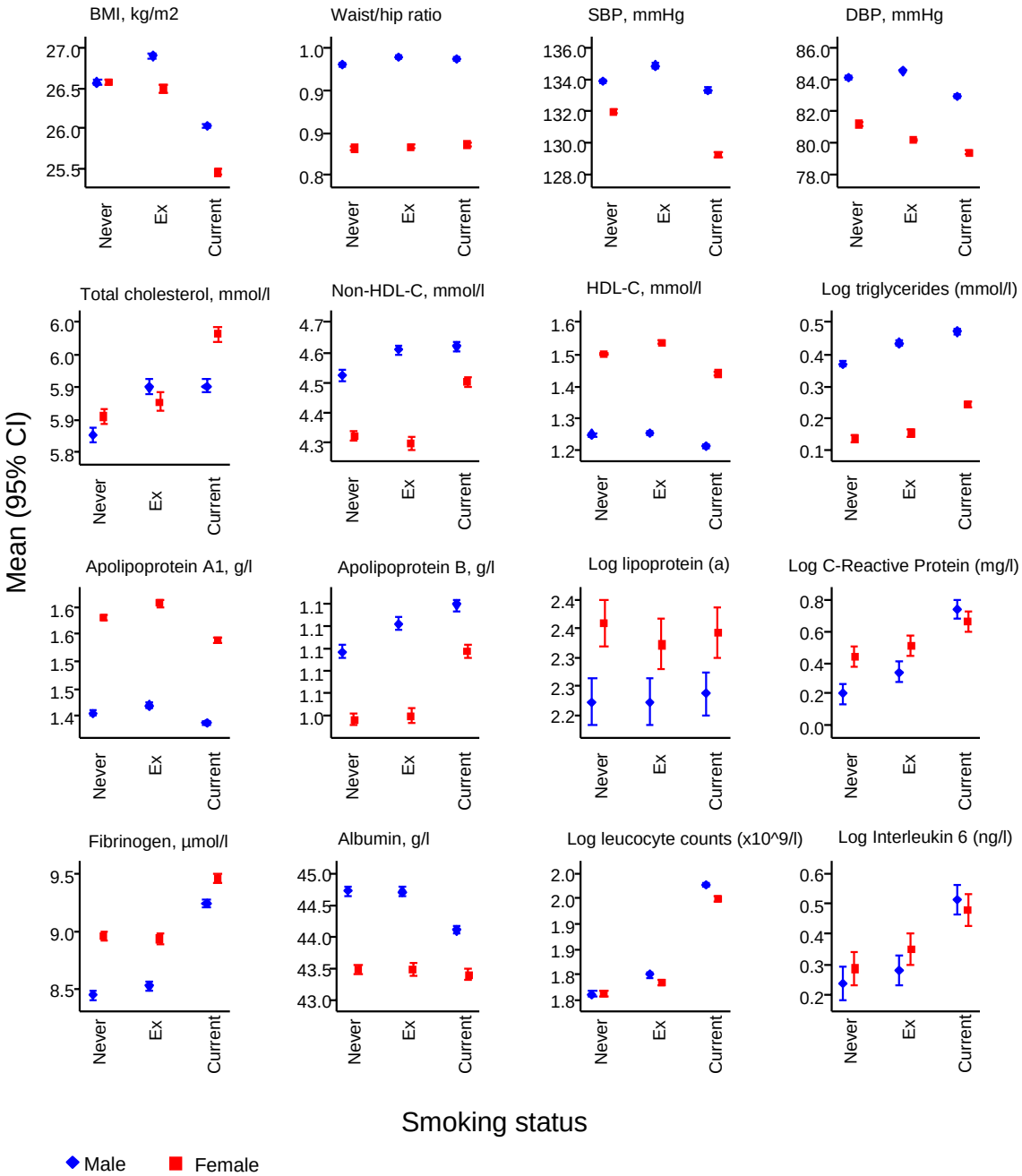
In total, 54 studies including 80,318 individuals provided information on number of years since stopping smoking. Cohort are grouped by decade of baseline survey, and within each decade, ordered by increasing number of years.

Figure 3.7: Smoking amount cigarette by sex and by duration of smoking



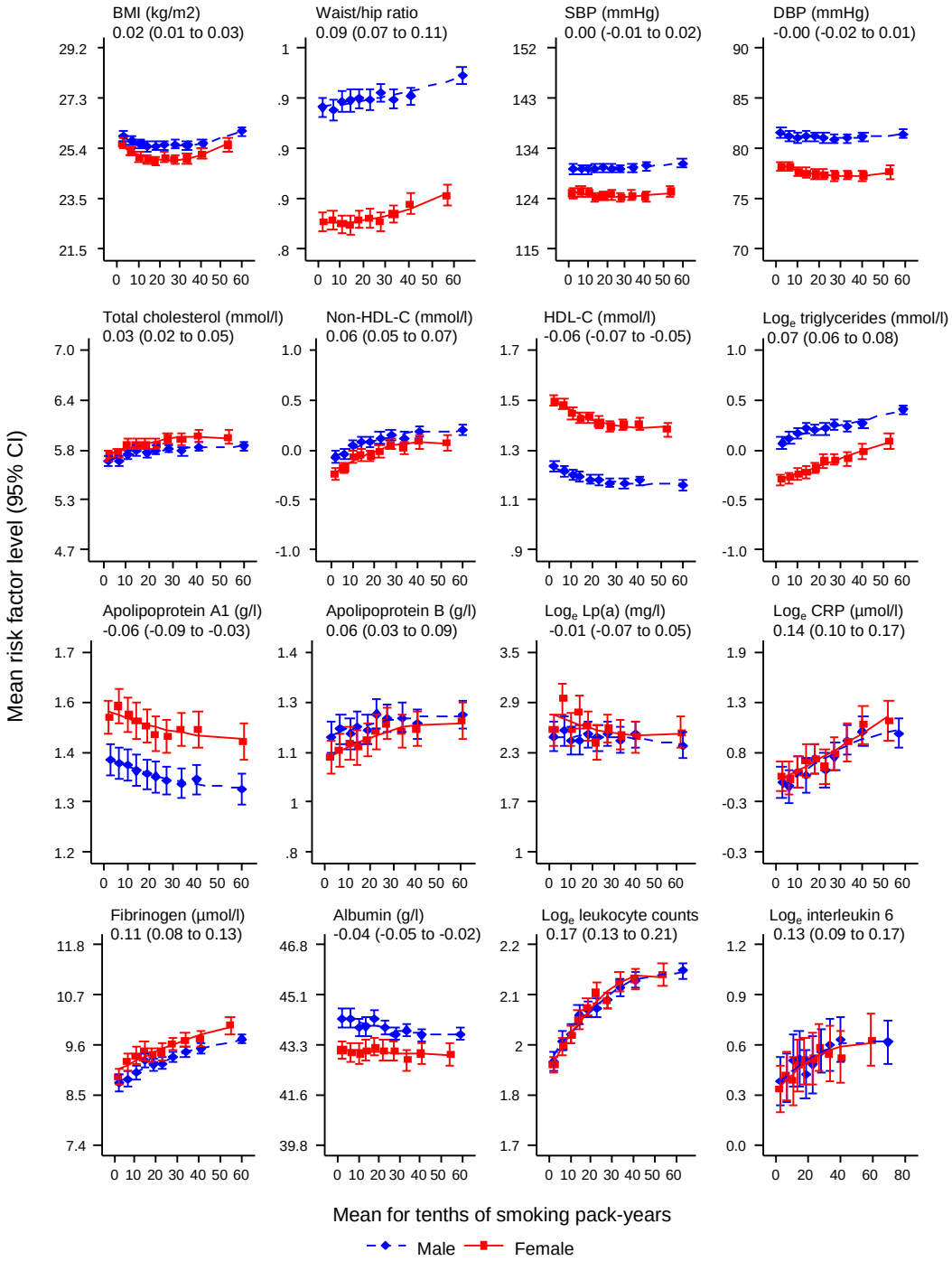
Mean baseline smoking amount of cigarette, pipes and cigar combined were plotted by sex against mean of each tenth of smoking duration amongst current smokers at baseline. Error bars represent 95% confidence intervals. Means and SD were adjusted for study and age 50 years old using mixed models (see **Methods** section). 44 studies and 82,922 current smokers at baseline provided information on both amount and duration at baseline.

Figure 3.8: Cross-sectional association between smoking status and continuous baseline characteristics



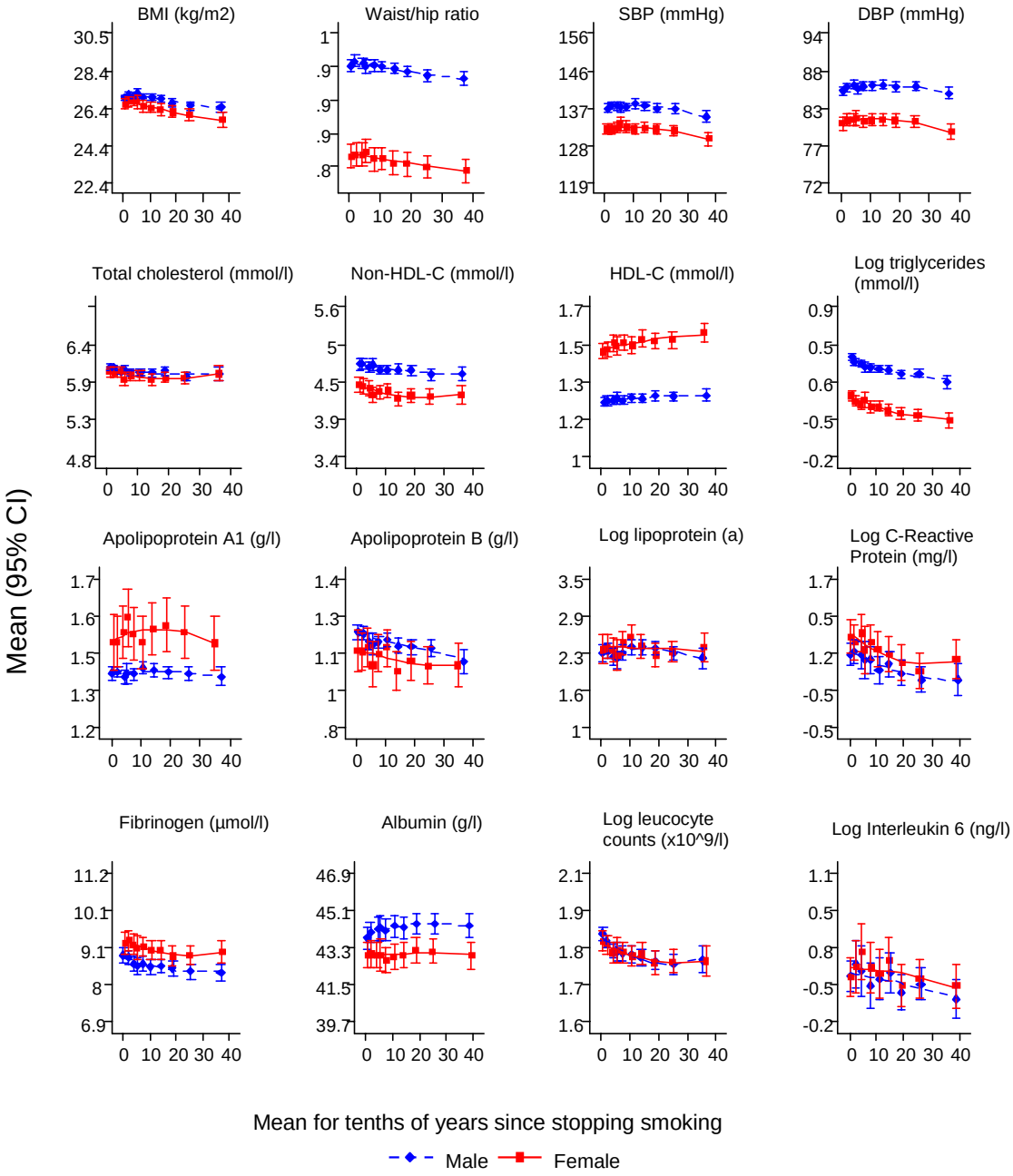
Mean risk factor levels were adjusted to age 50 years. Error bars represent 95% confidence intervals. Mean risk factors were estimated using a linear mixed model. The model was adjusted for cohort, *age*, *age*^{2}, *sex*, *age x sex*, *age*^{2} *x sex*, *smoking status x sex* and *smoking status x age* (where x denotes and interaction). Coefficients that were allowed to vary randomly across studies were *age, age*^{2}*, sex* and *smoking status.* The number of studies and number of individuals included in each graph can be found in the first two columns of **Table 2**.

Figure 3.9: Cross-sectional association between levels of smoking pack-years in current smokers and continuous baseline characteristics



Mean risk factor levels were adjusted for age 50 years. The values above each figure correspond to the age- and sex- adjusted partial correlation coefficient (95% CI) between risk factor and number of pack-years in current smokers, males and females combined. Error bars represent the 95% CIs. Mean risk factors were estimated using a linear mixed model. The model was adjusted for cohort, *age*, *age*^{2}, *sex*, *age x sex*, *age*^{2} *x sex*, *risk factor x sex* and *risk factor x age* (where x denotes and interaction). Coefficients that were allowed to vary randomly across studies were *age, age*^{2}, *sex* and *risk factor*. The number of studies and number of individuals included in each graph can be found in **Table 2**.

Figure 3.10: Cross-sectional association between number of years since quitting smoking and continuous baseline characteristics



Mean risk factor levels were adjusted for age 50 years. The values above each figure correspond to the age- and sex- adjusted partial correlation coefficient (95% CI) between risk factor and number of pack-years in current smokers, males and females combined. Error bars represent the 95% CIs. Mean risk factors were estimated using a linear mixed model. The model was adjusted for cohort, *age*, *age*^{2}, *sex*, *age x sex*, *age*^{2} *x sex*, *risk factor x sex* and *risk factor x age* (where x denotes and interaction). Coefficients that were allowed to vary randomly across studies were *age, age*^{2}*, sex* and *risk factor*.

Table 3.1: Summary of demographic and lifestyle covariates by smoking status at baseline

	Smoking status					Smoking pack-years in current smokers		
	No of studies	**No of subjects**	**Current smokers**	**Ex-smokers**	**Never smokers**	**No of studies**	**No of subjects**	**Difference (95% CI) in number of smoking pack-years compared to the reference category**
Age (years), mean (SD)	*114*	*929,335*	53.3 (8.5)	55 (8.6)	54.3 (10.2)	*44*	*82922*	4.51 (3.67 to 5.35)*
Sex, %	*114*	*929,335*				*44*	*82922*	
Male		*476,809*	41%	28%	32%		*50059*	1
Female		*452,526*	27%	18%	55%		*32863*	-5.18 (-17.31 to 6.96)
Race, %	*90*	*483,114*				*37*	*47099*	
White		*414,085*	38%	24%	38%		*36637*	1
Non white		*69,029*	33%	24%	43%		*10462*	-9.22 (-12.07 to -6.36)
Alcohol status, %	*86*	*432,223*				*37*	*69297*	
Not current		*153,086*	32%	21%	48%		*12960*	1
Current		*279,137*	38%	28%	34%		*56337*	-0.09 (-0.80 to 0.62)
History of diabetes, %	*103*	*749,098*				*39*	*76858*	
No		*714,983*	36%	26%	38%		*74509*	1
Yes		*34,115*	23%	36%	42%		*2349*	0.89 (-0.07 to 1.84)
Level of education reached, %	*59*	*337,521*				*26*	*57483*	
No schooling		*12,863*	29%	19%	52%		*3146*	1
Primary		*63,474*	36%	23%	41%		*12522*	0.86 (-0.70 to 2.42)
Secondary		*167,910*	41%	25%	34%		*31365*	0.66 (-0.46 to 1.77)
Vocational/University		*93,274*	30%	31%	39%		*10450*	-1.05 (-2.31 to 0.21)
Occupation or job, %	*51*	*327,152*				*25*	*36228*	
Not working		*81,454*	19%	14%	67%		*7835*	1
Manual		*84,173*	38%	17%	44%		*10034*	1.08 (-0.54 to 2.69)
Office		*9,401*	65%	19%	15%		*815*	-1.37 (-3.18 to 0.43)
Service		*120,953*	27%	21%	52%		*8776*	-0.65 (-3.15 to 1.84)
Student		*546*	33%	24%	43%		*62*	-2.74 (-7.24 to 1.77)
Other		*30,625*	35%	30%	35%		*8706*	-1.53 (-2.97 to -0.09)

Change in number of smoking pack-years compared to reference category - except for age where it is per 1 standard deviation higher levels of age -, adjusted for age and sex, pooled across studies using random effects meta-analysis.

Table 3.2: Summary of anthropometric covariates, blood pressure, lipids and inflammatory markers, by smoking status at baseline

	Correlation with smoking status					Correlation with smoking pack-years in current smokers		
	No of studies	***No of subjects***	**Mean (SD) or % in current smokers**	**Mean (SD) or % in ex-smokers**	**Mean (SD) or % in never smokers**	***No of studies***	***No of subjects***	**Difference (95% CI) in row variable per 1 SD increase in number of smoking pack-years**
Anthropometry								
Body Mass Index (kg/m2)	*114*	*929,335*	25.4 (3.8)	26.4 (4)	26.3 (4.2)	*44*	*82922*	0.07 (0.02 to 0.13)
Waist to Hip Ratio	*46*	*154,674*	0.888 (0.093)	0.905 (0.090)	0.866 (0.090)	*21*	*20203*	0.01 (0.00 to 0.01)
Blood pressure								
Systolic blood pressure	*111*	*754,077*	133.9 (18.4)	137.3 (18.3)	136.5 (19.5)	*44*	*80905*	0.10 (-0.18 to 0.39)
Diastolic blood pressure	*111*	*755,576*	80.9 (11.1)	82.8 (10.8)	82.2 (10.9)	*44*	*80879*	-0.04 (-0.21 to 0.13)
Lipid factors								
Total cholesterol (mmol/l)	*111*	*738,215*	5.8 (1.1)	5.9 (1.1)	5.8 (1.1)	*43*	*79994*	0.04 (0.02 to 0.06)
Non-HDL-C (mmol/l)	*93*	*381,742*	4.5 (1.2)	4.5 (1.1)	4.4 (1.1)	*39*	*63904*	0.07 (0.05 to 0.09)
HDL-C (mmol/l)	*93*	*382,079*	1.3 (0.4)	1.3 (0.4)	1.4 (0.4)	*39*	*63972*	-0.02 (-0.03 to -0.02)
Log_e triglyceride (mmol/l)	*93*	*574,463*	0.4 (0.5)	0.3 (0.5)	0.3 (0.5)	*37*	*59540*	0.04 (0.03 to 0.05)
Apo A1 (g/l)	*28*	*111,370*	1.4 (0.3)	1.5 (0.3)	1.5 (0.3)	*11*	*11667*	-0.02 (-0.03 to -0.01)
Apo B (g/l)	*28*	*113,902*	1.1 (0.3)	1.1 (0.3)	1.1 (0.3)	*11*	*11273*	0.02 (0.01 to 0.03)
Log_e Lp(a)	*29*	*90,111*	2.3 (1.3)	2.3 (1.2)	2.3 (1.2)	*10*	*8286*	-0.01 (-0.08 to 0.05)
Inflammatory markers								
Log_e CRP (mg/l)	*48*	*107,879*	0.8 (1.1)	0.6 (1.1)	0.4 (1.1)	*12*	*5471*	0.18 (0.11 to 0.24)
Fibrinogen (µmol/l)	*44*	*180,379*	9.6 (2.2)	9.1 (2.1)	8.9 (2)	*14*	*22768*	0.22 (0.16 to 0.28)
Albumin (g/l)	*35*	*122,447*	43.3 (3.5)	43.3 (3.6)	43.1 (3)	*13*	*16988*	-0.13 (-0.20 to -0.06)
Log_e leucocyte count(x10^9/l)	*33*	*109,974*	2 (0.3)	1.8 (0.3)	1.8 (0.3)	*14*	*13229*	0.05 (0.04 to 0.07)
Log_e Interleukin-6 (ng/l)	*10*	*16,205*	0.6 (0.6)	0.5 (0.7)	0.4 (0.6)	*6*	*2283*	0.08 (0.05 to 0.10)

Mean and SD for current, ex and never smokers were calculated within studies and pooled across studies using random effect meta-analysis. Change in row variable per 17 additional pack-years, adjusted for age and sex, and pooled across studies using random effect meta-analysis.

References

1. McBride PE. The health consequences of smoking. Cardiovascular diseases. *Med Clin North Am.* 1992;76:333-353.

2. Ambrose JA, Barua RS. The pathophysiology of cigarette smoking and cardiovascular disease: an update. *J Am Coll Cardiol.* 2004; 43:1731-1737.

3. Cullen P, Schulte H, Assmann G. Smoking, lipoproteins and coronary heart disease risk. Data from the Munster Heart Study (PROCAM). *Eur Heart J.* 1998;19:1632-1641.

4. Njolstad I, Arnesen E, Lund-Larsen PG. Smoking, serum lipids, blood pressure, and sex differences in myocardial infarction. A 12-year follow-up of the Finnmark Study. *Circulation.* 1996;93:450-456.

5. Smith CJ, Fischer TH. Particulate and vapor phase constituents of cigarette mainstream smoke and risk of myocardial infarction. *Atherosclerosis.* 2001;158:257-267.

6. Kannel WB, D'Agostino RB, Belanger AJ. Fibrinogen, cigarette smoking, and risk of cardiovascular disease: insights from the Framingham Study. *Am Heart J.* 1987;113:1006-1010.

7. Liu J, Liang Q, Frost-Pineda K et al. Relationship between Biomarkers of Cigarette Smoke Exposure and Biomarkers of Inflammation, Oxidative Stress, and Platelet Activation in Adult Cigarette Smokers. *Cancer Epidemiol Biomarkers Prev.* 2011;20:1760-1769.

8. Kaptoge S, White IR, Thompson SG et al. Associations of plasma fibrinogen levels with established cardiovascular disease risk factors, inflammatory markers, and other characteristics: individual participant meta-analysis of 154,211 adults in 31 prospective studies: the fibrinogen studies collaboration. *Am J Epidemiol.* 2007;166:867-879.

9. Gayen AK. The frequency distribution of the product-moment correlation coefficient in random samples of any size drawn from non-normal universes. *Biometrika.* 1951;38:219-247.

10. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ.* 2011;342.

11. Thun, M. J., Day-Lally, C., Myers, DG, Calle, EE, Flanders, WD, Zhu, B-P, Namboodiri, MM, and Heath, CW Jr. Trends in tobacco smoking and mortality from cigarette use in Cancer Prevention Studies I (1959 through 1965) and II (1982 through 1988). Changes in Cigarette-Related Disease Risks and Their Implication for Prevention and Control.Smoking and Tobacco Control Monograph No.8.Rockville (MD) [*a*], 305-382. 1997. Shopland DR, Burns DM, Garfinkel L, Samet JM.

12. World Health Organization. *The Tobacco Atlas*. 2002.

13. Pierce JP, Messer K, White MM, Cowling DW, Thomas DP. Prevalence of heavy smoking in California and the United States, 1965-2007. *JAMA.* 2011;305:1106-1112.

14. Kandel DB, Chen K. Extent of smoking and nicotine dependence in the United States: 1991-1993. *Nicotine Tob Res.* 2000;2:263-274.

15. Berrendero F, Robledo P, Trigo JM, Martin-Garcia E, Maldonado R. Neurobiological mechanisms involved in nicotine dependence and reward: participation of the endogenous opioid system. *Neurosci Biobehav Rev.* 2010;35:220-231.

16. Thun MJ, Apicella LF, Henley SJ. Smoking vs other risk factors as the cause of smoking-attributable deaths: confounding in the courtroom. *JAMA.* 2000;284:706-712.

17. Kenfield SA, Stampfer MJ, Rosner BA, Colditz GA. Smoking and smoking cessation in relation to mortality in women. *JAMA.* 2008;299:2037-2047.

18. Nettle D. Large Differences in Publicly Visible Health Behaviours across Two Neighbourhoods of the Same City. *PLoS One.* 2011;6:e21051.

19. Niederdeppe J, Kuang X, Crock B, Skelton A. Media campaigns to promote smoking cessation among socioeconomically disadvantaged populations: what do we know, what do we need to learn, and what should we do now? *Soc Sci Med.* 2008;67:1343-1355.

20. Nichter M, Nichter M, Carkoglu A, Lloyd-Richardson E. Smoking and drinking among college students: "it's a package deal". *Drug Alcohol Depend.* 2010;106:16-20.

21. Barrett-Connor E, Khaw KT. Cigarette smoking and increased central adiposity. *Ann Intern Med.* 1989;111:783-787.

22. Lissner L, Bengtsson C, Lapidus L, Bjorkelund C. Smoking initiation and cessation in relation to body fat distribution based on data from a study of Swedish women. *Am J Public Health.* 1992;82:273-275.

23. Akbartabartoori M, Lean ME, Hankey CR. Relationships between cigarette smoking, body size and body shape. *Int J Obes (Lond).* 2005;29:236-243.

24. Symons JD, Stebbins CL. Hemodynamic and regional blood flow responses to nicotine at rest and during exercise. *Med Sci Sports Exerc.* 1996;28:457-467.

25. Benowitz NL, Kuyt F, Jacob P, III. Influence of nicotine on cardiovascular and hormonal effects of cigarette smoking. *Clin Pharmacol Ther.* 1984;36:74-81.

26. Craig WY, Palomaki GE, Haddow JE. Cigarette smoking and serum lipid and lipoprotein concentrations: an analysis of published data. *BMJ.* 1989;298:784-788.

27. U.S.Department of Health and Human Services. *How tobacco smoke causes disease: the biology and behavioral basis of smoking-attributable disease: a report of the surgeon general*. 2010.

28. Bazzano LA, He J, Muntner P, Vupputuri S, Whelton PK. Relationship between cigarette smoking and novel risk factors for cardiovascular disease in the United States. *Ann Intern Med.* 2003;138:891-897.

29. Danesh J, Lewington S, Thompson SG et al. Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *JAMA.* 2005;294:1799-1809.

30. Smith FB, Lee AJ, Fowkes FG, Price JF, Rumley A, Lowe GD. Hemostatic factors as predictors of ischemic heart disease and stroke in the Edinburgh Artery Study. *Arterioscler Thromb Vasc Biol.* 1997;17:3321-3325.

31. Kaptoge S, Di AE, Lowe G et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *Lancet.* 2010;375:132-140.

32. Caraballo RS, Giovino GA, Pechacek TF et al. Racial and ethnic differences in serum cotinine levels of cigarette smokers: Third National Health and Nutrition Examination Survey, 1988-1991. *JAMA.* 1998;280:135-139.

Chapter 4: Cigarette smoking and the risk of cardiovascular diseases, lung cancer deaths and all-cause mortality in developed countries

Summary

Cigarette smoking is a well-known risk factor for cardiovascular diseases thanks to numerous prospective cohort studies, mainly conducted in the developed world. The present analyses used the Emerging Risk Factors Collaboration and aimed to synthesize the available evidence on the association between smoking and the risk of CVD in the developed world using appropriate adjustment for cardiovascular risk factors. In addition, they investigated areas of this association which remain controversial or based on insufficient evidence, such as the presence of a higher risk in female versus male current smokers; and the association of smoking with subtypes of CVD. They also contained novel analyses on potential interplays between smoking and other predominant risk factors with CVD risk; and an estimation of risk of life lost to smoking in women. Finally, these analyses looked at the effect of stopping smoking, in terms of the number of years necessary for an ex-smoker to reach the level of CVD risk of a never smoker. For comparison, associations with lung cancer deaths were also presented.

The ERFC had data with smoking status from 114 prospective cohort studies on 929,335 individuals, with an average follow up of 14.2 years; who experienced in total 46,576 cardiovascular deaths. After adjustment for age and body mass index, the risk ratios for current versus never smokers were for coronary heart diseases 1.99 (95% CI: 1.86, 2.13), and for cardiovascular deaths 2.01 (1.89, 2.14). All types of CVD, including pulmonary embolism, heart failure and cardiac dysrhythmia were associated with smoking. The effect of smoking was strongest on the risk of fatal aortic aneurysm (OR: 4.34; 3.33 to 5.66) and deaths from peripheral vascular diseases (OR: 3.81; 2.67 to 5.45). There was no evidence of confounding by other lifestyle and biochemical risk factors. Associations were stronger in younger age groups and in women, and for myocardial infarction, in non-diabetics. However, in absolute terms, more men and diabetics experienced a CVD event as the result of smoking than women and non-diabetics. Associations with numbers of amount and duration were curvilinear and converged toward a plateau above approximately 20 pack-years for the risk of CHD and all CVD. Cessation was rapidly beneficial, with an

80% decrease in CVD risk within the first 5 years of cessation. However, the risk only became non-significant after 20 years of cessation. Men who were current smokers and 50 years old lost an average of 7.5 years, and women of the same age lost 4.6 years due to smoking. These results confirm and enhance previous epidemiological evidence on the association between cigarette smoking and CVD.

4.1 Background

Smoking remains the leading preventable cause of disease, disability and death in developed countries, and a large majority of smokers are cigarette smokers. Considerable epidemiological evidence has been accumulated on the association between cigarette smoking and the risk of cardiovascular diseases and death [1]. Recently, a very large-scale analysis reported on the risk associated with smoking and the benefits of cessation using a database which included 1.3 million women recruited between 1996 and 2001 and follow-up for up to 8 years [2]. However, some areas have not been investigated in enough detail and others remain controversial.

Firstly, some types of CVD are rarer than others and studies have often been underpowered to investigate the effect of smoking on these types, namely subarachnoid haemorrhage, aortic aneurysm, heart failure, cardiac dysrhythmia, pulmonary embolism and peripheral vascular disease. Current evidence is mostly derived from literature based meta-analysis with relatively few cases [3-5], which sometimes also include retrospective studies subject to recall bias [6], and are generally unable to consistently adjust for other risk factors of CVD across substantially heterogeneous studies [7].

A second area of uncertainty relates to the interplay between cigarette smoking and other well known risk factors of CVD. At a time when the prevalence of glycaemia and diabetes is rising because of general ageing of the population [8], obesity is spreading as a result of poor diet and low physical activity [9], and total cholesterol remains highest worldwide in high income countries [10], it is important to assess the joint effect with smoking of these well known risk factors of CVD [11-13]. The presence of an effect modification by sex also remains controversial. It has been suggested by a recent meta-analysis that women experience an increase in risk 25% higher when smoking than men [14]. However, this meta-analysis was literature based, adjustment was inconsistent across studies included, and when pooling the age-only adjusted RR ratios, the effect modification by sex was non-significant.

Thirdly, the shape of association between smoking dose, smoking duration and the risk of CVD is still controversial. A plateau effect at high smoking intensities has been found in some but not all studies [15]. The effect of duration of smoking has been difficult to assess as it is nearly collinear with age, most smokers starting in their teenage years [16]. The common use of smoking pack-years to quantify total exposure

to smoking makes the assumption that an increase of 1 cigarette per day is equivalent to 1 additional year of smoking and, whilst it may be appropriate as a crude measure of total exposure to smoking and when assessing the risk of lung cancer [17, 18], its relevance when assessing CVD risk has yet to be demonstrated.

Fourthly, the rapidity of the benefits obtained by cessation is uncertain. Some studies have estimated a halving in CHD risk within the first 3 to 5 years [19, 20] whilst others have observed a slower decline [21]. The estimated time for levels of CVD risk to come back to that of a never smoker have also oscillated between 10 years [20] and more than 20 years [22]. In the ERFC, information was available in 80,318 individuals from 54 studies on number of years since quitting smoking, enabling more reliable estimates of risk reduction than previously possible.

For continuing smokers, quantifying the reduction in life expectancy as a result of smoking relies on precise estimates of its association with disease and death, and in particular with CVD. Published estimates of an average of 10 years of life lost as a result of smoking were based on a study of British Male doctors born before 1930, with lifelong smoking habits likely to differ from more recent generations [23]. These estimates were also based on men only. Nowadays, smoking rates in women are close to or have reached the same rates as men in some Western countries [24]. As women tend to live longer than men and have been conversely shown to experience greater increase in risk when smoking [14], the question of years of life lost to smoking needs to be addressed in women. A recent study estimated a similar number of years of life lost for women compared to men but this finding has not yet been replicated [2].

The objective of this Chapter is to address these uncertainties and provide better estimates of risk of CVD than previously possible using data from up to 114 prospective cohort studies including nearly 1 million participants. The main focus of this Chapter is on CVD, but some sections present results on lung cancer deaths for comparison. In addition, the effect of smoking on life expectancy is considered, and the respective share of CVD and cancer in tobacco related burden is assessed using ERFC findings.

4.2 Methods

4.2.1 The dataset

The subset of the Emerging Risk Factors Collaboration (ERFC) dataset with information on smoking has been described in **Chapter 2**. There were 114 studies with information on smoking status which including 929,335 individuals, 103 studies with information on smoking pack-years (defined as the number of packs of 20 cigarettes per day times the number of years), 84 studies for amount, 44 studies for duration, 44 studies for starting age and 54 studies for stopping age, also expressed as number of years since quitting smoking. Three studies conducted in developing countries were excluded from the analysis. Harmonization of the coding for smoking status into 3 categories, current, ex- and never smokers, was done by the data management team, in correspondence with the individual study investigators. This analysis was done, combining all type of smoking, because some studies did not provide information on type of smoking product used, and in order to maximize power. However, 72% of individuals had information on cigarette smoking status and information was mostly concordant: over 90% of current smokers of any type were current cigarette smokers and over 98% of ex-smokers of any type were ex-cigarette-smokers. When smoking status was provided by the studies with no mention of the type of smoking, cigarette smoking was assumed.

4.2.2 Statistical methods

Details of the statistical methods have been reported previously [25]. Following the example of previous reports published by the ERFC [26-31], I assessed associations of smoking status, pack-years, type, cessation and CVD. Participants were censored at the first non-fatal MI or stroke and at any cause of death. Sensitivity analyses excluding studies which recorded only fatal CVD events were conducted when comparing the effect for fatal versus non-fatal MI or stroke. To avoid reverse causality, sensitivity analyses were done excluding the first 5 years of follow-up.

Relative risks (RRs) were estimated using a 2-stage approach. In the first stage, RRs were calculated separately for each study. For prospective cohort studies and trials, a Cox proportional hazards regression model was fitted to estimate the hazard ratio for each study, stratifying, where appropriate, by sex and trial arm [32]. Separately for

each study *s=1 …S*, with strata *k=1 …K*$_s$ (for most studies *K*$_s$=2, just for the two sexes) and individuals *i=1 …n*$_s$, with exposure of interest *E*$_{si}$ and other covariate *X*$_{si}$, the hazard at time *t* after baseline was modeled as:

$$\log(h_{ski}(t \mid E_{si}, X_{si})) = \log h_{0sk}(t) + \beta_s E_{si} + \gamma_s X_{si} \qquad (1)$$

The β_s is the parameter of interest, and corresponds to the log of the hazard ratios per unit increase in continuous variables for study *s*, or for one category versus the reference group in the case of categorical variables. For smoking status, never smokers were chosen as the reference group. To check whether the assumption of proportional hazards held, a time-dependent interaction term was added to the model for each study and its pooled estimate was tested for significance. For nested case-control studies, odds ratios were calculated fitting a logistic regression rather than a Cox model. Odds ratios were assumed to approximate hazard ratios in that case, which is reasonable because cardiovascular disease incidence can be considered as rare [33]. To characterize shapes of associations for continuous variables such as pack-years, years since quitting and age stopped, RRs were computed by quintiles or predefined categories of the distribution. To investigate effect modification, interaction terms were added to the model. For the sex (or race) interactions, the dataset was reduced to studies recording both Males and Female (both White and non-White), in order to avoid spurious differences in hazard ratios due to between-studies heterogeneity rather than within-studies heterogeneity.

The second stage involved a multivariate random-effects meta-analysis of study-specific hazard ratios or odds ratios, to obtain an overall “risk ratio”. Compared to fixed effects, a random effects meta-analysis incorporates heterogeneity between studies [25, 34]. A fixed effects model assumes that a single parameter RR is common to all studies, while a random effects model makes the assumption that the underlying RR follows a random distribution. The latter is more appropriate for studies such as the ones involved in the ERFC, which differ considerably in design and beyond measurable covariates included in the model (for example, decade of recruitment, country of recruitment, study design, etc.), and are therefore likely to truly vary in terms of underlying RR. Writing the variance of the estimated β_s as *v*$_s$, the random-effects meta-analysis model is given by:

$$\begin{aligned} \hat{\beta}_s &= \beta_s + \varepsilon_s, \quad \text{where } \varepsilon_s \sim N(0, v_s) \\ \beta_s &= \beta + \eta_s, \quad \text{where } \eta_s \sim N(0, \tau^2) \end{aligned} \qquad (2)$$

where β represents the average log HR. The true heterogeneity between studies, beyond variation in the estimates due to chance, is represented by the variance τ^2, and the percentage of variance in the estimates that is attributable to between-study variation as opposed to sampling variation is expressed by I^2 statistic [35]. An I^2 close to 0% suggests that variability in study estimates is entirely due to chance, and an I^2 >50% suggests substantial heterogeneity between studies. RR pooled by random effect meta-analysis provides an estimate of the average or, in other terms, expected RR of a study (with allowance for random noise around this average), rather than the common RR across studies which would be provided assuming a fixed effects model. Specific sources of heterogeneity were explored by fitting interactions with study-specific risk factors such as decade of enrollment and region of the world (Western Europe, North America or other).

Another aspect of the regressions and meta-analyses performed was that they were both multivariate. Compared to univariate models, multivariate models take into account the correlation between the exposure of interest and covariates. Multivariate meta-analysis is a meta-analysis of β of main effect and coefficients of covariates in study-specific models at the same time [36]. Multivariate meta-analyses are well designed for my analysis, because I can assume that the effect of current versus never smoking and the effect of ex versus never smokers will be correlated. They are also useful for shape analyses, as coefficients for different quintiles are likely to be correlated; and when testing for interaction, as interactions and main effects are likely to be correlated. To reduce excessive computation time, the method of moment was used to estimate pooled RRs by multivariate meta-analysis, rather than maximum likelihood methods [36].

To assess shapes, pooled estimates within quintiles of continuous variables were plotted against the pooled mean of the quintile, obtained by random effect univariate meta-analysis of study-specific means. For example, for number of years since quitting smoking in relation to CVD risk, quintiles were defined using individuals who declared being ex-smokers at baseline and provided information on number of years since stopping smoking. For the association with number of pack-years smoking, quintiles were defined using baseline pack-years information available for current smokers.

To enable graphical comparison of RRs between any two groups, and not only with the reference group, 95% confidence intervals for the RRs were represented using "floating absolute" variances, and the size of the box representing RRs was chosen to be proportional to the inverse of the '"floating variance" [37]. RRs were at least adjusted for age and body mass index. To explore potential biological pathways' underlying associations, hazard ratios for smoking status were further adjusted for systolic blood pressure, history of diabetes, BMI, waist circumference, waist-to-hip ratio, total and high density lipoprotein cholesterol, triglyceride, C-reactive protein, fibrinogen, alcohol consumption, or socioeconomic indicators, i.e. educational attainment and occupational category. The association with number of years since stopping smoking was further adjusted for age starting smoking and past number of pack-years. For the purpose of investigating interactions, RRs were adjusted for age, body mass index, history of diabetes, systolic blood pressure and total cholesterol, and the dataset was reduced to individuals with information on all these covariates.

To estimate absolute rather than relative risks, event rates were estimated in the ERFC dataset by bands of 5 years of age at risk (40-45, 45-50, 50-55, 55-60, 60-65, 65-70, 70-75, 75-80, 80-85, 85+) and by sex, for current and never smokers separately, on the model of the Prospective Studies Collaboration [38]. Age at risk is defined as age of occurrence of an event and requires the data to be split into 5 year age bands so that individuals contribute to follow-up until they are censored or until they experience a first CVD event or death. Practically, event rates at 40 years or older were calculated by applying RR adjusted for age, BMI, history of diabetes, total cholesterol and systolic blood pressure, and stratified, where appropriate, by sex and trial arm, to the event rate in the ERFC in current smokers. For Body Mass Index, RRs were only adjusted for age at baseline. Never smokers were chosen as the reference group, except for sex, where current smokers were chosen as the reference group. RRs were based on 21,946 fatal and non-fatal MI (15,759 events for the graph showing the interaction with sex, selecting only studies including both Males and Females) and 10,243 cerebrovascular events (7,732 events for the graph showing the interaction with sex). The age at risk and sex standardized event rate for fatal and non-fatal MI was 0.00216 in never smokers and 0.004371 for current smokers in the ERFC dataset. The age at risk and sex standardized event rate for fatal and non-fatal cerebrovascular events was 0.001204 for never smokers and 0.00201 for current smokers.

Age at risk and sex specific rates were combined into an overall rate using calibrations provided by the European standard population (a notional population of 2 million people commonly used to standardise rates) [39]. In a third step, these rates were applied to estimated RRs.

I followed the example of a previous ERFC report on diabetes to estimate survival curves for current, ex- and never smokers and corresponding years of life lost [30]. Briefly, estimates of cumulative survival from 35 years of age and older among current, ex and never smokers at baseline were calculated by applying hazard ratios, specific to age at risk and sex, for cause-specific mortality associated with smoking status, to cause-specific rates of death at 35 years of age and older, for residents of the European Union. Analyses were carried out in Stata release 11 (StataCorp).

4.3 Results

Description of smoking variables and correlations with other baseline characteristics has been presented in **Chapter 3**. Information was available in up to 114 prospective cohort studies including 929,335 individuals; who experienced 40,218 incident CHD events including 16,390 non-fatal MI and 11,585 coronary deaths; 17,445 strokes including 8,866 non-fatal and 8,579 fatal strokes; and 11,511 other vascular events, during an average follow-up time of 14 years.

4.3.1 Current versus never smoking and the risk of CVD

Current smokers experienced roughly a doubling in risk of fatal and non-fatal CVD. RRs adjusted for age, BMI, sex and study, were 2.01 (1.89, 2.14) for all cardiovascular deaths and 1.99 (1.86, 2.13) for fatal and non-fatal MI, and 1.64 (1.54; 1.75) for all cerebrovascular events, compared to never smokers (**Figure 4.1**). Restricting the data to studies which recorded both fatal and non-fatal MI, RR was 1.89 (1.74; 2.06) for non-fatal MI and 2.10 (1.84; 2.39) for fatal MI (**Table 4.1**). Subsidiary analyses excluding the first 5 years of follow-up, and censoring at death rather than at first cardiovascular event or death only, did not materially impact the RRs. RRs for ischaemic stroke and haemorrhagic stroke were nearly the same (1.72; 1.54 to 1.93 versus 1.64; 1.45 to 1.85), while RR for subarachnoid haemorrhage was stronger: 2.69 (2.26; 3.20). Amongst other vascular diseases, the associations for heart failure, sudden death, cardiac dysrhythmia and pulmonary embolism were all significant with increases in risk comprising between 60% and 130%. The strongest associations were observed for aortic aneurysm (RR: 4.34; 3.33-5.66), and

peripheral vascular deaths (RR: 3.81; 2.67-5.45). Further adjustment for blood pressure, history of diabetes, alcohol consumption, levels of education and occupation, as well as lipids and inflammatory markers did not appreciably alter these associations (**Table 4.2**).

There was substantial heterogeneity between studies for the associations with fatal and with non-fatal MI. Values of I^2 were above 50%, indicating that most of the variability between studies could not be explained by chance. Looking at study specific determinants, I observed a small variation of RRs across regions of location of the studies and according to decade of start of enrolment of the study (p-values of interaction < 0.0001) (**Figures 4.2 & 4.3**). Pooled RRs were 2.02 (1.90; 2.19) in Western European countries, 1.75 (1.53; 2.00) in North America, and 1.52 (1.41; 1.65) in "other countries" which grouped studies from Australia, New Zealand, Israel and Japan. The association was 1.63 (1.44; 1.85) for studies started in the 1960s while it was above 2.01 (1.78; 2.27) for studies started in the 1970s. It was 1.95 (1.79; 2.13) for studies started in the 1980s and 1.91 (1.66; 2.20) for studies started in the 1990s.

The proportional hazard assumption was tested for departure from proportionality. P-values were statistically significant for fatal and non-fatal MI and all CVD. However, for these outcomes, RRs decreased only very slightly with the number of years of follow-up. For fatal and non-fatal MI, the difference in the RR for current versus never smokers was -0.10 (-0.13; -0.08) and for all CVD it was -0.09 (-0.12; -0.06) (**Table 4.3**).

Because of the relatively low number of events, it was not possible to investigate the effect of smoking other than that smoking status presented above on the risk of subtypes of CVD other than CHD and all cerebrovascular events.

4.3.2 Interplay of smoking with other risk factors on CVD risk

Individuals with a lower absolute risk, such as younger age groups, women, non-diabetics and individuals with low blood pressure levels were most harmed by smoking, in terms of CVD risk (**Figure 4.4 & 4.5**). Individuals ≥60 years old experienced a 27% (95%CI: 21%; 33%) lower RR of CHD for current versus never smokers, compared to individuals <60 years old. The difference in pooled RR between individuals aged ≥60 years versus <60 years old was 17% (95% CI: 9%;

25%) for all cerebrovascular events. Regarding sex, women had a stronger increase in risk when smoking compared to men by approximately 13%. However, the interaction was only borderline significant (95% CI for CHD: 4%-21% and for stroke: 2%-21%). For diabetes, an effect modification of smoking status according to history of diabetes was observed for CHD, with a 30% (21%; 38%) reduction in RR for diabetics compared to non-diabetics; but not for cerebrovascular events. There was no effect modification according to BMI or total cholesterol. The interaction with blood pressure for CHD death and non-fatal MI was statistically significant (P-value<0.001) but modification was marginal: per SD increase in SBP, smoking relation to CVD risk decreased by less than 10%. Subsidiary analyses where the models testing for effect modification with sex, diabetes, BMI, SBP and total cholesterol were adjusted for an additional interaction term between age and smoking in order to take into account potential confounding by an interaction with age yielded the same results.

The age at risk and sex standardized rate of fatal and non-fatal MI in the ERFC was 2.16/1,000 for never smokers and 4.37/1,000 for current smokers. For fatal and non-fatal cerebrovascular events, rates were 1.20/1,000 for never smokers and 2.01/1,000 for current smokers. Using these rates and applying them to RRs for current versus never smokers, a significantly greater number of CHD events were caused by smoking amongst individuals with a higher absolute risk such as older age groups, men, diabetics, and individuals in the top third of BMI, SBP and total cholesterol (**Figure 4.6a**). Individuals with higher baseline risk (such as individuals in the top third of SBP levels and diabetics) also experienced a greater increase in the number of cerebrovascular events when smoking than individuals with lower baseline risk (**Figure 4.6b**).

4.3.3 Dose-response relationship

Relationships between smoking intensity, duration and pack-years and the risk of CHD were all non-linear (**Figure 4.7**). Individuals smoking 15-20 CPD had a RR of 2.50 (2.34; 2.66), 20-25 CPD a RR of 2.54 (2.30; 2.81) and >25 CPD a RR of 2.93 (2.63; 3.26), compared to never smokers. Even smoking at low intensities was significantly associated with an approximate doubling in risk of CHD compared to never smokers and was significantly higher than the risk of individuals who stopped smoking: individuals smoking <10 CPD had a RR for CHD of 1.91(1.76; 2.08), and ex-smokers had a RR of 1.20 (1.09; 1.32) when comparing to never smokers. The association with duration was as follows: Individuals who had smoked <17 years had

a risk of 1.89 (1.68; 2.13), and individuals who had smoked 23-30 years experienced a risk of 2.60 (2.40; 2.80), with reference to never smokers. The shape of association for pack-years was approximately linear up to 25 pack-years, with a RR of 2.54 (2.33; 2.77) in individuals reporting 17-25 pack-years and then a tailoring above 25 pack-years, with reference to never smokers.

For cerebrovascular events, there was evidence of a dose-response relationship between smoking intensity and cerebrovascular events, but somewhat less strong than for CHD. Individuals smoking less than 10 CPD had a RR of 1.34 (1.09; 1.65) for ischaemic stroke and 1.61 (1.20; 2.15) for haemorrhagic stroke, compared to never smokers, and these were significantly greater than the risks for ex-smokers (0.93; 0.76 to 1.16 for ischaemic stroke and 1.16; 0.83 to 1.62 for haemorrhagic stroke). For 15-20 CPD, RRs were 2.00 (1.68; 2.36) for ischaemic stroke and 1.56 (1.14; 2.17) for haemorrhagic stroke. For 20-25 CPD, they were 2.12 (1.53; 2.93) for ischaemic stroke and 2.42 (1.56; 3.78) for haemorrhagic stroke and for >25 CPD 1.65 (1.24; 2.19) and 2.45 (1.60; 3.74) respectively. The association with duration was approximately flat, with RRs oscillating around 1.8 for ischaemic stroke and 2 for haemorrhagic stroke. Regarding pack-years, a positive and approximately linear association was apparent for ischaemic strokes and all cerebrovascular events combined, apart from the top quintile where RRs were slightly lower compared to the top 4th quintile.

4.3.4 Smoking cessation and CVD risk

Stopping smoking translated into clear benefits in terms of CVD risk when compared to current smokers. Ex-smokers experienced an excess risk for CVD which varied depending on the type of CVD and ranged from 4% to 44% (**Figure 4.1**). RRs were 1.13 (1.10; 1.17) for CHD, 1.17 (1.03; 1.33) for death from heart failure; 1.44 (1.17; 1.77) for death from aortic aneurysm; 1.04 (099; 1.09) for all cerebrovascular events; 1.04 (0.89; 1.22) for sudden death; 1.10 (0.97; 1.47) for deaths from cardiac dysrhythmia; and 1.39 (0.91; 2.13) for deaths from peripheral vascular diseases. Further adjustment for blood pressure, history of diabetes, lifestyle risk factors, lipids and inflammatory markers did not appreciably alter these associations for CHD (**Table 4.4**).

Cessation rapidly reduced CHD risk, from 1.95 (1.78; 2.13) in current smokers to 1.22 (1.14; 1.31) for ex-smokers who had stopped ≤3 years before enrolment in the

study, and became non-significant only 20 years after cessation with a RR of 1.00 (0.93; 1.05) (**Figure 4.8**, please note that RRs were plotted against the median number of years since stopping smoking, which for the category of ≥20 years was approximately 25 years before entry into the baseline). Individuals who had stopped 4-9 years before enrolment had a pooled RR of 1.21 (1.11; 1.31) and individuals who had stopped 10-19 years before enrolment a RR of 1.18 (1.10; 1.26). For the risk of all CVD deaths combined, RRs were 1.87 (1.74; 2.01) for current smokers, 1.23 (1.15; 1.32) for individuals who had stopped ≤3 years before enrolment, 1.15 (1.07; 1.23) for those who had stopped 4-9 years before enrolment, 1.08 (1.03; 1.14) for those who had stopped 10-19 years before enrolment and 0.95 (0.91; 1.00) for those who had stopped 20 years and more before enrolment. The decrease in risk for lung cancer deaths was more gradual and past smokers remained subjected to an excess risk even 20 years after cessations. RRs were 12.78 (10.98; 14.86) for current smokers, and respectively 8.99 (7.23; 11.18), 5.96 (5.03; 7.07), 4.21 (3.53; 5.03); 2.01 (1.64; 2.44) for individuals who had stopped 0-3, 4-9, 10-19 and 20+ years before entry into the study. Further adjustment for age starting smoking and past number of pack-years slightly attenuated the associations (**Table 4.5**).

4.3.5 Number of years of life lost attributable to smoking

Continuous smoking translated into a substantial loss in terms of years of life, for both men and women. RRs for all causes of deaths were 2.02 (1.96; 2.08) in men and 1.88 (1.82; 1.93) in women for current versus never smokers, with important heterogeneity between studies for both sexes (I^2 equal to 70% for men and 65% for women). Amongst individuals who were resurveyed, the proportion of never smokers who became current smokers during follow-up was below 1% even after 40 years follow-up, the proportion of individuals who were current smokers and became ex-smokers was <10% until 15 years of follow-up where it became >10%, and reached 28% in individuals who had been resurveyed 30 to 40 years after baseline entry (**Table 4.6**). A small percentage (mostly ≤5%) of individuals had started and stopped smoking at the time of the next resurvey, which was indicated by a switch in their status from never to ex-smokers.

At baseline age of 40, 50, 60 and 70 years old, men who were also current smokers lost respectively 7.7 years, 7.5 years, 6 years and 4.6 years of life compared to never smokers. The corresponding years of life lost for women smokers versus never smokers at these ages were 5.2 years, 4.6 years, 4.2 years and 3.1 years. Median

survival was 70 years old for male current smokers whilst it was 77 years old for male current smokers (**Figure 4.9a**). For women, median survival was 76 years old amongst current smokers whilst it was 81 years old amongst never smokers. At baseline age 50, in men, 28.5% of all deaths were attributable to CVD, 40.9% to cancer deaths and 28.7% to deaths which were not vascular or cancer related and the rest to unknown causes (**Figure 4.9b**). In women aged 50, vascular deaths represented 30.0%, cancer deaths 35.7% and non-vascular non-cancer deaths 32.2%, the rest being from unknown causes. The proportion of deaths attributable to CVD decreased with ageing while the proportion attributed to cancer deaths remained similar and the proportion attributed to non-vascular non-cancer causes increased, representing 51.7% at age 90 (31.1% for cancer deaths and 16.7% for CVD deaths).

For ex- versus never smokers, the difference was lower but still represented respectively 1.6 years, 1.5 years, 1.4 years and 1.1 years at ages 40, 50, 60 and 70 years old in men. In women with age 40, 50, 60 and 70 at baseline, the loss in number of years of life was respectively 0.7, 0.65, 0.60 and 0.46. Median survival for ex-smokers was 74 years old for men and 81 years for women (**Figure 4.10a**). A majority of the excess deaths of ex- versus never smokers was attributed to causes other than CVD. At age 50, in men, 23.7% of deaths amongst ex-smokers were due to CVD, 51.7% to cancer and 23.7% to non CVD non cancer deaths. In women at the same age, the proportions were 14% for CVD, 6.06% for cancers and 25.3% for non-CVD non-cancer deaths. These proportions remained similar in different age groups (**Figure 4.10b**).

4.4 Discussion

I studied smoking in relation to several major vascular and non-vascular outcomes in approximately 1 million adults from developed countries. The present analysis distinguishes itself by the large number of events and relatively long follow-up, with more than 70,000 cardiovascular events occurring during 15 million person years follow-up [40]. The ERFC censored individuals at their first non-fatal MI and stroke event as well as CVD mortality [20, 41]. My dataset included more than three times as many women participants as the largest study on the effect of smoking in women [20]. Contrasting with most published studies which have been questionnaire-based [20, 41, 42]; information was available on several blood based factors relevant to CVD such as blood pressure, lipids and inflammatory markers, enabling appropriate adjustment

and tests of interaction. The advantage of an individual participant meta-analysis like the ERFC, compared to literature-based meta-analyses, is that it allows consistent adjustment across studies and investigation of the sources of heterogeneous effects across studies. Bias from misreporting of smoking status in sick individuals was minimized by involving data from prospective cohort studies. The use of a prospective design and appropriate adjustment for a range of relevant risk factors increase the likelihood that RRs demonstrate causality between smoking and CVD [43]. Over 90% of current smokers (98% of past smokers) were current cigarette (respectively past cigarette) smokers, and therefore these findings describe the relationship between cigarette smoking and CVD.

4.4.1 Main findings

Smoking was associated with a wide range of cardiovascular events, both non-fatal and fatal. The effect of current smoking on vascular diseases and deaths was relatively homogenous, oscillating between 1.5 and 2.1, with stronger effects on the risks of subarachnoid haemorrhage, aortic aneurysm and peripheral vascular deaths, which may point toward a hypertensive effect of smoking [44] in addition to its atherosclerotic and thrombotic effects [45, 46]. The association with death caused by rupture of an aortic aneurysm was high and was higher than previously reported [47]. Second after aortic aneurysm, smoking had the strongest effect on death from peripheral vascular diseases. This is in agreement with the most recent meta-analysis which was literature-based, and included only 4 prospective cohort studies with 50 times less individuals than the current collaboration [4]. Regarding subarachnoid haemorrhage, I obtained estimates similar to that found by a meta-analysis of 14 longitudinal studies, which included, in total, 892 cases, compared to 1,313 in the ERFC [5]. The association between smoking and pulmonary embolism death was nearly as strong as the association with MI, contrary to previous belief that the RR is about half of the RRs for the more common forms of CVD which are MI or stroke [3, 48]. The association between smoking and atrial fibrillation (the most common type of sustained dysrhythmia) has been conflicting in the literature [7]. In the ERFC, smoking increased the risk of death from cardiac dysrhythmia by ~80% and the strength of association was similar to that of CHD and stroke. Evidence on smoking and death from heart failure has so far been limited and in the ERFC the RR for current versus never smokers was twice as high as previous estimates [49].

The relationship between smoking and MI or ischaemic stroke has been reported by numerous studies. The association was not significantly different for fatal versus non-fatal events, even if smokers experiencing an MI or stroke have been shown to have lower levels of atherosclerosis and less filled atherosclerotic plaques than non-smokers [50]. In the Nurses' Health Study, current smokers had a multivariate HR of 3.26 (2.97; 3.59) compared to never smokers for CHD and 2.81 (2.32; 3.41) for cerebrovascular diseases, which is higher than my estimates of 1.99 (1.86, 2.13) and 1.64 (1.54; 1.75) respectively [20]. Length of follow-up was longer (24 years versus around 14 years in the ERFC), which could affect RRs for CVD which are diseases of middle age and generally happen several decades after smoking initiation. Misclassification due to never smokers starting to smoke or current smokers at baseline stopping during follow-up may have attenuated associations.

There are several proposed mechanisms by which smoking could cause CVD. As developed in the **Introduction**, nicotine, carbon monoxide and oxidant gases are constituents of cigarette smoke and have been shown to promote atherosclerosis and thrombosis. Cigarette smoking is thought to contribute to the process of atherosclerosis by promoting inflammation, insulin sensitivity and lipid abnormalities; and to produce acute myocardial infarction by adversely affecting the balance of demand for myocardial oxygen and nutrients with myocardial blood supply [16]. One mechanism that links smoking to ischaemic stroke is structural artery wall damage and carotid atherosclerosis, leading to thrombosis or embolic phenomena [51]. The mechanism of thrombogenesis is a short-term effect of smoking, which includes increased levels of inflammatory markers as shown in the previous Chapter, and is supported by the reduction in risk of cerebrovascular diseases observed after cessation of smoking. For subarachnoid haemorrhage, there is evidence for an association of smoking with aneurysm formation, growth and rupture [52]. For haemorrhagic stroke, possible mechanisms are still speculative and may be mediated through arterial wall damage [53] and blood pressure [54].

The effect of smoking was very slightly stronger in women, younger age groups, non-diabetics, non-obese and non-hypertensive individuals. These findings suggest that smoking may act as a trigger of acute thrombotic events even in individuals with low levels of atherosclerosis. It has been suggested that women smoking are more at risk than men smoking. Hormonal treatment in post-menopausal women and different vessel composition would be reasons for this difference. My findings show a

modestly increased RR in women compared to men and which was borderline significant. Therefore, I did not replicate results from a recent literature based meta-analysis claiming a 25% greater risk of MI in women smoker compared to men smoker [14]. Inconsistent adjustment across studies included in the literature based meta-analysis and consistent adjustment in the ERFC make ERFC findings more reliable. For other risk factors, interplays were small except for age where RRs were about ¼ smaller for individuals ≥60 years old compared to individuals aged 40-59 years old.

I confirmed previous findings of a nonlinear association between smoking intensity and CVD risk [20, 55]. The first few cigarettes accounted for most of the increased risk and there was no "safe level" of smoking in terms of CVD risk. This suggests that the underlying biochemical and cellular processes may become saturated with small doses of toxic components from cigarette smoking. In that respect, it has been shown that passive smokers, who experienced lower levels of smoke than smokers of 1-5 cigarettes per day still have increased levels of platelet activation similar to that of active smokers[56]. Another reason may be that heavy smokers take lighter and shorter puffs than lighter smokers, therefore inhaling less smoke and absorbing less nicotine and carbon monoxide into the blood stream. In a study measuring cotinine, a metabolite of nicotine, levels increased most strongly from 0 to 10 CPD, reaching a plateau above 30 CPD [57]. The association between smoking duration and CHD was relatively weaker than that of intensity and it was approximately flat for cerebrovascular diseases. This suggests that smoking has mostly a short term impact on CVD risk rather than a long term damaging effect [51]. Because of the shapes of association between intensity and duration with CVD risk, I would not advocate summarizing smoking exposure with the use of pack-years when considering CVD risk, as had been done when investigating lung cancer association, where both duration and intensity have an approximately linear relationship with disease [17].

Cessation of smoking was rapidly beneficial. Ex-smokers experienced a risk which was not significantly higher than that of never smokers for all CVD subtypes, except for heart failure and aortic aneurysm. Smoking marijuana has been shown to act as a trigger of MI, but not smoking cigarettes [58]. Damage to the arterial wall caused by smoking takes time to heal, and excess risk still remained significantly higher than

that of never smokers for the next 10 years for all CVD and for the next 20 years for CHD only.

Estimates of number of years of life lost to smoking in current smokers (being at their maximum at age 40 years old with 7.7 years lost) were lower than previously published estimates of 10 years in men [23]. However, the previous estimate published by the British Doctor's Study was based on men born at the end of the 19th century and first half of the 20th century whose mortality rates and smoking habits differed from more recent generations. For women, the present study is the first to my knowledge to give estimates of years of life lost as a result of smoking. Reliability of these findings is warranted by the large scale of the data and the approximately equal proportion of men and women enrolled at baseline. For current smokers, vascular deaths accounted for slightly less than a third of excess deaths in both men and women, while cancer accounted for over a third of excess deaths. In ex-smokers, most excess deaths were due to cancer rather than CVD, which was a consequence of my finding that smoking effects are more easily undone on the cardiovascular system than on cancer conditions, in particular lung cancer.

4.4.2 Limitations

Despite its several strengths, this analysis contains several limitations. First, I conducted an independent participant meta-analysis, and heterogeneity between studies may affect estimates of the risk. However, I investigated the extent of heterogeneity using the I^2 statistic and attempted to explain some of this heterogeneity by grouping RRs according to geographical and other study characteristics. The definition of smoking status and assessments of duration and amount may have differed in studies, but data managers involved prior to the analyses corresponded with authors and attempted to resolve main discrepancies. Studies included in the ERFC were conducted over several decades during which cigarettes have changed in their chemical composition and new forms have appeared such as "filtered" and "light" cigarettes. However, violation of the proportional hazard assumption was marginal, showing no clinically significant decay of the RRs over time. In addition, residual bias could persist due to unmeasured or imprecisely measured confounding factors, for example alcohol consumption and socioeconomic factors. Some confounding variables such as menopause and hormone therapy were not available.

Analyses comparing the risk of current, ex and never smokers were not adjusted for amount and duration of smoking, in order to maximize power of the analyses. This may have introduced biases in the results asindividuals are likely to differ in terms of amount and duration of smoking. For example, women in the ERFC dataset are likely to have been smoking for shorter duration and lower amounts than their male counterparts, and this may distort the interaction between sex and smoking status on the risk of CVD. The same applies for individuals with higher body mass index, high blood pressure or high cholesterol levels.

Apart from for coronary disease and stroke, follow-up for other cardiovascular events and for lung cancer was only up to a fatal event. It is plausible that individuals experiencing a non-fatal cardiovascular event or lung cancer would modify their smoking habit. To limit the influence of potential prior diseases on baseline measurements, subsidiary analysis was done excluding the first five years of follow-up and RRs remain unchanged (**Table 4.1**).

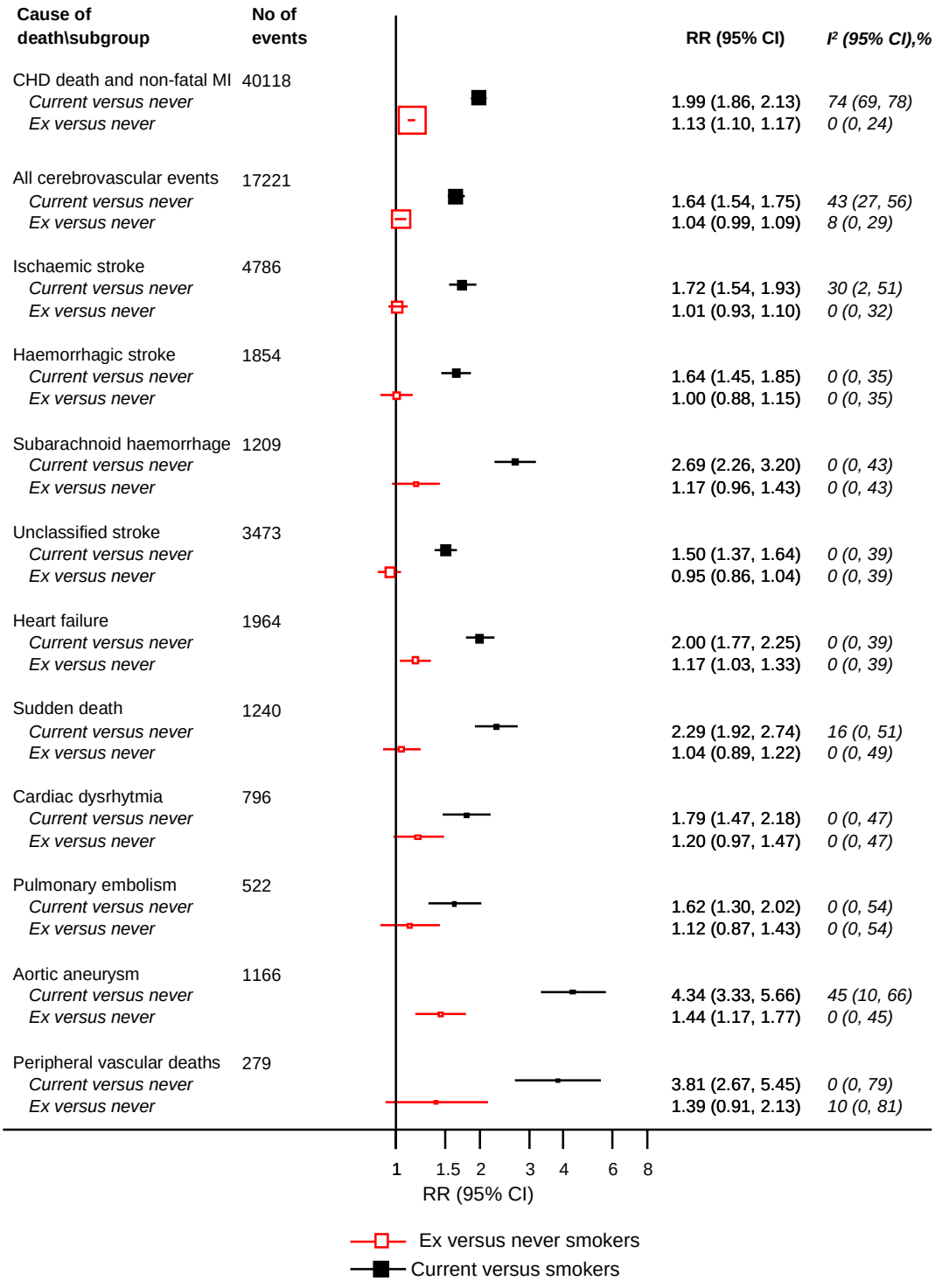
Baseline smoking status was used. Measurement error could be introduced if current smokers attempt to quit during follow-up and never smokers initiate smoking. In my dataset, amongst individuals who provided resurvey information, more than 80% had a stable smoking status <20 years of follow-up, 78% after 20-30 years of follow-up and 66% after 30-40 years of follow-up (**Table 4.6**). In particular, the proportion of never smokers initiating smoking was <1% independent of the time of follow-up. An increasing proportion of current smokers became ex-smokers: the proportion was <10% within the first 15 years of follow-up and reached 28% after 30-40 years of follow-up.

Taking into account measurement error in a categorical variable such as smoking status remains statistically challenging [59] and was beyond the scope of this thesis. Sex and age specific mortality rates for the overall population were used instead of rates in lifelong never smokers for the estimation of number of years lost as a result of smoking, which may have led to underestimation of the effect of smoking. Finally, European rates were used as reference rates, which may not be appropriate in the case of non-European studies included in the ERFC.

4.5 Conclusion

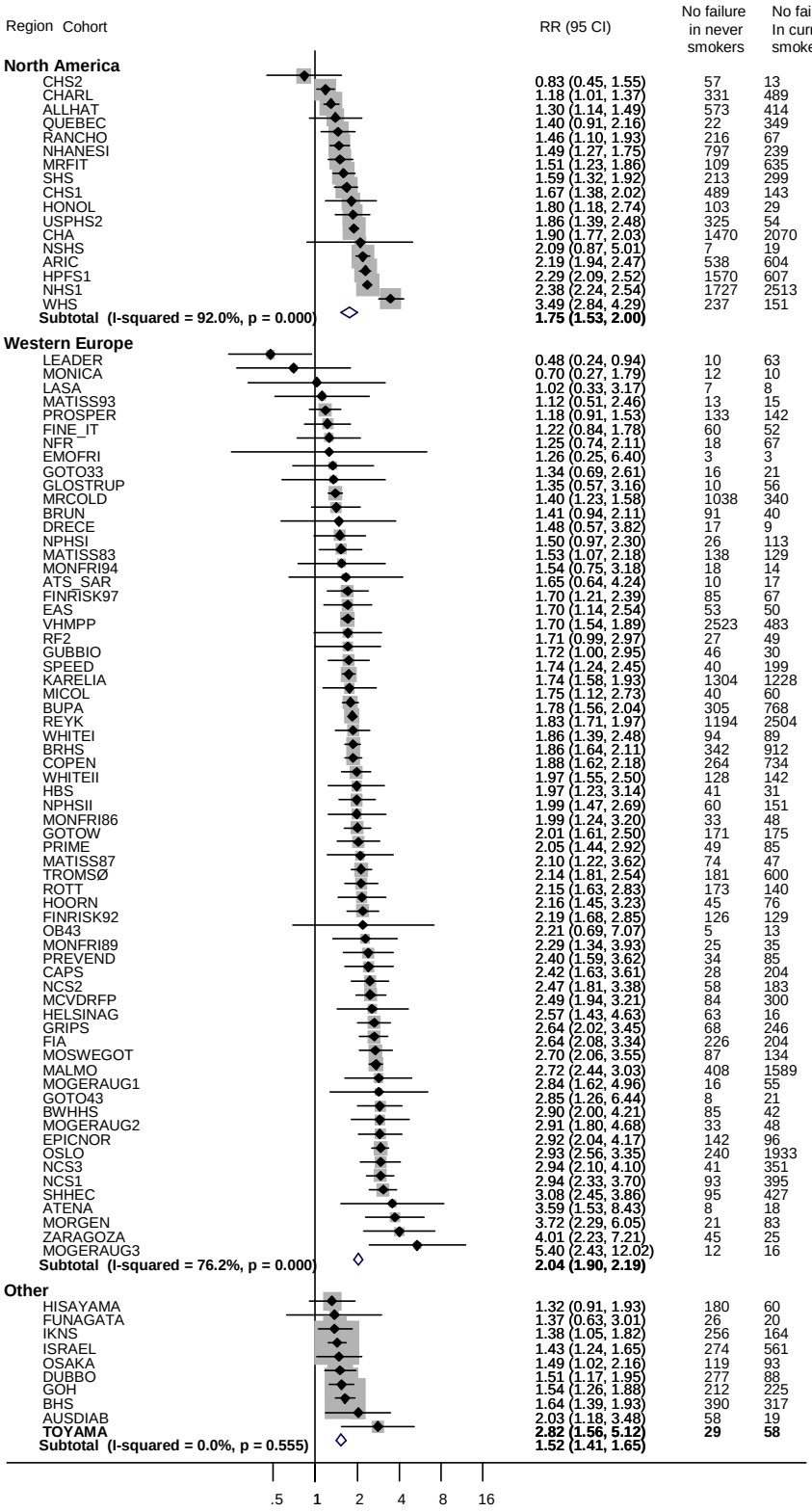
Smoking was associated with substantial CVD burden and reduction in life expectancy in both men and women. All types of CVD were associated with smoking. Women, younger age groups and non-diabetics experienced the greatest increase in risk of CVD when smoking, but in absolute terms, smoking caused more CVD events amongst men, older age groups and diabetics. Rather than reducing intake, the only safe option in terms of CVD risk was to stop smoking, which was associated with a rapid drop in CVD risk.

Figure 4.1: Risk ratios for non-fatal MI, non-fatal stroke and cardiovascular deaths by smoking status



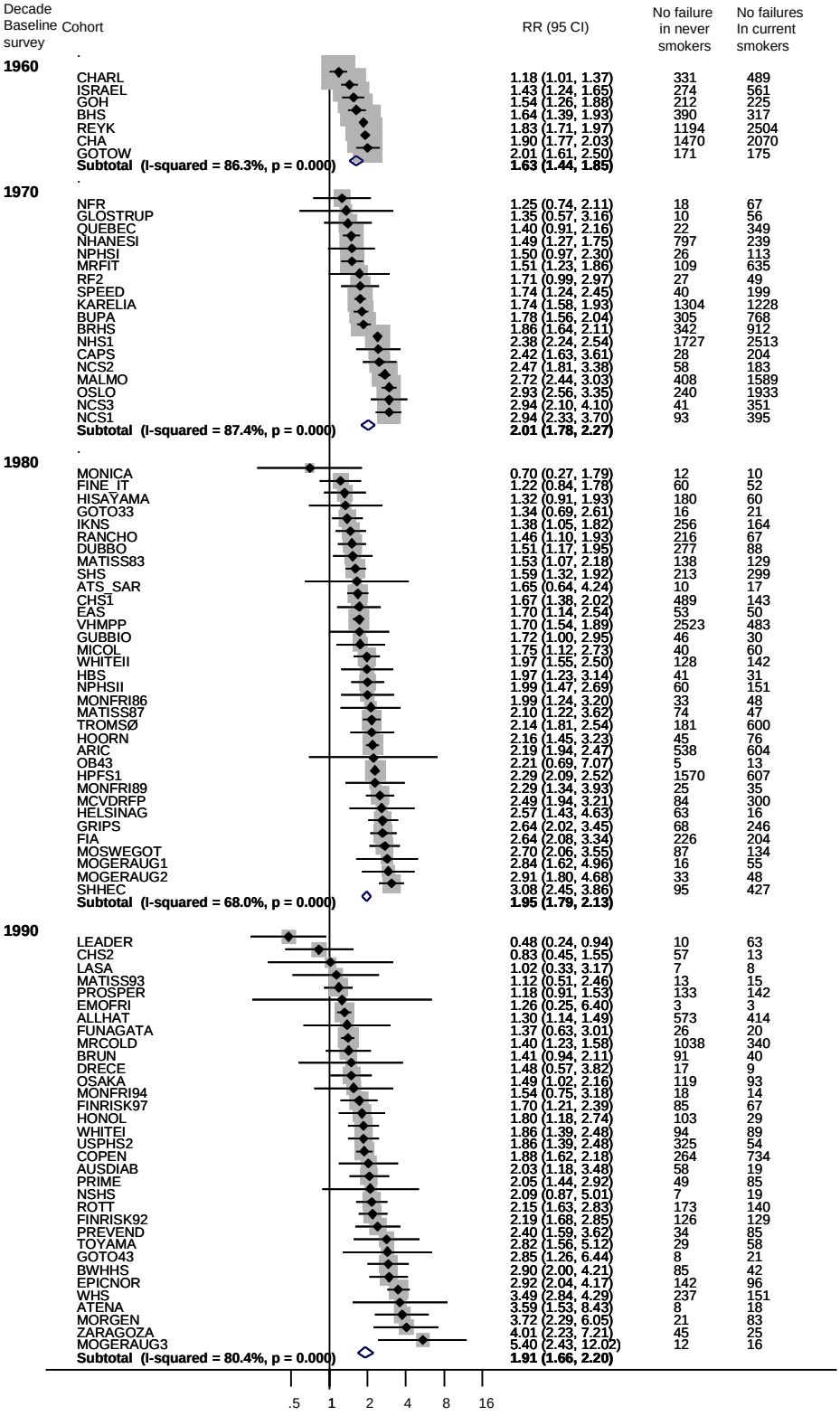
Study-specific $\log_e$ risk ratios adjusted for baseline age and body mass, and stratified, where appropriate, by sex and trial arm, were combined using a multivariate random-effects meta-analysis. Studies with fewer than 10 cases were excluded from the analysis of the outcome. Sizes of the data markers are proportional to the inverse of the variance of the log risk ratios. RRs were calculated with reference to never smokers. Individuals were censored at their first non-fatal myocardial infarction or stroke and at death.

Figure 4.2: Forest plot of risk ratios for fatal and non-fatal MI for current versus never smokers by region where studies were conducted



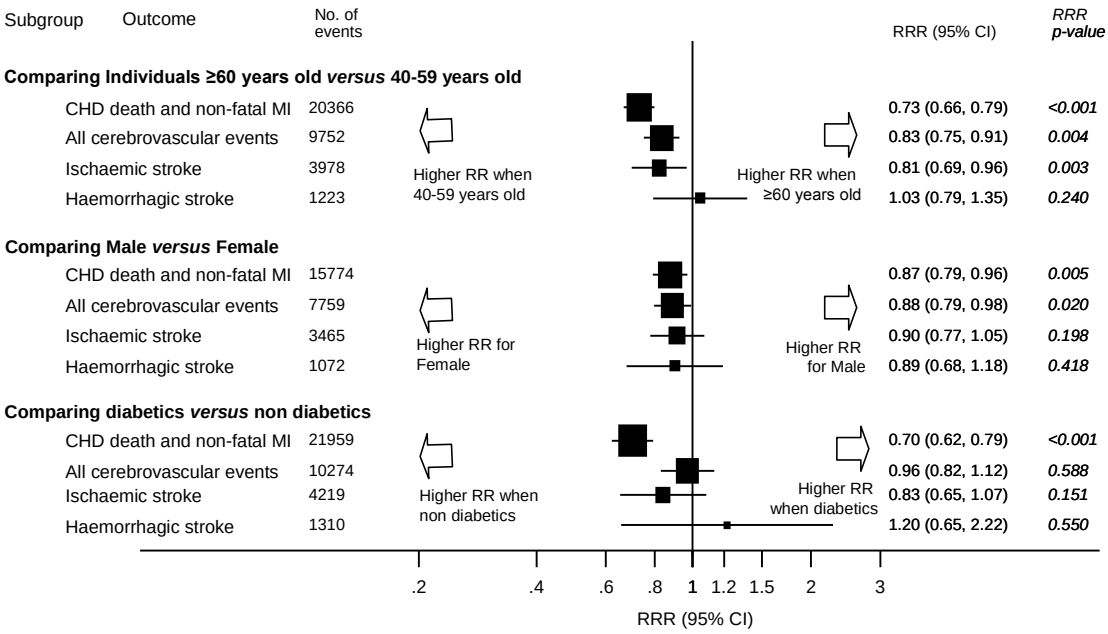
Boxes sizes are inversely proportional to study weights within each subgroup. Weights are from a random effect univariate meta-analysis. Study specific estimates were adjusted for baseline age and body mass index and stratified, when appropriate, by sex and trial arm. Studies with less than 10 events were excluded.

Figure 4.3: Forest plot of risk ratios for fatal and non-fatal MI for current versus never smokers by baseline decade of survey



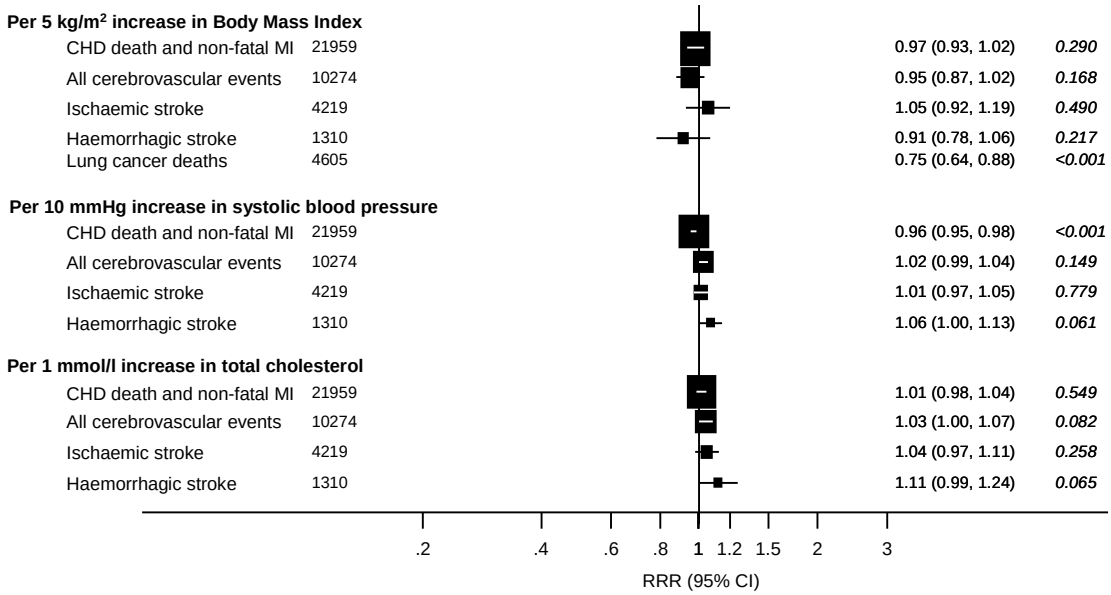
Boxes sizes are inversely proportional to study weights within each subgroup. Weights are from a random effect univariate meta-analysis. Study specific estimates were adjusted for baseline age and body mass index and stratified, when appropriate, by sex and trial arm. Studies with less than 10 events were excluded.

Figure 4.4: Adjusted relative risk ratios for major vascular morbidity and mortality and for current smokers versus never by levels of categorical CVD risk factors



RRR: Ratio of risk ratio; 95% CI: 95% confidence interval; p-value: p-value of 1 degree of freedom test of interaction between baseline characteristics (either categories for a categorical variable or in their continuous form for numeric variables) and smoking status. Study-specific $\log_e$ risk ratios adjusted for baseline age body mass index, history of diabetes, systolic blood pressure and total cholesterol, and stratified, where appropriate, by sex and trial arm, were combined using a multivariate random-effects meta-analysis. Studies with fewer than 5 cases were excluded from the analysis of the outcome. Sizes of the data markers are proportional to the inverse of the variance of the log risk ratios. RRs were calculated with reference to never smokers. Individuals were censored at their first non-fatal myocardial infarction or stroke and at death. For the interaction with sex, the data was restricted to studies including both Male and Females. All cerebrovascular events include ischaemic stroke, haemorrhagic stroke, unclassified strokes and subarachnoid haemorrhage.

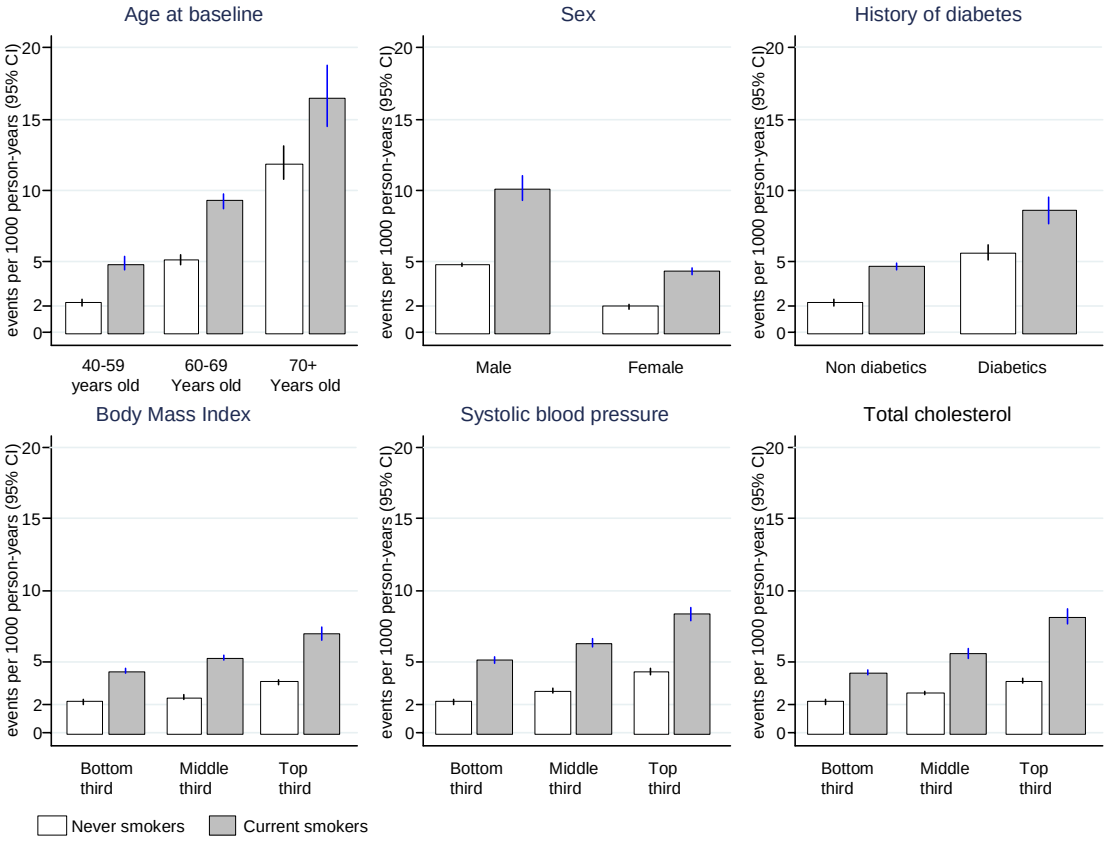
Figure 4.5: Adjusted relative risk ratios for major vascular morbidity and non-vascular mortality and for current smokers versus never by levels of continuous CVD risk factors



RRR: Ratio of risk ratio; 95% CI: 95% confidence interval; p-value: p-value of 1 degree of freedom test of interaction between baseline characteristics (either categories for a categorical variable or in their continuous form for numeric variables) and smoking status. Study-specific loge risk ratios adjusted for baseline age body mass index, history of diabetes, systolic blood pressure and total cholesterol, and stratified, where appropriate, by sex and trial arm, were combined using a multivariate random-effects meta-analysis. Studies with fewer than 5 cases were excluded from the analysis of the outcome. Sizes of the data markers are proportional to the inverse of the variance of the log risk ratios. RRs were calculated with reference to never smokers. Individuals were censored at their first non-fatal myocardial infarction or stroke and at death. For the interaction with sex, the data was restricted to studies including both Male and Females. All cerebrovascular events include ischaemic stroke, haemorrhagic stroke, unclassified strokes and subarachnoid haemorrhage.

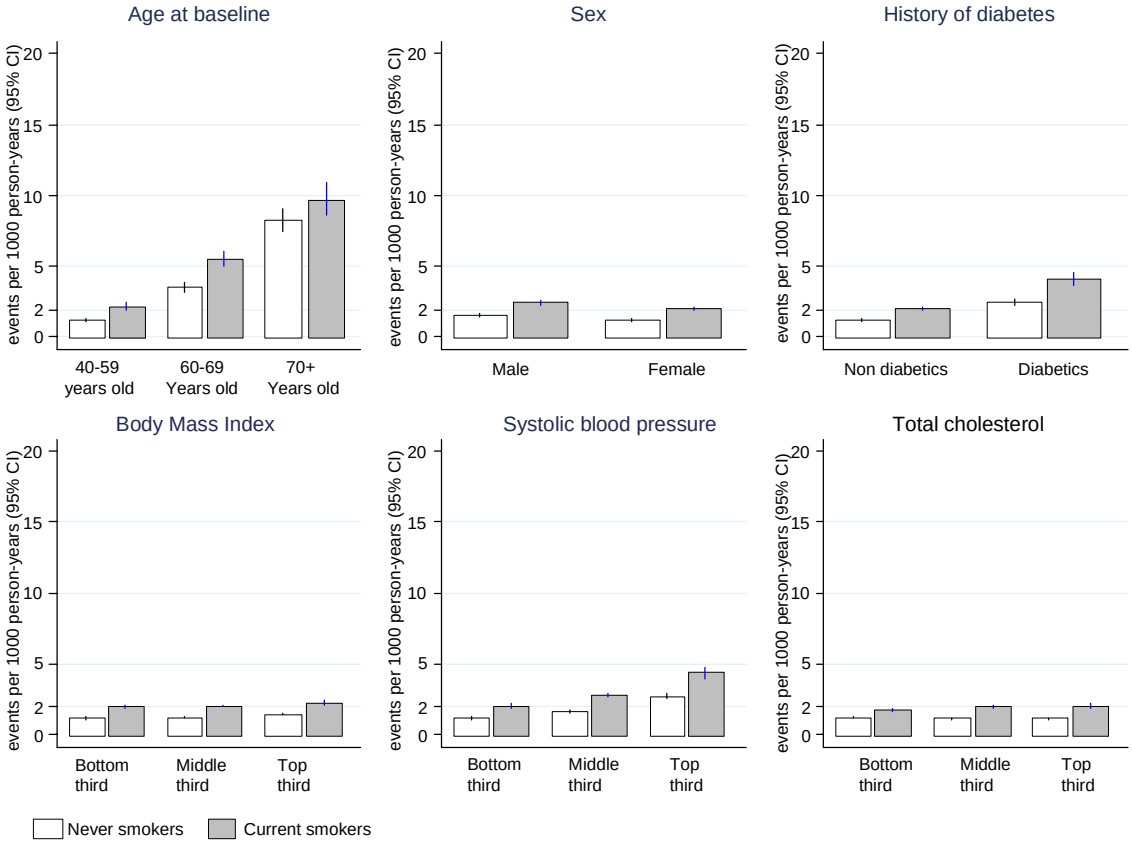
Figure 4.6 a-b: Absolute risk of fatal or non-fatal MI for current smokers and never, baseline characteristics, representing event rates on absolute scale using bar plots

a) Fatal and non-fatal MI



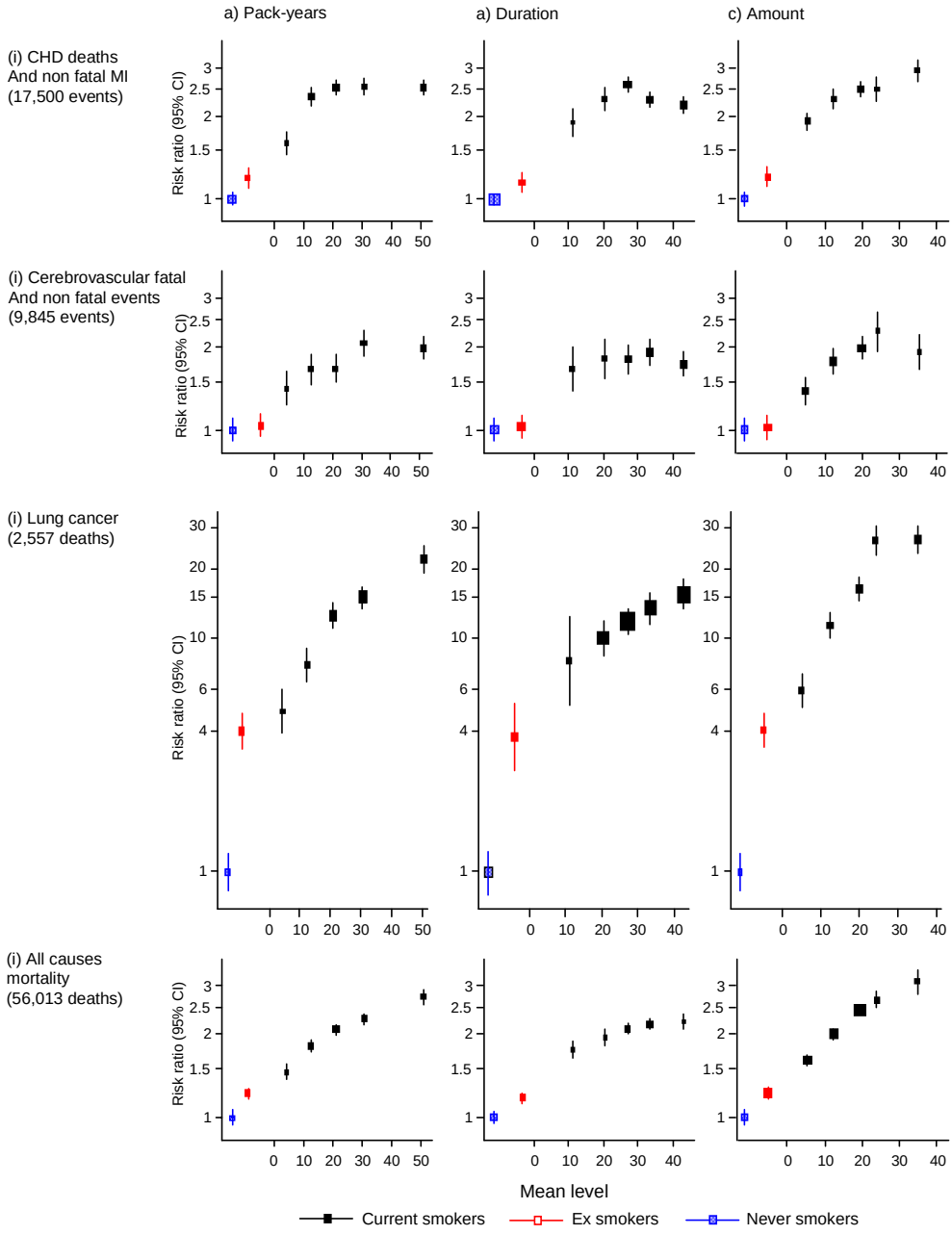
Please see the Method section for an explanation of this graphs.

b) fatal and non-fatal cerebrovascular events



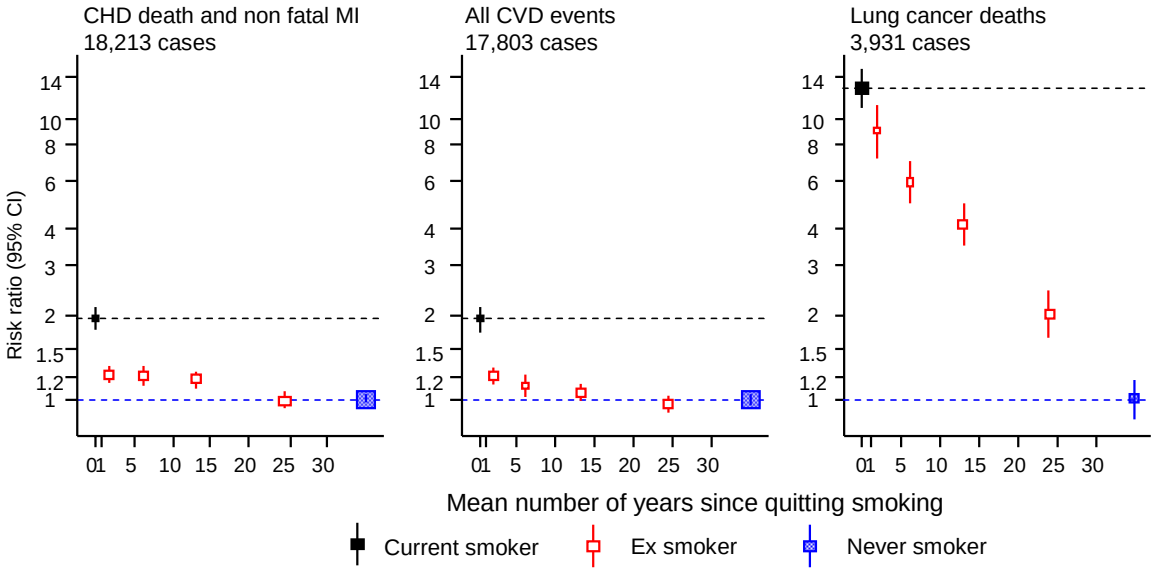
Please see the Method section for an explanation of this graphs.

Figure 4.7: Risk ratios for fatal and non-fatal MI, fatal and non-fatal cerebrovascular events, lung cancer deaths and all causes of death by smoking status and quintiles of smoking pack-years, duration (years) and amount (cigarettes per day)



Study-specific $\log_e$ risk ratios adjusted for baseline age and body mass index, and stratified, where appropriate, by sex and trial arm; were combined using a multivariate random-effects meta-analysis. Studies with fewer than 5 cases were excluded from the analysis of the outcome. Sizes of the data markers are proportional to the inverse of the variance of the log risk ratios. RRs were calculated with reference to never smokers. Confidence intervals are derived from "floating absolute variances". "Never" correspond to never smokers and to 0 pack-years/years/cigarettes per day. "Ex" corresponds to "ex-smokers". All cerebrovascular events include ischaemic stroke, haemorrhagic stroke, unclassified strokes and subarachnoid haemorrhage. Data restricted to studies with information on amount, duration and pack-years. Exclusive pipe or cigar smokers were considered as non-smokers for this analysis.

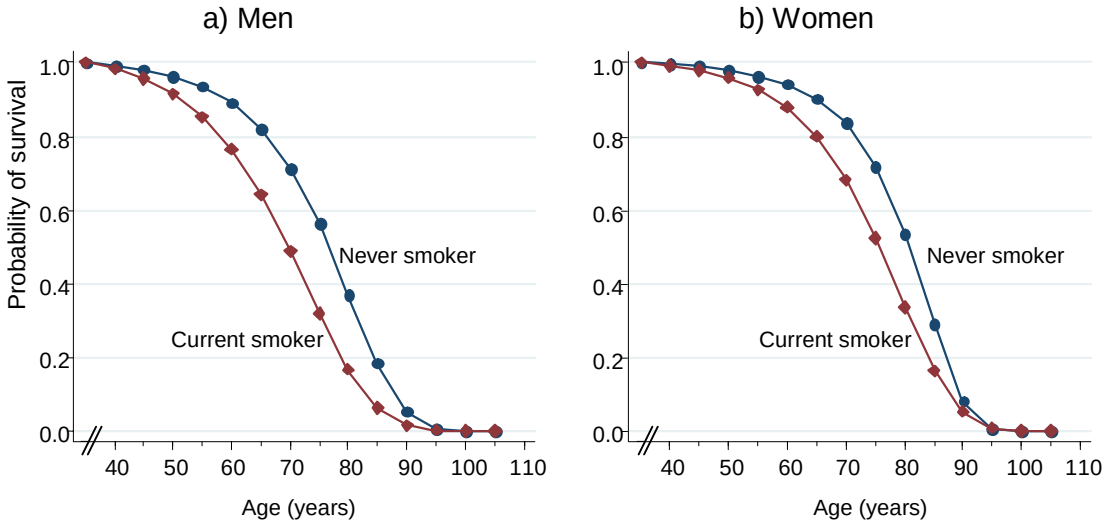
Figure 4.8: Risk ratios for CHD death and non-fatal MI, all fatal CVD and lung cancer death according to the number of years since quitting smoking



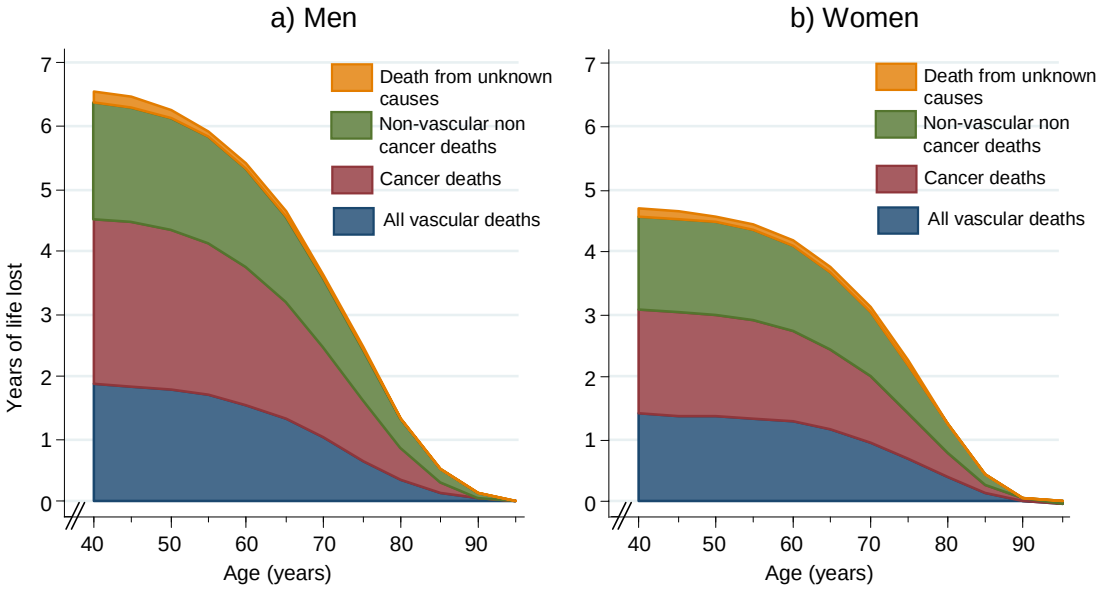
Study-specific $\log_e$ risk ratios adjusted for baseline age and body mass index, and stratified, where appropriate, by sex and trial arm; were combined using a multivariate random-effects meta-analysis. Studies with fewer than 10 cases were excluded from the analysis of the outcome. Sizes of the data markers are proportional to the inverse of the variance of the log risk ratios. RRs were calculated with reference to never smokers. Confidence intervals are derived from "floating absolute variances". Note "All CVD events" refers to all cardiovascular deaths.

Figure 4.9: Smoking and survival, according to sex and smoking status in current versus never smokers

a) Estimated survival of current versus never smokers



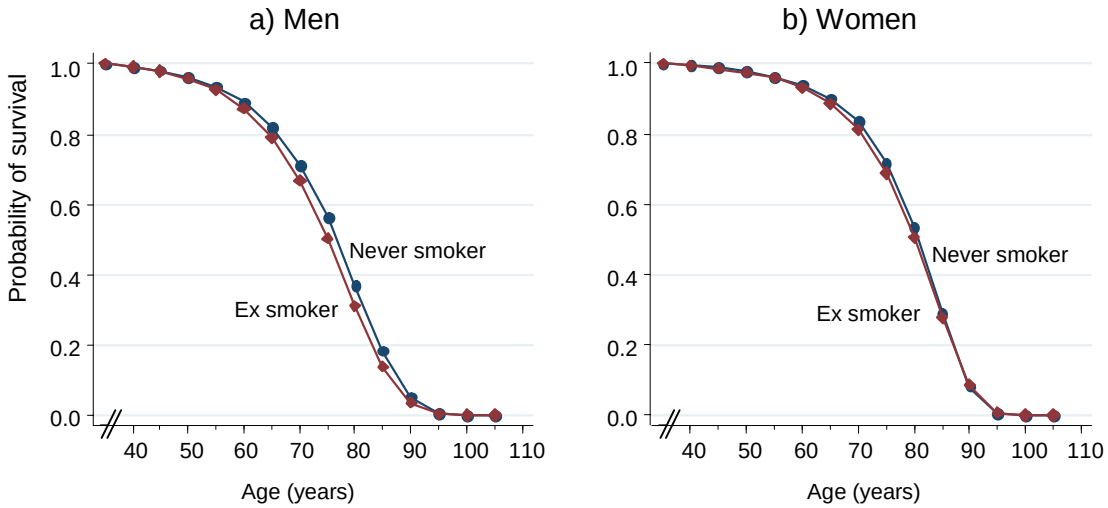
b) Estimated future years of life lost owing to smoking in current versus never smokers



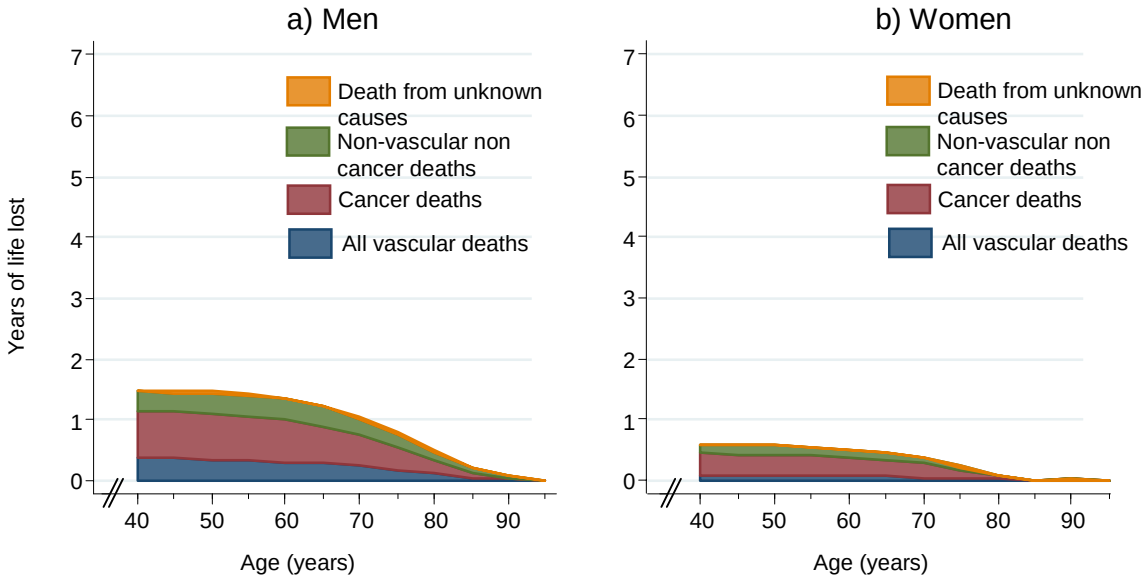
Panel a) shows estimated survival curves that were plotted by applying hazard ratios for death from any cause (specific for sex and age at risk) from the present analyses to mortality data for the European Union in 2000. Panel B shows the estimated number of years of life lost owing to smoking. Participants with known pre-existing cardiovascular disease at baseline were excluded from these analyses.

Figure 4.10: Smoking and survival, according to sex and smoking status in ex- versus never smokers

a) Estimated survival of ex versus never smokers



b) Estimated future years of life lost owing to smoking in ex- versus never smokers



Panel a) shows estimated survival curves that were plotted by applying hazard ratios for death from any cause (specific for sex and age at risk) from the present analyses to mortality data for the European Union in 2000 [39]. Panel b) shows the estimated number of years of life lost owing to smoking. Participants with known pre-existing cardiovascular disease at baseline were excluded from these analyses.

Table 4.1: Risk ratios for major outcomes for current and ex versus never smokers, adjusted for age, sex and BMI under different conditions

			Current versus never smoker		Ex versus never smokers	
Analysis	**Outcome**	**No of events**	**RR (95% CI)**	**I2 (95% CI)**	**RR (95% CI)**	**I2 (95% CI)**
Including everyone						
	Coronary heart disease*	40118	1.99 (1.86, 2.13)	*74 (69, 78)*	1.13 (1.10, 1.17)	*0 (0, 24)*
	Ischaemic stroke*	3803	1.74 (1.57, 1.93)	*7 (0, 37)*	1.04 (0.94, 1.15)	*6 (0, 36)*
	CVD deaths	43861	2.01 (1.89, 2.14)	*72 (66, 77)*	1.08 (1.05, 1.12)	*7 (0, 28)*
Excluding the first 5 years of follow-up						
	Coronary heart disease*	29820	1.91 (1.78, 2.04)	*70 (62,76)*	1.09 (1.05, 1.13)	*2 (0, 22)*
	Ischaemic stroke*	3262	1.69 (1.46, 1.96)	*54 (22,73)*	1.06 (0.94, 1.19)	*0 (0, 50)*
	CVD deaths	36543	1.98 (1.85, 2.11)	*70 (63,76)*	1.07 (1.03, 1.11)	*8 (0, 31)*
Excluding studies recording only fatal MI						
	Non-fatal MI	15997	1.89 (1.74, 2.06)	*61 (50,70)*	1.13 (1.08, 1.19)	*2(0, 22)*
	Fatal MI	6935	2.10 (1.84, 2.39)	*57 (42,68)*	1.11 (1.03, 1.20)	*0(0, 31)*
Excluding studies recording only fatal Ischaemic strokes						
	Fatal Ischaemic stroke	2271	1.63 (1.43, 1.85)	*0 (0, 85)*	1.05 (0.86, 1.27)	*0(0, 85)*
	Non-fatal Ischaemic stroke	64	1.59 (0.74, 3.43)	*0 (0, 52)*	1.08 (0.49, 2.36)	*33(0, 63)*
Censoring individuals at death only rather than at first non-fatal MI/stroke or death						
	Fatal MI	28532	2.08 (1.93, 2.25)	*71 (65,77)*	1.14 (1.09, 1.19)	*9 (0, 30)*
	Fatal Ischaemic stroke	1334	1.71 (1.38, 2.12)	*39 (6,61)*	0.94 (0.80, 1.11)	*0 (0, 40)*
	All fatal CVD	52698	1.95 (1.83, 2.07)	*75 (70,80)*	1.07 (1.03, 1.12)	*25 (4, 41)*
Restricting to studies including both Males and Females and fitting and interaction with sex						
In Females	Coronary heart disease*	8569	2.12 (1.87, 2.39)	47 (30, 59)**	1.10 (1.03, 1.18)	0 (0, 28)**
	Ischaemic stroke*	1505	1.73 (1.51, 1.98)	0 (0, 28)**	1.00 (0.84, 1.18)	0 (0, 28)**
	CVD deaths	10557	1.95 (1.77, 2.15)	37 (16, 52)**	0.99 (0.92, 1.07)	0 (0, 28)**
In Males	Coronary heart disease*	16924	1.99 (1.86, 2.14)	49 (33, 61)	1.15 (1.09,1.21)	0 (0, 28)
	Ischaemic stroke*	1586	1.59 (1.39, 1.82)	0 (0, 38)	0.96 (0.82,1.11)	3 (0, 39)
	CVD deaths	15072	2.09 (1.95, 2.25)	43 (24, 57)	1.06 (1.00,1.12)	0 (0, 29)

*Includes both fatal and non-fatal events

** I^2 for the interaction term (the effect for Female was computed adding the interaction term on top of the effect for Male). Risk ratios were adjusted for age at baseline and BMI and stratified, where appropriate, by sex and trial arm. Studies with fewer than 10 events for a specific outcome were excluded from the analysis of the outcome.

Table 4.2: : Risk ratios for coronary heart disease, ischaemic stroke and all cardiovascular mortality for current versus never smokers, with progressive adjustment for baseline levels of biological, socioeconomic and behavioural risk factors

	Coronary heart disease*			Ischaemic stroke*			All CVD deaths		
Progressive adjustment	*No of events*	RR (95% CI)	*I^2 (95%CI) %*	*No of events*	RR (95% CI)	*I^2 (95%CI) %*	*No of events*	RR (95% CI)	*I^2 (95%CI) %*
Age, sex and BMI	*28,095*	2.03 (1.88, 2.18)	*72 (65, 77)*	*4,184*	1.66 (1.47, 1.87)	*31 (0, 53)*	*24,547*	2.04 (1.88, 2.20)	*69 (61, 75)*
Plus systolic blood pressure	*28,095*	2.07 (1.92, 2.23)	*72 (66, 77)*	*4,184*	1.69 (1.50, 1.90)	*31 (0, 53)*	*24,547*	2.08 (1.92, 2.25)	*69 (61, 75)*
Plus history of diabetes	*28,095*	2.05 (1.91, 2.21)	*71 (64, 76)*	*4,184*	1.70 (1.52, 1.92)	*32 (0, 53)*	*24,547*	2.08 (1.92, 2.24)	*69 (61, 75)*
Plus total cholesterol	*28,095*	2.01 (1.88, 2.16)	*68 (60, 74)*	*4,184*	1.72 (1.53, 1.94)	*31 (0, 53)*	*24,547*	2.08 (1.93, 2.24)	*67 (58, 73)*
Additional adjustment									
Lifestyle factors									
Age, sex and BMI	*19,175*	2.02 (1.87, 2.19)	*60 (46, 70)*	*3,126*	1.75 (1.49, 2.05)	*46 (16, 65)*	*15,230*	2.08 (1.90, 2.28)	*60 (46, 71)*
Plus education	*19,175*	1.98 (1.84, 2.14)	*57 (42, 69)*	*3,126*	1.73 (1.48, 2.03)	*46 (16, 65)*	*15,230*	2.07 (1.89, 2.27)	*60 (45, 71)*
Age, sex and BMI	*14,423*	1.90 (1.75, 2.06)	*51 (32, 65)*	*1,733*	1.52 (1.33, 1.74)	*0 (0, 47)*	*12,933*	1.87 (1.70, 2.06)	*58 (40, 70)*
Plus occupation/job	*14,423*	1.87 (1.72, 2.03)	*49 (29, 64)*	*1,733*	1.46 (1.10, 1.94)	*0 (0, 47)*	*12,933*	1.86 (1.68, 2.06)	*57 (39, 69)*
Age, sex and BMI	*13,159*	1.96 (1.80, 2.13)	*55 (41, 65)*	*2,916*	1.74 (1.49, 2.02)	*28 (0, 53)*	*10,934*	1.87 (1.70, 2.06)	*51 (35, 63)*
Plus alcohol consumption	*13,159*	2.01 (1.84, 2.19)	*56 (42, 66)*	*2,916*	1.79 (1.54, 2.07)	*24 (0, 51)*	*10,934*	1.86 (1.68, 2.06)	*52 (37, 64)*
Lipids									
Basic model[$]	*11,446*	1.95 (1.76, 2.15)	*58 (44, 68)*	*3,203*	1.75 (1.58, 1.94)	*0 (0, 41)*	*6,282*	1.98 (1.78, 2.22)	*43 (22, 58)*
Plus non-HDL-C, HDL-C & $\log_e$ triglycerides[+]	*11,446*	1.91 (1.75, 2.08)	*50 (33, 63)*	*3,203*	1.76 (1.57, 1.97)	*0 (0, 41)*	*43,861*	2.03 (1.81, 2.27)	*44 (23, 59)*
Inflammatory markers									
Basic model[$$]	*7,240*	1.87 (1.63, 2.15)	*60 (43, 72)*	*2,150*	1.72 (1.32, 2.23)	*63 (37, 78)*	*4,260*	2.11 (1.84, 2.42)	*41 (12, 61)*
Plus $\log_e$ CRP	*7,240*	1.70 (1.48, 1.95)	*59 (43, 71)*	*2,150*	1.59 (1.24, 2.04)	*59 (30, 76)*	*4,260*	1.87 (1.65, 2.13)	*34 (1, 56)*
Basic model[$$]	*6,748*	2.08 (1.87, 2.30)	*44 (18, 62)*	*2,445*	1.76 (1.46, 2.12)	*44 (9, 65)*	*4,259*	2.23 (1.93, 2.59)	*55 (35, 69)*
Plus fibrinogen	*6,748*	1.95 (1.76, 2.15)	*39 (10, 59)*	*2,445*	1.68 (1.40, 2.01)	*40 (4, 63)*	*4,259*	2.04 (1.78, 2.35)	*49 (25, 65)*

*Includes both fatal and non-fatal events. $: The basic model is adjusted for age, sex, systolic blood pressure, history of diabetes and body-mass index. $$: The basic model is adjusted for age, sex, systolic blood pressure, history of diabetes, body-mass index and total cholesterol. +: total cholesterol was not included in this model. RRs were computed within cohorts and combined using random effect meta-analysis. Study-specific RRs were adjusted as shown and stratified where appropriate by sex and trial arm. Studies recording fewer than 5 events during follow-up were excluded from the analysis of that outcome.

Table 4.3: Tests of the proportional hazard assumption

	Current versus never smokers		Ex versus never smokers		
Outcome	**Change (95% CI) in RRs per 5 years additional follow-up**	***I^2 (95% CI)***	**Change (95% CI) in RRs per 5 years additional follow-up**	***I^2 (95% CI)***	**Overall P-value**
All cardiovascular	-0.09 (-0.12; -0.06)	26 (4, 43)	-0.03 (-0.05; -0.01)	2 (0, 21)	0
CHD death and non-fatal MI	-0.10 (-0.13; -0.08)	7 (0, 29)	-0.04 (-0.07; -0.02)	0 (0, 26)	0
All cerebrovascular events	-0.07 (-0.12; -0.01)	33 (11, 49)	0.00 (-0.04; 0.04)	0 (0, 28)	0.065
Ischaemic stroke	-0.06 (-0.15; 0.04)	9 (0, 37)	0.06 (-0.04; 0.18)	0 (0, 34)	0.29
Haemorrhagic stroke	0.05 (-0.05; 0.16)	0 (0, 36)	0.09 (-0.04; 0.24)	0 (0, 37)	0.236
Subarachnoid haemorrhage	-0.19 (-0.33; -0.03)	14 (0, 47)	-0.17 (-0.30; -0.03)	0 (0, 45)	0.005
Unclassified stroke (fatal)	-0.04 (-0.13; 0.06)	21 (0, 49)	-0.03 (-0.11; 0.07)	0 (0, 42)	0.605
Heart failure (fatal)	-0.03 (-0.17; 0.14)	17 (0, 47)	-0.01 (-0.23; 0.25)	36 (0, 59)	0.923
Sudden death (fatal)	-0.15 (-0.25; -0.04)	0 (0, 51)	0.02 (-0.11; 0.17)	0 (0, 58)	0.028
Cardiac dysrhythmia (fatal)	-0.16 (-0.29; 0.01)	0 (0, 48)	-0.10 (-0.27; 0.11)	0 (0, 49)	0.108
Pulmonary embolism (fatal)	0.06 (-0.10; 0.25)	0 (0, 54)	0.08 (-0.11; 0.31)	0 (0, 54)	0.565
Aortic aneurysm (fatal)	0.12 (-0.05; 0.33)	10 (0, 45)	0.09 (-0.10; 0.33)	0 (0, 50)	0.285
Peripheral vascular disease (fatal)	-0.20 (-0.48; 0.24)	23 (0, 68)	-0.33 (-0.53; -0.04)	0 (0, 79)	0.058

RR: Risk ratio. The proportional hazard assumption was tested by adding a time dependant interaction to study specific models, with adjustment for age and body mass index, and stratification, when appropriate, by sex and trial arm. The change in RRs per 5 years additional follow-up was obtained by taking the pooled time dependant interaction estimate multiplied by 5, exponentiating it and subtracting 1. Overall p-values are derived from a 2 degree of freedom test of nullity of the time dependant interactions, testing jointly the significance of current versus never and ex versus never interaction estimates. All cerebrovascular events include ischaemic stroke, haemorrhagic stroke, unclassified strokes and subarachnoid haemorrhage.

Table 4.4: : Risk ratios for coronary heart disease, ischaemic stroke and all cardiovascular mortality for ex versus never smokers, with progressive adjustment for baseline levels of biological, socioeconomic and behavioural risk factors

	Coronary heart disease*			Ischaemic stroke*			All CVD deaths		
Progressive adjustment	*No of events*	RR (95% CI)	*I^2 (95%CI) %*	*No of events*	RR (95% CI)	*I^2 (95%CI) %*	*No of events*	RR (95% CI)	*I^2 (95%CI) %*
Age, sex and BMI	*28,095*	1.14 (1.10, 1.18)	*1 (0, 26)*	*4,184*	1.01 (0.92, 1.11)	*7 (0, 35)*	*24,547*	1.08 (1.02, 1.14)	*12 (0, 33)*
Plus systolic blood pressure	*28,095*	1.15 (1.10, 1.19)	*7 (0, 29)*	*4,184*	1.01 (0.92, 1.11)	*6 (0, 34)*	*24,547*	1.08 (1.03, 1.14)	*11 (0, 33)*
Plus history of diabetes	*28,095*	1.14 (1.09, 1.18)	*7 (0, 29)*	*4,184*	1.03 (0.94, 1.14)	*5 (0, 33)*	*24,547*	1.08 (1.02, 1.13)	*12 (0, 33)*
Plus total cholesterol	*28,095*	1.12 (1.08, 1.17)	*11 (0, 32)*	*4,184*	1.04 (0.94, 1.15)	*6 (0, 34)*	*24,547*	1.07 (1.02, 1.13)	*13 (0, 34)*
Additional adjustment									
Lifestyle factors									
Age, sex and BMI	*19,175*	1.15 (1.10, 1.20)	*4 (0, 30)*	*3,126*	0.94 (0.85, 1.05)	*8 (0, 39)*	*15,230*	1.06 (1.00, 1.12)	*4 (0, 28)*
Plus education	*19,175*	1.14 (1.07, 1.21)	*2 (0, 24)*	*3,126*	0.95 (0.84, 1.08)	*7 (0, 38)*	*15,230*	1.08 (1.02, 1.15)	*12 (0, 37)*
Age, sex and BMI	*14,423*	1.10 (1.05, 1.16)	*0 (0, 34)*	*1,733*	0.97 (0.81, 1.17)	*25 (0, 56)*	*12,933*	1.08 (0.99, 1.17)	*28 (0, 51)*
Plus occupation/job	*14,423*	1.10 (1.03, 1.18)	*0 (0, 34)*	*1,733*	0.98 (0.74, 1.30)	*26 (0, 57)*	*12,933*	1.07 (0.99, 1.16)	*27 (0, 50)*
Age, sex and BMI	*13,159*	1.10 (1.04, 1.17)	*15 (0, 37)*	*2,916*	0.97 (0.86, 1.09)	*0 (0, 39)*	*10,934*	1.06 (1.00, 1.13)	*7 (0, 31)*
Plus alcohol consumption	*13,159*	1.12 (1.05, 1.19)	*17 (0, 39)*	*2,916*	0.99 (0.88, 1.12)	*0 (0, 39)*	*10,934*	1.10 (1.02, 1.18)	*8 (0, 32)*
Lipids									
Basic model[$]	*11,446*	1.15 (1.09, 1.22)	*0 (0, 30)*	*3,203*	1.09 (0.97, 1.22)	*0 (0, 41)*	*6,282*	1.09 (1.01, 1.18)	*6 (0, 33)*
Plus non-HDL-C, HDL-C & $\log_e$ triglycerides[+]	*11,446*	1.15 (1.09, 1.22)	*0 (0, 30)*	*3,203*	1.11 (0.99, 1.24)	*0 (0, 41)*	*6,282*	1.09 (1.00, 1.19)	*2 (0, 26)*
Inflammatory markers									
Basic model[$$]	*7,240*	1.14 (1.06, 1.22)	*0 (0, 36)*	*2,150*	1.14 (1.00, 1.31)	*3 (0, 52)*	*4,260*	1.16 (1.03, 1.30)	*16 (0, 45)*
Plus $\log_e$ CRP	*7,240*	1.11 (1.04, 1.19)	*0 (0, 36)*	*2,150*	1.13 (0.99, 1.28)	*0 (0, 51)*	*4,260*	1.13 (1.00, 1.27)	*17 (0, 46)*
Basic model[$$]	*6,748*	1.15 (1.06, 1.25)	*8 (0, 38)*	*2,445*	1.02 (0.90, 1.16)	*0 (0, 44)*	*4,259*	1.12 (1.02, 1.23)	*0 (0, 37)*
Plus fibrinogen	*6,748*	1.15 (1.07, 1.24)	*7 (0, 37)*	*2,445*	1.01 (0.88, 1.14)	*0 (0, 44)*	*4,259*	1.11 (1.01, 1.22)	*0 (0, 37)*

*Includes both fatal and non-fatal events. $: The basic model is adjusted for age, sex, systolic blood pressure, history of diabetes and body-mass index. $$: The basic model is adjusted for age, sex, systolic blood pressure, history of diabetes, body-mass index and total cholesterol. +: total cholesterol was not included in this model. RRs were computed within cohorts and combined using random effect meta-analysis. Study-specific RRs were adjusted as shown and stratified where appropriate by sex and trial arm. Studies recording fewer than 5 events during follow-up were excluded from the analysis of that outcome.

Table 4.5: Risk ratios for CVD and fatal and non-fatal MI by number of years since stopping smoking, with progressive adjustment for age started smoking and past number of pack-years.

Outcome	Model adjusted for	*No of events*	RRs (95% CI) for current versus never smokers	RRs (95% CI) for ex versus never smokers			
				<5 years	5-10 years	10-20 years	≥20 years
CHD deaths and non-fatal MI	Ages & BMI	*15578*	2.16 (2.03, 2.29)	1.36 (1.20, 1.55)	1.23 (1.06, 1.43)	1.21 (1.07, 1.36)	1.02 (0.92, 1.14)
	+ Age started	*15578*	2.04 (1.92, 2.18)	1.28 (1.12, 1.45)	1.15 (0.98, 1.34)	1.13 (1.00, 1.27)	0.94 (0.84, 1.06)
	Ages & BMI	*24905*	2.01 (1.89, 2.14)	1.40 (1.26, 1.56)	1.18 (1.04, 1.34)	1.03 (0.92, 1.15)	0.85 (0.76, 0.96)
	+ Past number of pack-years	*24905*	1.78 (1.65, 1.92)	1.23 (1.10, 1.38)	1.06 (0.93, 1.21)	0.94 (0.84, 1.06)	0.82 (0.73, 0.93)
All CVD events	Ages & BMI	*28769*	1.92 (1.76, 2.10)	1.26 (1.12, 1.41)	1.12 (0.99, 1.26)	1.09 (0.99, 1.20)	0.93 (0.84, 1.02)
	+ Age started	*28769*	1.87 (1.77, 1.98)	1.20 (1.09, 1.33)	1.05 (0.93, 1.18)	1.02 (0.93, 1.12)	0.86 (0.78, 0.94)
	Ages & BMI	*44547*	2.07 (1.92, 2.24)	1.64 (1.49, 1.80)	1.27 (1.14, 1.43)	1.00 (0.90, 1.11)	0.93 (0.84, 1.04)
	+ Past number of pack-years	*44547*	1.66 (1.55, 1.77)	1.24 (1.12, 1.36)	0.99 (0.88, 1.11)	0.83 (0.75, 0.92)	0.85 (0.76, 0.95)
Lung cancer deaths	Ages & BMI	*2301*	14.81 (11.88, 18.47)	10.49 (7.98, 13.78)	6.40 (4.57, 8.97)	3.84 (2.76, 5.34)	2.09 (1.47, 2.97)
	+ Age started	*2301*	10.80 (8.55, 13.64)	7.38 (5.55, 9.82)	4.41 (3.11, 6.24)	2.64 (1.88, 3.71)	1.38 (0.96, 1.99)
	Ages & BMI	*15039*	2.29 (2.10, 2.50)	1.93 (1.68, 2.22)	1.37 (1.15, 1.63)	1.21 (1.05, 1.41)	1.11 (0.95, 1.31)
	+ Past number of pack-years	*15039*	1.66 (1.50, 1.84)	1.34 (1.15, 1.56)	1.01 (0.84, 1.21)	0.98 (0.84, 1.15)	1.04 (0.89, 1.23)

No: Number; BMI: Body Mass Index; CHD: Coronary heart disease; CVD: Cardiovascular diseases; RR: Risk ratio; Study-specific loge risk ratios stratified, where appropriate, by sex and trial arm; were combined using a multivariate random-effects meta-analysis. Studies with fewer than 10 individuals experiencing an outcome were excluded from the analysis of that outcome.

Table 4.6: Proportion of individuals changing their smoking status compared to baseline according to time since baseline

Time between baseline and resurvey	Smoking status				
	Stable status	Never smoker becomes current smoker	Current smoker becomes ex-smoker	Never smokers becomes Ex-smoker	***Total No individual resurveys***
<5 years	88.3%	0.1%	7.0%	4.6%	***236,639***
5-10 years	88.3%	0.3%	7.9%	3.5%	***158,646***
10-15 years	89.2%	0.6%	8.8%	1.4%	***34,846***
15-20 years	85.6%	0.4%	12.3%	1.4%	***12,835***
20-30 years	77.9%	0.3%	20.5%	1.3%	***10,715***
30-40 years	66.0%	0.8%	28.1%	5.1%	***2,093***
Total No individual resurveys	***400,794***	***1,083***	***36,695***	***17,202***	***455,774***
Total proportions	***88.0%***	***0.2%***	***8.1%***	***3.8%***	***100%***

No: Number. Percentages are row percentages and are computed based on 455,774 individual resurveys. Individuals who were never smokers and changed their status to ex-smoker at the next resurvey were understood to have started and stopped smoking during the time interval between the two surveys, or to have answered the question incorrectly.

Reference List

(1) Dept.of Health and Human Services. The health consequences of smoking: a report of the Surgeon General. 2004.

(2) Pirie K, Peto R, Reeves GK, Green J, Beral V. The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. *Lancet* 2012 October 26.

(3) Ageno W, Becattini C, Brighton T, Selby R, Kamphuisen PW. Cardiovascular risk factors and venous thromboembolism: a meta-analysis. *Circulation* 2008 January 1;117(1):93-102.

(4) Willigendael EM, Teijink JA, Bartelink ML et al. Influence of smoking on incidence and prevalence of peripheral arterial disease. *J Vasc Surg* 2004 December;40(6):1158-65.

(5) Feigin VL, Rinkel GJE, Lawes CMM et al. Risk Factors for Subarachnoid Hemorrhage. *Stroke* 2005 December 1;36(12):2773-80.

(6) Pomp ER, le CS, Rosendaal FR, Doggen CJ. Risk of venous thrombosis: obesity and its joint effect with oral contraceptive use and prothrombotic mutations. *Br J Haematol* 2007 October;139(2):289-96.

(7) D'Alessandro A, Boeckelmann I, Hammwhoner M, Goette A. Nicotine, cigarette smoking and cardiac arrhythmia: an overview. *Eur J Cardiovasc Prev Rehabil* 2011 May 25.

(8) Danaei G, Finucane MM, Lu Y et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *Lancet* 2011 July 2;378(9785):31-40.

(9) Finucane MM, Stevens GA, Cowan MJ et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet* 2011 February 12;377(9765):557-67.

(10) Farzadfar F, Finucane MM, Danaei G et al. National, regional, and global trends in serum total cholesterol since 1980: systematic analysis of health examination surveys and epidemiological studies with 321 country-years and 3.0 million participants. *Lancet* 2011 February 12;377(9765):578-86.

(11) Sarwar N, Gao P, Seshasai SR et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. *Lancet* 2010 June 26;375(9733):2215-22.

(12) Di AE, Sarwar N, Perry P et al. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA* 2009 November 11;302(18):1993-2000.

(13) Wormser D, Kaptoge S, Di AE et al. Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. *Lancet* 2011 March 26;377(9771):1085-95.

(14) Huxley RR, Woodward M. 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* 2011 August 10.

(15) Teo KK, Ounpuu S, Hawken S et al. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet* 2006 August;%19;368(9536):647-58.

(16) U.S.Department of Health and Human Services. *How tobacco smoke causes disease: the biology and behavioral basis of smoking-attributable disease: a report of the surgeon general*. 2010.

(17) Lubin JH, Alavanja MC, Caporaso N et al. Cigarette smoking and cancer risk: modeling total exposure and intensity. *Am J Epidemiol* 2007 August 15;166(4):479-89.

(18) Samet JM, Thun MJ, de Gonzalez AB. Models of smoking and lung cancer risk: a means to an end. *Epidemiology* 2007 September;18(5):649-51.

(19) Teo KK, Ounpuu S, Hawken S et al. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet* 2006 August;%19;368(9536):647-58.

(20) Kenfield SA, Stampfer MJ, Rosner BA, Colditz GA. Smoking and smoking cessation in relation to mortality in women. *JAMA* 2008 May 7;299(17):2037-47.

(21) Doll R, HILL AB. Mortality in relation to smoking: ten years' observations of British Doctors. *Br Med J* 1964 May 30;1(5395):1399-410.

(22) Cook DG, Shaper AG, Pocock SJ, Kussick SJ. Giving up smoking and the risk of heart attacks. A report from The British Regional Heart Study. *Lancet* 1986 December 13;2(8520):1376-80.

(23) Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ* 2004 June 26;328(7455):1519.

(24) European Union Directorate-General for Health and Consumers. Data and Information on Women's Health in the European Union. 2009.

(25) Thompson S, Kaptoge S, White I, Wood A, Perry P, Danesh J. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. *Int J Epidemiol* 2010 October;39(5):1345-59.

(26) Emerging Risk Factors Collaboration. Lipoprotein(a) concentration and the risk of coronary heart disease, stroke, and nonvascular mortality. *JAMA* 2009 July 22;302(4):412-23.

(27) Emerging Risk Factors Collaboration. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA* 2009 November 11;302(18):1993-2000.

(28) Emerging Risk Factors Collaboration. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *Lancet* 2010 January 9;375(9709):132-40.

(29) Emerging Risk Factors Collaboration. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. *Lancet* 2010 June 26;375(9733):2215-22.

(30) Emerging Risk Factors Collaboration. Diabetes mellitus, fasting glucose, and risk of cause-specific death. *N Engl J Med* 2011 March 2;364(9):829-41.

(31) Emerging Risk Factors Collaboration. Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. *Lancet* 2011 March 11.

(32) Prentice RL, Kalbfleisch JD. Hazard rate models with covariates. *Biometrics* 1979 March;35(1):25-39.

(33) Kirkwood B, Sterne J. Logistic regression: controlling for confounding and other extensions. In: Blackwell Science, editor. *Medical statistics*. 2 ed. Massachusetts, USA: 2003. p. 212-3.

(34) Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011 February 10;342.

(35) Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002 June 15;21(11):1539-58.

(36) Jackson D, White IR, Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Stat Med* 2009 April 30.

(37) Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med* 1991 July;10(7):1025-35.

(38) Whitlock G, Lewington S, Sherliker P et al. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* 2009 March 28;373(9669):1083-96.

(39) West Midlands Public Health Observatory. MetaData; European Standard Population. 2011.

(40) Thun MJ, Day-Lally C, Myers D et al. Trends in tobacco smoking and mortality from cigarette use in Cancer Prevention Studies I (1959 through 1965) and II (1982 through 1988). Changes in Cigarette-Related Disease Risks and Their Implication for Prevention and Control.Smoking and Tobacco Control Monograph No.8.Rockville (MD) [*a*], 305-382. 1997. Shopland DR, Burns DM, Garfinkel L, Samet JM.

(41) Thun MJ, Apicella LF, Henley SJ. Smoking vs other risk factors as the cause of smoking-attributable deaths: confounding in the courtroom. *JAMA* 2000 August 9;284(6):706-12.

(42) Doll R, Peto R. Mortality in relation to smoking: 20 years' observations on male British doctors. *Br Med J* 1976 December 25;2(6051):1525-36.

(43) Last J. *A Dictionary of Epidemiology*. 4th edition ed. Oxford University Press; 2001.

(44) Symons JD, Stebbins CL. Hemodynamic and regional blood flow responses to nicotine at rest and during exercise. *Med Sci Sports Exerc* 1996 April;28(4):457-67.

(45) Michael PR. Cigarette smoking, endothelial injury and cardiovascular disease. *Int J Exp Pathol* 2000 August;81(4):219-30.

(46) Burke A, Fitzgerald GA. Oxidative stress and smoking-induced vascular injury. *Prog Cardiovasc Dis* 2003 July;46(1):79-90.

(47) Cornuz J, Sidoti PC, Tevaearai H, Egger M. Risk factors for asymptomatic abdominal aortic aneurysm: systematic review and meta-analysis of population-based screening studies. *Eur J Public Health* 2004 December;14(4):343-9.

(48) Lowe GD. Common risk factors for both arterial and venous thrombosis. *Br J Haematol* 2008 March;140(5):488-95.

(49) Hoffman RM, Psaty BM, Kronmal RA. Modifiable risk factors for incident heart failure in the coronary artery surgery study. *Arch Intern Med* 1994 February 28;154(4):417-23.

(50) Metz L, Waters DD. Implications of cigarette smoking for the management of patients with acute coronary syndromes. *Prog Cardiovasc Dis* 2003 July;46(1):1-9.

(51) Kurth T, Kase CS, Berger K, Schaeffner ES, Buring JE, Gaziano JM. Smoking and the risk of hemorrhagic stroke in men. *Stroke* 2003 May;34(5):1151-5.

(52) Juvela S, Poussa K, Porras M. Factors affecting formation and growth of intracranial aneurysms: a long-term follow-up study. *Stroke* 2001 February;32(2):485-91.

(53) Howard G, Wagenknecht LE, Burke GL et al. Cigarette smoking and progression of atherosclerosis: The Atherosclerosis Risk in Communities (ARIC) Study. *JAMA* 1998 January 14;279(2):119-24.

(54) Omvik P. How smoking affects blood pressure. *Blood Press* 1996 March;5(2):71-7.

(55) Price JF, Mowbray PI, Lee AJ, Rumley A, Lowe GD, Fowkes FG. Relationship between smoking and cardiovascular risk factors in the development of peripheral arterial disease and coronary artery disease: Edinburgh Artery Study. *Eur Heart J* 1999 March;20(5):344-53.

(56) Celermajer DS, Adams MR, Clarkson P et al. Passive smoking and impaired endothelium-dependent arterial dilatation in healthy young adults. *N Engl J Med* 1996 January 18;334(3):150-4.

(57) Caraballo RS, Giovino GA, Pechacek TF et al. Racial and ethnic differences in serum cotinine levels of cigarette smokers: Third National Health and Nutrition Examination Survey, 1988-1991. *JAMA* 1998 July 8;280(2):135-9.

(58) Nawrot TS, Perez L, Kunzli N, Munters E, Nemery B. Public health importance of triggers of myocardial infarction: a comparative risk assessment. *Lancet* 2011 February 26;377(9767):732-40.

(59) White IR, Frost C, Tomar SL. Correcting for measurement error in binary and continuous variables using replicates. *Statistics in Medicine* 2001;20:3441-57.

Chapter 5: Pipe or cigar smoking and the risk of cardiovascular diseases and lung cancer deaths in developed countries

Summary

The effects of types of smoking other than cigarettes have remained relatively little studied due to lower prevalence in Western countries. The large scope of the ERFC provides information on pipe and cigar smoking on a unique scale. In total, 83 studies provided information on at least cigarette smoking status, 22 studies on pipe status and 20 studies on cigar status. The prevalence of current smoking of cigarettes only was 34%, cigar only 8% and pipe only 5% in the subsets of data which provided information on these types of smoking. Cigar and pipe smoking was mainly a male habit. Current cigarette-only smokers experienced a total of 13,077 fatal and non-fatal MI, current pipe-only smokers 712 such events, current cigar-only smokers 356 such events; and current pipe or cigar on top of cigarette smokers, a total of 678 events. All forms of smoking were significantly associated with CVD risk. Compared to never smokers, RRs for all CVD events were 1.31 (1.19; 1.44) for cigar-only smokers, 1.68 (1.56; 1.81) for pipe-only smokers; 1.97 (1.85; 2.11) for cigarette-only users and 1.95 (1.81; 2.10) for individuals who were secondary pipe or cigar smokers (and also smoked cigarettes). All these RRs were higher than the risk of past smokers, which was 1.09 (1.05; 1.13) compared to never smokers. I observed a nonlinear dose response relationship between amount of pipe and cigar smoking with the risk of MI and all CVD, with strongest increases in risk experienced at low smoking intensities. The strength of association of pipe or cigar smoking with risk of CVD was not affected by adjustment for conventional lifestyle and biochemical risk factors of CVD. These novel findings represent the first demonstration of the effects of pipe or cigar smoking on the risk of major cardiovascular events, using large scale prospective cohort studies and after appropriate adjustment for cardiovascular risk factors. As cardiovascular diseases remain the main cause of death in developed countries, it is important that public health policies address pipe and cigar smoking in addition to cigarette smoking.

5.1 Background

The prevalence of pipe and cigar smoking has been decreasing over the past 50 years in Western Europe. In 2007, the proportion of British men smoking cigars attained 2%, compared to 34% in 1974, and the number of women smoking cigars was very low and scarcely measurable [1]. In the same year, only 1% of men in the UK said they smoked a pipe, and they were almost all aged 50 and over. In contrast, there is evidence of a rise in cigar smoking in the USA, where they represent an increasingly popular alternative to cigarettes, especially amongst women and younger age groups, to the extent that the proportion of teenagers initiating cigar smoking is presently higher than the proportion initiating cigarettes [2]. Whilst the proportion of the US population declaring themselves as current cigar users decreased by 66% between 1964 and 1992, it has risen by nearly 50% between 1993 and 1996, and continues to increase [3]. In 2008, 5% of American smoked cigars and 0.7% smoked tobacco in pipes, representing respectively 13 million and nearly 2 million individuals[4]. This resurgence of cigar smoking has been attributed to campaigns of glamorization of the image of cigars as a substitute for cigarettes [5]. In a study conducted in the USA, UK, Australia and Canada, O'Connor showed that a quarter of smokers in these countries believe that pipes, cigars or roll-your-own cigarettes are safer than manufactured cigarettes[6].

However, this belief is based on relatively scarce epidemiological evidence, and the relationship between cigar and pipe smoking and CVD risk remains uncertain. Early studies such as the British Doctors Study found a non-significant increased risk when smoking cigar, while cigarette smoking was a highly significant risk factor [7]. In this study, pipe and cigar smokers were defined as those who had never smoked cigarettes. These primary smokers tend not to inhale and so are exposed to a relatively low amount of tar and other harmful constituents of the tobacco smoke compared with cigarette smokers. Nowadays, as most cigar smokers are former cigarette smokers, they are likely to have transferred their inhalation techniques, despite the greater irritancy of the pipe and cigar smoke, and it is plausible that their risk would have reached significant levels. Two more studies based in the USA and conducted in 1960's and 1980's found that cigar smokers face a CHD risk around 1/3 higher than those who never smoked [8,9]. In a study on over 7,000 men in the UK, pipe or cigar smokers combined experienced a 70% higher risk of major CHD events, whether or not they had smoked cigarettes previously, but data was too scarce to disentangle the effect of cigars from that of pipes [10]. In a study in Norway, men who

switched from smoking cigarettes only to smoking pipes only experienced the same risk of death as people who remained cigarette only users, and individual who had only ever smoked pipe had double the risk of dying from all causes compared to never smokers [11].

In this context, the aim of this **Chapter** is to investigate the association between pipe and cigar smoking with CVD risk in Western populations in more detail than has been previously possible. The association with lung cancer is also briefly presented for comparison.

5.2 Methods

5.2.1 The dataset

Study design and statistical methods are detailed in **Chapter 1**. The dataset was restricted to studies which provided information on at least cigarette smoking status. Individuals with missing information on pipe (respectively cigar) smoking status because (1) the study did not ask this information to participants, (2) the study did not provide it to the ERFC coordinating centre or (3) individuals refused to provide the information, were coded as non-current pipe (respectively cigar) smokers (**Table 5.1**). This is a reasonable assumption because cigar and pipe smoking are relatively rare habits in the general population and studies included in the ERFC were prospective cohorts sampled from the general population [12]. In a subsidiary analysis, only studies which provided information on all three types of smoking, namely cigarettes, pipes and cigars were used. All individuals who gave information on whether they were current, ex or never pipe or cigar smokers also provided information on cigarette smoking status. Smoking type was defined as “never smoker”, “ex-smoker”, “current cigarette smoking only”, “current pipe smoking only”, “current cigarette and cigar/pipe smoking”. “Primary pipe or cigar smokers” were defined as individuals who currently smoked pipes or cigars but not cigarettes, “secondary pipe or cigar smokers” as individuals who currently smoked pipes or cigars as well as cigarettes, and “exclusive pipe or cigar smokers” as individuals who currently smoked pipes or cigars and had never smoked cigarettes. Outcomes of interest were fatal and non-fatal MI, fatal and non-fatal cerebrovascular events, all fatal and non-fatal CVD events and fatal lung cancer. Individuals were censored at their first non-fatal MI or cerebrovascular event and at any cause of death.

5.2.2 Statistical methods

Statistical methods are the same as the methods presented in **Chapter 3**. Risk Ratios (RRs) and 95% Confidence Intervals were estimated using Cox proportional hazard models for prospective cohort studies and clinical trials; and using logistic models for nested case-control studies. Models were adjusted at least for age and Body Mass Index and stratified by sex and trial arm where appropriate. For graphical representation, estimates were represented by boxes with sizes inversely proportional to their variances, and confidence intervals were represented using floating absolute variances [13]. Robustness of estimates was tested by progressive adjustment for traditional cardiovascular risk factors, restricting the dataset to studies providing information on these additional variables. To investigate heterogeneity of effects across studies, I^2 value and its 95% confidence interval was calculated for each estimate. Subgroup analysis was done by age group, history of diabetes, use of alcohol, race and geographical region. Interaction between smoking type and other risk factors on raising the risk for CVD were tested using formal tests of interaction. All estimates were computed in two steps. In a first step, adjusted estimates were calculated for each study and in a second step, study-specific estimates were combined using random effect multivariate meta-analyses. Estimates and tests with P-value of significance <0.001 were emphasized.

5.3 Results

5.3.1 Description of the dataset

In total, 83 studies provided information on at least cigarette smoking status, 22 studies on pipe status and 20 studies on cigar status (**Figure 5.1**). Assuming that individuals with missing information on cigar or pipe status were not cigar or pipe users, information was available for (in non-overlapping subsets of data) 235,776 cigarette only smokers, 6,255 pipe only smokers, 4,162 cigar only smokers and 5,875 smokers of pipes or cigars on top of cigarettes. During an average follow-up time of 15.6 years, current cigarettes only smokers experienced a total of 13,077 fatal and non-fatal MI, current pipe only smokers 712 such events, current cigar only smokers 356 such events; and current pipe or cigar on top of cigarette smokers, a total of 678 events. There were 15 additional studies with information on pipe or cigar smoking, including 2,350 current pipe or cigar users, for which information on whether either pipe or cigar was smoked was not available. These individuals

experienced 202 fatal and non-fatal MI events during an average follow-up of 12.5 years.

5.3.2 Baseline characteristics of pipe and cigar smokers

Half of participants who smoked either pipes only or cigars only were past cigarette smokers. More than a third (37%) of individuals who smoked pipes or cigars also declared themselves to be current cigarette smokers. While 44% of current cigarette only smokers were women, the proportions were <1% of pipe and 7% of cigar only smokers (**Table 5.1a**). Pipe and cigar smokers were also overwhelmingly White (respectively 99% and 97%) whilst non White individuals represented 12% of cigarette smokers (6% Asian or Oriental and 4% Black). Pipe and cigar smokers were ~20% more likely to be current alcohol drinkers than cigarette only smokers. A higher proportion of pipe only users had not reached secondary education (36% versus 22% for cigarette users), whilst the level of education of cigar smokers was similar to that of cigarette smokers. Individuals in manual jobs and in services were over-represented amongst cigar smokers (37% and 41% respectively versus 30% and 28% for cigarette smokers).

Levels of conventional risk factors were slightly higher in cigar smokers, and to a lesser extent in pipe smokers (**Table 5.1b**). Compared to cigarette only users, cigar smokers had higher levels of anthropometric (Body Mass Index and Waist to Hip Ratio), hemodynamic (systolic and diastolic blood pressure), lipid (total cholesterol, non HDL-cholesterol, Apolipoprotein B, $\log_e$ triglycerides, $\log_e$ Lp(a)), and some inflammatory markers ($\log_e$ CRP). Pipe only smokers had slightly raised WHR and blood pressure levels, but levels of other risk factors were similar to that of cigarette only smokers.

5.3.3 Pipe, cigar and cigarette smoking in relation to CVD risk

Pipe, cigar and cigarette smoking were all significantly associated with the risk of fatal and non-fatal cardiovascular events (**Figure 5.2**). The highest risk was experienced by cigarette, then pipe, then cigar smokers. Compared to never smokers, RRs for fatal and non-fatal myocardial infarction were 1.35 (1.20, 1.51) for cigar only smokers; 1.84 (1.69, 2.01) for pipe only smokers, 2.11 (1.95, 2.28) for cigarette only smokers. The excess risk of pipe and cigar smokers who also smoked cigarettes was not significantly different from that of cigarette only users: RR of 2.05 (1.87, 2.24) compared to never smokers. Compared to never smokers, RRs for all

CVD events were 1.31 (1.19; 1.44) for cigar only smokers, 1.68 (1.56; 1.81) for pipe only smokers; 1.97 (1.85; 2.11) for cigarette only users and 1.95 (1.81; 2.10) for individuals who were secondary pipe or cigar smokers (and also smoked cigarettes). All these RRs were higher than the risk of past smokers, which was 1.09 (1.05; 1.13) for all CVD events compared to never smokers.

In a subsidiary analysis restricting the data to studies which provided information on all three smoking types to enable a within study only comparison of the association with cigarettes, cigars and pipes, associations were marginally changed. RRs for fatal and non-fatal MI became 2.04 (1.62; 2.58) for cigarette only users, 1.28 (1.10; 1.48) for cigar only users, 1.80 (1.57; 2.07) for pipe only users and 1.79 (1.40; 2.29) for individuals combining cigarettes with pipe or cigar use, compared to never smokers (**Figure 5.3**). Results were also similar to those found previously for all CVD events combined.

The association was similar for current pipe or cigar smokers who were past cigarette smokers and for smokers who had always only smoked cigars or pipes, and it was significantly lower than that of current cigarette smokers (**Figure 5.4**). RRs for fatal and non-fatal MI were 1.56 (1.38; 1.77) for exclusive pipe/cigar smokers and 1.71 (1.56; 1.88) for individuals who had switched from cigarettes to pipes or cigars, compared to never smokers. For all cardiovascular events, RRs were respectively 1.49 (1.34; 1.65) and 1.57 (1.45; 1.70), compared to never smokers.

The association between pipe and cigar smoking differed according to type of cardiovascular event (**Figure 5.5**). It was slightly higher for fatal than for non-fatal myocardial infarction: 1.71 versus 1.57 for primary pipe or cigar smoking and 2.30 versus 1.96 for secondary pipe or cigar smoking (on top of cigarette smoking), compared to never smokers. The risk of MI was also higher than that of cerebrovascular events; which was 1.28 (1.13; 1.46) for primary and 1.95 (1.61; 2.34) for secondary pipe or cigar smokers, compared to never smokers. The association with other types of cardiovascular events, which included aortic aneurysm, heart failure, pulmonary embolism, sudden death and other complication of the heart, was similar in strength to that of myocardial infarction: 1.75 (1.52; 2.02) for primary and 2.27 (189; 2.73) for secondary pipe or cigar smoking. Levels of heterogeneity between studies were low and $I^2 < 50\%$ for all types of smoking and all outcomes studied.

Estimates remained significant after adjusting for systolic blood pressure, history of diabetes, non HDL-cholesterol, HDL-cholesterol and log triglycerides levels (**Table 5.2a**). For the risk of fatal and non-fatal MI, RRs for current pipe or cigar smokers versus never smokers attenuated from 1.58 to 1.53, for current cigarette only from 2.05 to 1.99; and for current cigarette and pipe or cigar from 2.16 to 2.12. Estimates of the risk of all CVD combined were also altered only marginally (**Table 5.2b**). Adjustment for these risk factors explained some of the heterogeneity between studies: I^2 for the risk of MI for current cigarette smokers versus never smokers was reduced from 40 to 21 and from 49 to 33 for the risk of all cardiovascular events. Adjustment for alcohol consumption (categorized as never or past alcohol drinker versus current alcohol drinker), for occupation (categorized as not working, manual, services, office, student and other) and for level of education reached (no schooling, primary, secondary, vocational or university), in subsets which provided information on these covariates, did not significantly modify estimates for the risk of fatal and non-fatal MI and for the risk of all CVD events.

5.3.4 Pipe, cigar and cigarette smoking in relation to the risk of lung cancer death

Significant increases in risk of lung cancer death were also observed for all smoking forms (**Figure 5.6**). The risk was strongest for cigarette (RR: 13.13; 11.72 to 14.71), then pipe (10.05; 8.15 to 12.38), then cigar (6.54; 5.00 to 8.55) smokers, compared to never smokers. Secondary pipe or cigar smokers experienced a risk similar to that of exclusive cigarette users (RR of 13.90 versus 13.13). In a subsidiary analysis where data were restricted to studies which provided information on all three smoking types to enable a within study only comparison of the association, RRs for lung cancer deaths were even higher for cigarette only and pipe only smokers (**Figure 5.7**). Individuals who had switched from cigarettes to pipes or cigars had a RR higher than exclusive pipe or cigar smokers and lower than current cigarette only smokers: 9.77 versus 5.65 and 12.79; but differences were non-significant due to relatively small numbers of events in each sub-category resulting in wide confidence intervals (**Figure 5.8**).

5.3.5 Effects modifications

RRs for MI and all cardiovascular events did not differ significantly by geographical region (Western Europe/ North America / Other) (p-value of interaction: 0.28) (**Figure 5.9 a-b**). The effect of being primary pipe/cigar smoker versus never/ex-smoker was

highly significant amongst White with a RR of 1.41 (1.30; 1.53) and non-significant amongst non-White (0.91; 0.58 to 1.42). There was evidence of a stronger effect of cigarette smoking in younger age groups (RRs of all CVD events of 2.71, 2.05, 1.80 and 1.10 in individuals aged 20-39, 40-59, 60-69 and 70+ years old respectively) and there was a similar trend for primary pipe or cigar smoking (RRs of 3.52, 1.19, 1.00 and 2.01 for the same age categories) but it was not significant for the latest. Smoking cigarettes was associated with the greatest increase in risk amongst non-diabetics (RRs of 1.93 versus 1.51), but there was no such effect modification for primary pipe or cigar smokers (RRs of 1.45 versus 1.46).

5.3.6 Dose-response relationship

There was evidence of a non-linear dose response relationship with pipe and cigar smoking combined for fatal and non-fatal MI, after excluding current cigarette smokers (**Figure 5.10**). Compared to never smokers of any type, primary pipe or cigar smokers experienced RRs of 1.43 (1.29; 1.64) for 1-5 cigarettes equivalent per day, 1.61 (1.40; 1.86) for 5-10 cigarettes equivalent per day, 1.59 (1.27; 1.99) for 10-15 cigarettes equivalent per day and 1.86 (1.54; 2.25) for ≥15 cigarettes equivalent per day. When considering all cardiovascular events combined, the association was slightly higher, especially at low intensity of pipe or cigar smoking. Compared to never smokers, RRs were 1.27 (1.13; 1.43); 1.43 (1.27; 1.61); 1.55 (1.28; 1.87) and 1.73 (1.49; 2.02) for <5, 5-10, 10-15 and ≥15 cigarettes equivalent per day respectively.

5.4 Discussion

My analysis contains novel findings on the association between smoking types other than cigarettes in the Western world. The risk factor profile of cigar only or pipe only smokers was not more favourable than that of cigarette only smokers when comparing to never smokers. For some risk factors such as non-HDL-c and $\log_e$ triglycerides, unadjusted levels were even higher in cigar only smokers than in cigarette smokers. In addition, this analysis was well powered to show, for the first time using large scale prospective cohort studies and appropriate adjustment for cardiovascular risk factors, that both pipe and cigar smoking are associated with the risk of major CVD events. As CVD remains the main cause of death in developed countries, it is important that public health policies address pipe and cigar smoking in addition to cigarette smoking.

5.4.1 Main findings

After adjustment for risk factors, age and body mass index (and stratification for sex and trial arm), the risk for pipe and cigar smokers was significantly higher than that of never and of ex-smokers. Compared to never smokers, RRs for all CVD events were 1.31 (1.19; 1.44) for cigar only smokers, 1.68 (1.56; 1.81) for pipe only smokers; 1.97 (1.85; 2.11) for cigarette only users and 1.95 (1.81; 2.10) for individuals who were secondary pipe or cigar smokers (also smoked cigarettes). My estimate for cigar smoking is in agreement with previous estimates in the US population [14,15]. These findings were also compatible with results on pipe and cigar smoking combined from a study of the British population [16]. Cigar and pipe smoking exert an independent effect on cardiovascular risk which is likely to be similar to the way cigarette smoking acts on the vessel wall. Cigar and pipe smoke have been shown to contain the same toxic and carcinogenic compounds as cigarette smoke [17]. The mainstream smoke from cigars which is drawn into the mouth from the butt end has been shown to contain greater concentration of nicotine and carbon monoxide than the mainstream smoke from cigarettes [18]. Persons who smoke ≥4 cigars per day are exposed to an increased amount of smoke equivalent to the smoke of 10 cigarettes per day [19], and in this analysis I have shown that there is a positive dose-response relationship between pipe and cigar amount and risk of CVD which mimics that with cigarette smoking discussed in **Chapter 4,** and makes it more likely that the relationship between cigar and pipe smoking and CVD risk is causal..

The smoke of pipes or cigars is alkaline and generally not inhaled [20]. Early studies such as the British Doctors Study [21] did not observe an increased risk in cigar and pipe smokers, and this was attributed to the fact that at the time pipe and cigar smokers rarely inhaled and smoked smaller quantities of tobacco than cigarette smokers. In my dataset, 53% of cigar only smokers and 45% of pipe only smokers at baseline declared being ex-cigarette smokers. Individuals who switch from cigarettes to pipes or cigars are more likely to inhale smoke than exclusive pipe or cigar smokers [22]. An earlier study found a risk of death similar in exclusive pipe smokers, in cigarette only smokers and in individuals who switched from cigars to pipes [23]. Here, the CVD risk of exclusive pipe smokers was similar to that of individuals who used to smoke cigarettes whilst a difference was observable on the risk of lung cancer

deaths. This finding warrants further study in the Western population to better understand the role of smoke inhalation in promoting CVD risk.

Associations were not strongly modified upon adjustment for a range of biochemical and lifestyle risk factors of CVD. In the Cancer Prevention Study, Henley & al found a significant increase in risk in exclusive pipe smokers of both coronary heart disease and cerebrovascular diseases compared to never smokers, even after adjusting for a range of confounders including alcohol consumption, educational level, body mass index, occupation and diet [24]. In another study of males from the USA, cigar smoking was associated with a moderate but significant increase in the risk of coronary heart disease after multiple adjustment, but not with cerebrovascular disease [25]. In **Chapter 4** of this thesis, I have shown that the association with smoking was not importantly confounded by medical and behavioural risk factors. The present results show that the association of cigarettes, pipes or cigars with CVD risk is independent of lower socio-economic status, increased alcohol consumption and adverse medical history of smokers.

5.4.2 Strengths and limitation

This study contains several strengths. Despite the fact that data came from studies which differ in design, decade of enrolment, geographical region etc., study specific RRs were comparable and the heterogeneity between studies for RRs with pipes and cigars was non-significant. In a subsidiary analysis, the dataset was restricted to studies providing information on all 3 types of smoking analysed and results were largely unchanged. There was no significant difference in RRs according to region of the world and therefore these results may apply to all developed countries. Due to the large size and long follow-up of my data, it was possible to estimate the risk due to pipes and cigars separately, stratifying by past smoking use of cigarettes and by current complementary cigarette use. Analyses were adjusted for a range of potential confounders in a systematic way for all studies, providing more reliable estimates than those found by pooling literature based estimates in a meta-analysis. In particular, adjustment for conventional biochemical factors (blood pressure, cholesterol and triglycerides level, history of diabetes, alcohol status, education and occupation) did not attenuate the associations.

Despite its strengths, my analysis contains limitations. Individuals with missing information regarding cigar or pipe status were considered non-current cigar and

non-current pipe smokers, introducing measurement error. However, this is unlikely to have led to substantial under-estimation in my estimate, which was in broad agreement with previous estimates, especially for cigarette smoking [26-28]. Pipe and cigar smokers were mainly Men and White, and there was a difference in RRs between White and non-White, even if it was non-significant, indicating that these results may not be generalizable to non-White populations. Pipe and cigar amounts were harmonized by the ERFC data management team and expressed in number of cigarettes equivalent per day using conversion factors (see **Chapter 2**) which may not be appropriate for some studies, introducing measurement error and biasing results. Finally, there were not enough women smoking pipes or cigars to be able to test for a difference in RRs according to gender.

5.5 Conclusion

Cigar smoking confers a moderately higher risk of CVD than never smokers. Pipe smoking confers a risk close to that of cigarette smokers, but as amount and duration of pipe and cigars were not available, it was not possible to find out whether the lower risk experienced by pipe and cigar smoker in the data was a reflection of lower intensity of this type of smoking compared to cigarettes. The resurgence of cigar smoking in the USA over the past decades is a matter of substantial concern. Rather than switching from cigarette to either pipe or cigar, individuals should be encouraged to stop all forms of smoking.

Figure 5.1: Flow chart of study selection in the Emerging Risk Factors Collaboration Dataset

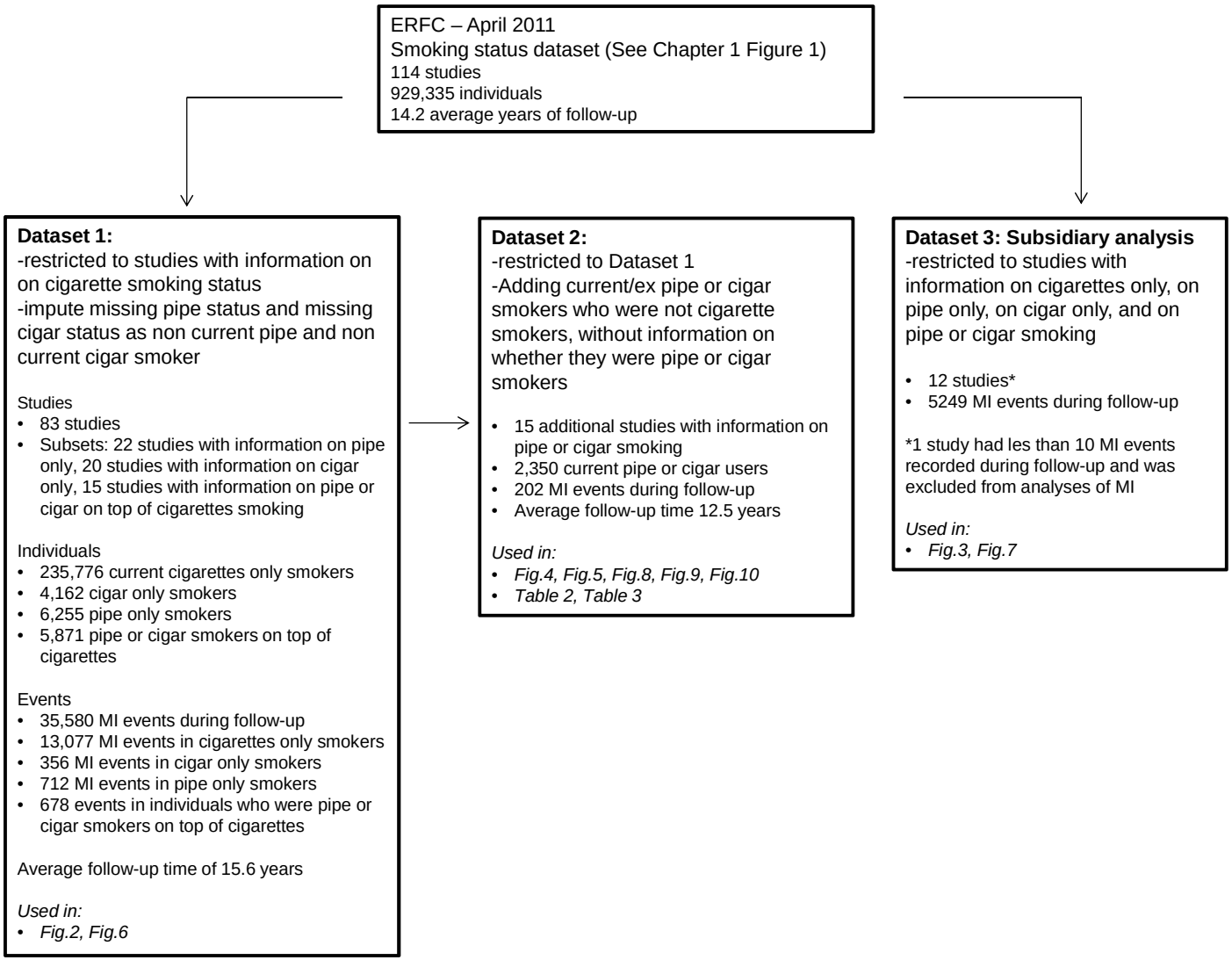
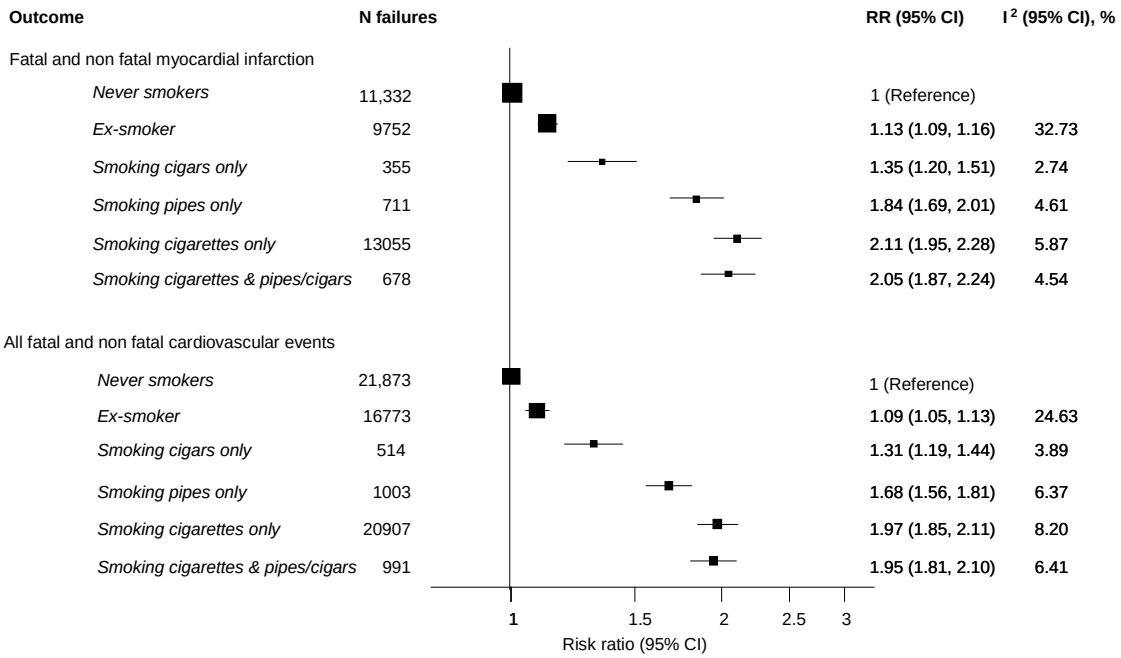
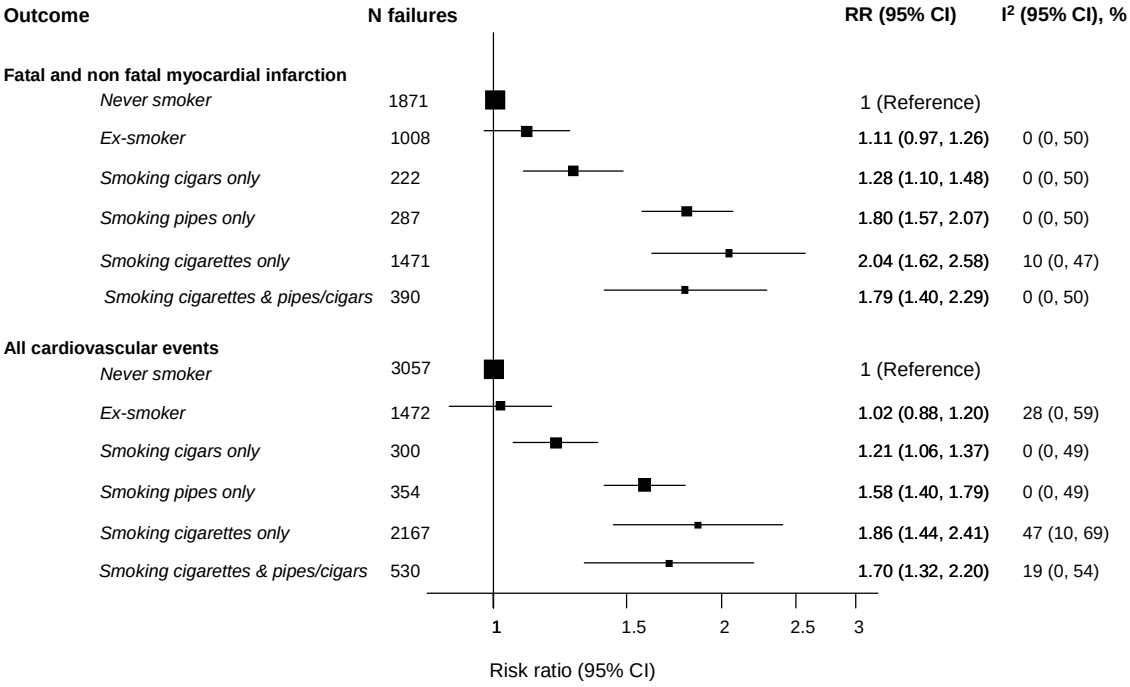


Figure 5.2: Risk ratios for fatal and non-fatal MI, and all cardiovascular events combined, for current and ex cigarette, pipe and cigar smokers versus never smokers



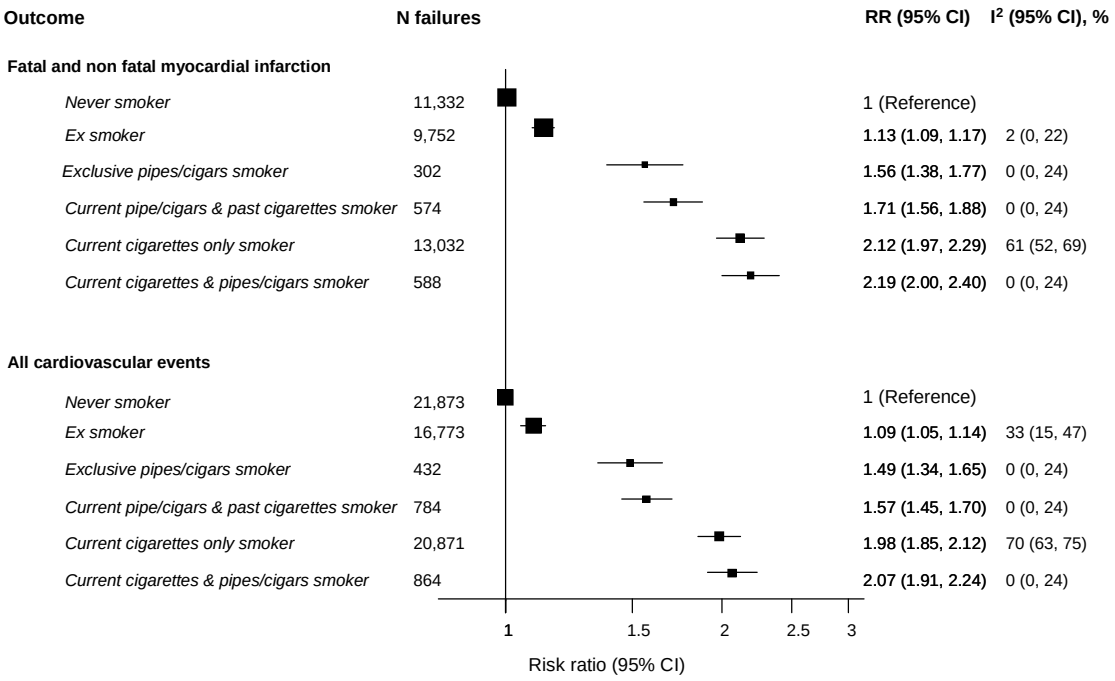
Study-specific risk ratios were adjusted for age at baseline and body mass index and stratified, where appropriate, by sex and trial arm. Studies with fewer than 10 events for a specific outcome were excluded from the analysis of that outcome. Study specific risk ratios were pooled by random effect meta-analysis and boxes are proportional to the inverse of the variance of the estimate, except for the reference group – never smokers – where the box is purely illustrative.

Figure 5.3: Risk ratios for cardiovascular events for current and ex cigarette, pipe and cigar smokers versus never smokers, restricting the dataset to 12 studies (11 studies for fatal and non-fatal MI risk) which provided information on cigarettes, on cigar and on pipe smoking status



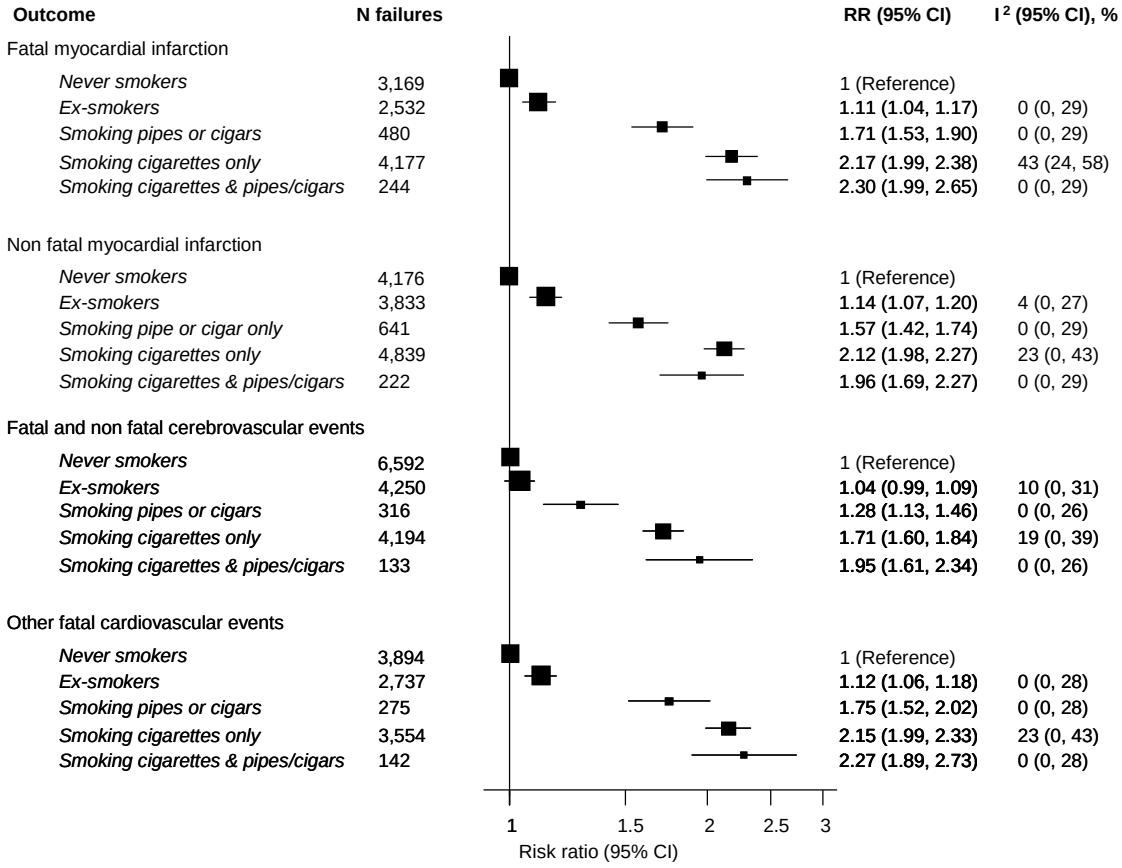
Study-specific risk ratios were adjusted for age at baseline and body mass index and stratified, where appropriate, by sex and trial arm. Studies with fewer than 10 events for a specific outcome were excluded from the analysis of that outcome, leaving 11 studies for the risk of fatal and non-fatal MI and 12 studies for the risk of all CVD events. “All CVD events” includes all non-fatal MI recorded in individual studies as well as all CVD causes of deaths. Study specific risk ratios were pooled by random effect meta-analysis and boxes are proportional to the inverse of the variance of the estimate, except for the reference group – never smokers – where the box is purely illustrative.

Figure 5.4: Risk ratios for cardiovascular events for current and ex cigarettes, pipe and cigar smokers, according to whether they used to smoke cigarettes



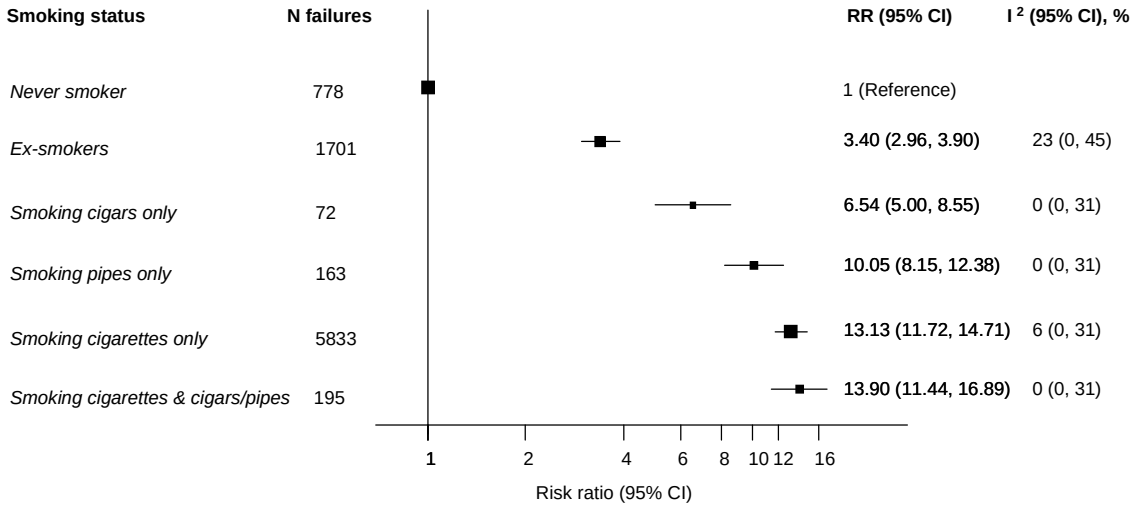
Study-specific risk ratios were adjusted for age at baseline and body mass index and stratified, where appropriate, by sex and trial arm. Studies with fewer than 10 events for a specific outcome were excluded from the analysis of that outcome. Study specific risk ratios were pooled by random effect meta-analysis and boxes are proportional to the inverse of the variance of the estimate, except for the reference group – never smokers – where the box is purely illustrative.

Figure 5.5: Risk ratios for subtypes of cardiovascular diseases for current and ex cigarette and pipe or cigar smokers versus never smokers.



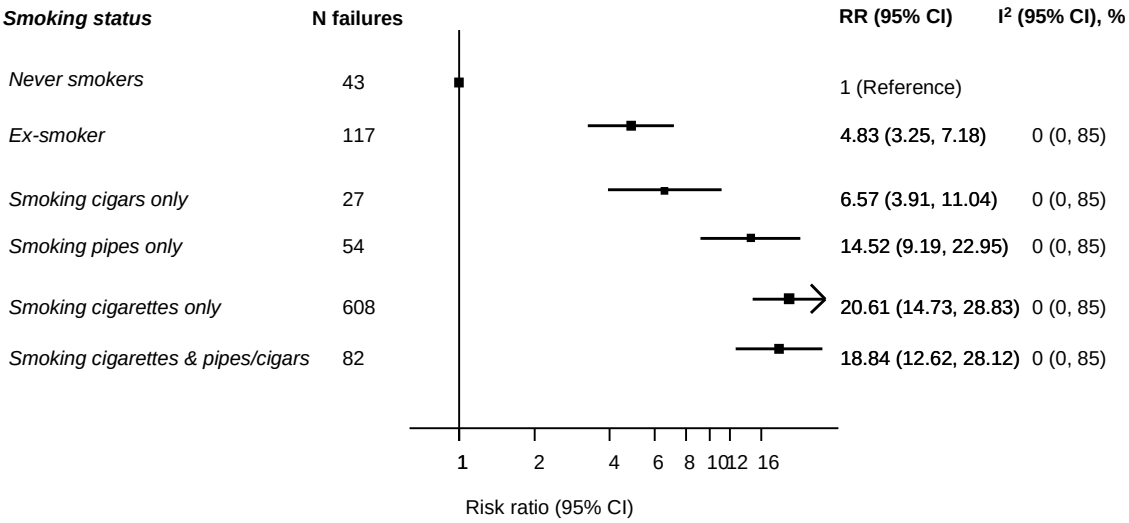
Study-specific risk ratios were adjusted for age at baseline and body mass index and stratified, where appropriate, by sex and trial arm. Studies with fewer than 10 events for a specific outcome were excluded from the analysis of that outcome. Study specific risk ratios were pooled by random effect meta-analysis and boxes are proportional to the inverse of the variance of the estimate, except for the reference group – never smokers – where the box is purely illustrative. Other fatal cardiovascular events included all cardiovascular deaths not attributable to either a myocardial or a cerebrovascular infarction.

Figure 5.6: Risk ratio for fatal lung cancer for current and ex cigarette, pipe and cigar smokers versus never smokers



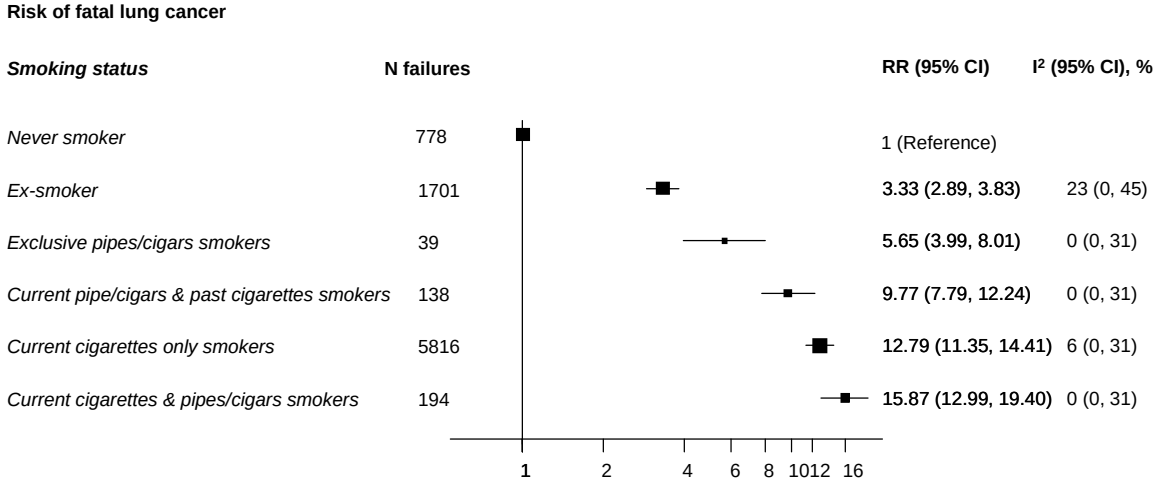
Study-specific risk ratios were adjusted for age at baseline and body mass index and stratified, where appropriate, by sex and trial arm. Studies with fewer than 10 events for a specific outcome were excluded from the analysis of that outcome. Study specific risk ratios were pooled by random effect meta-analysis and boxes are proportional to the inverse of the variance of the estimate, except for the reference group – never smokers – where the box is purely illustrative.

Figure 5.7: Risk ratio for fatal lung cancer for current and ex cigarette, pipe and cigar smokers versus never smokers restricting the dataset to studies which provided information on cigarettes, cigar and pipe smoking status



Study-specific risk ratios were adjusted for age at baseline and body mass index and stratified, where appropriate, by sex and trial arm. Studies with fewer than 10 events for a specific outcome were excluded from the analysis of that outcome, leaving 4 studies for the analysis of the risk of lung cancer deaths. Study specific risk ratios were pooled by random effect meta-analysis and boxes are proportional to the inverse of the variance of the estimate, except for the reference group – never smokers – where the box is purely illustrative.

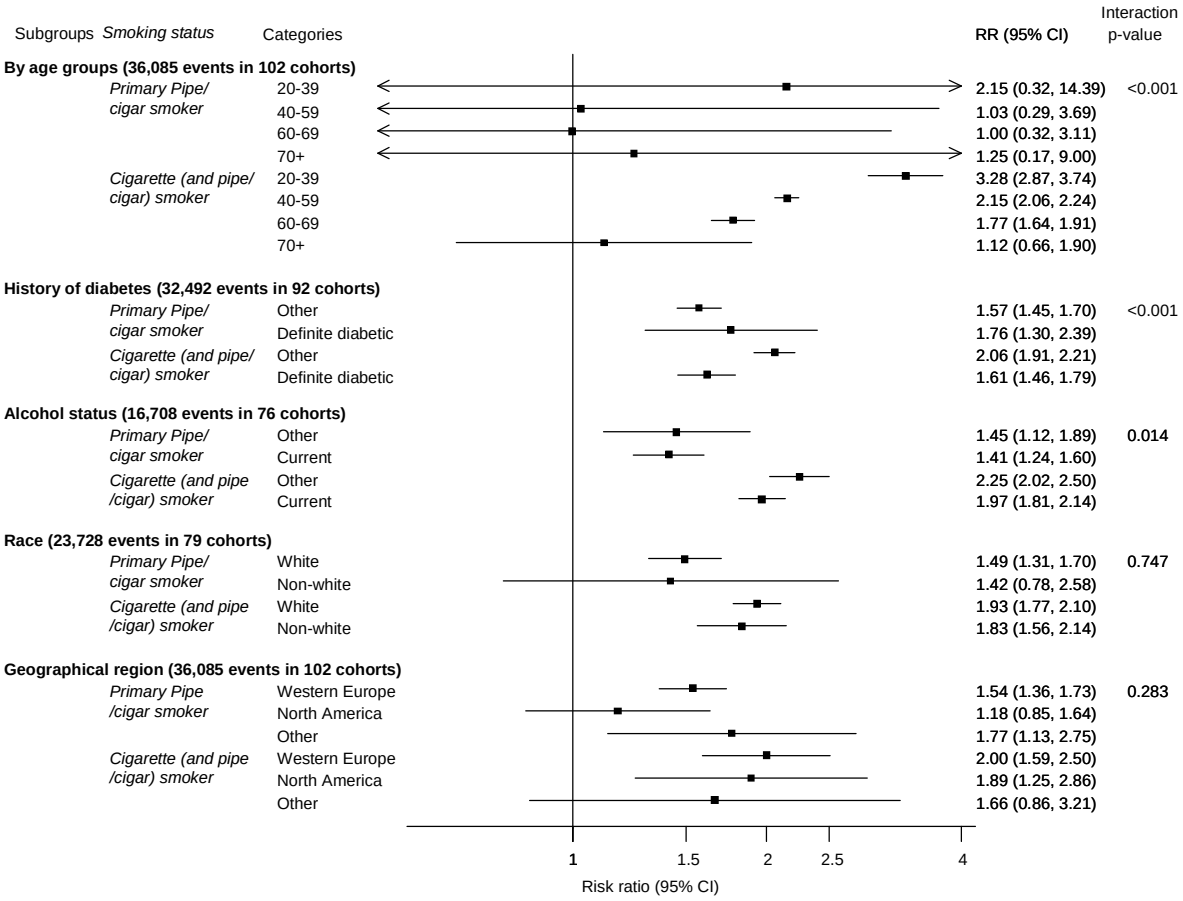
Figure 5.8: Risk ratio for fatal lung cancer for current and ex cigarettes, pipe and cigar smokers according to whether they used to smoke cigarettes



Study-specific risk ratios were adjusted for age at baseline and body mass index and stratified, where appropriate, by sex and trial arm. Studies with fewer than 10 events for a specific outcome were excluded from the analysis of that outcome. Study specific risk ratios were pooled by random effect meta-analysis and boxes are proportional to the inverse of the variance of the estimate, except for the reference group – never smokers – where the box is purely illustrative.

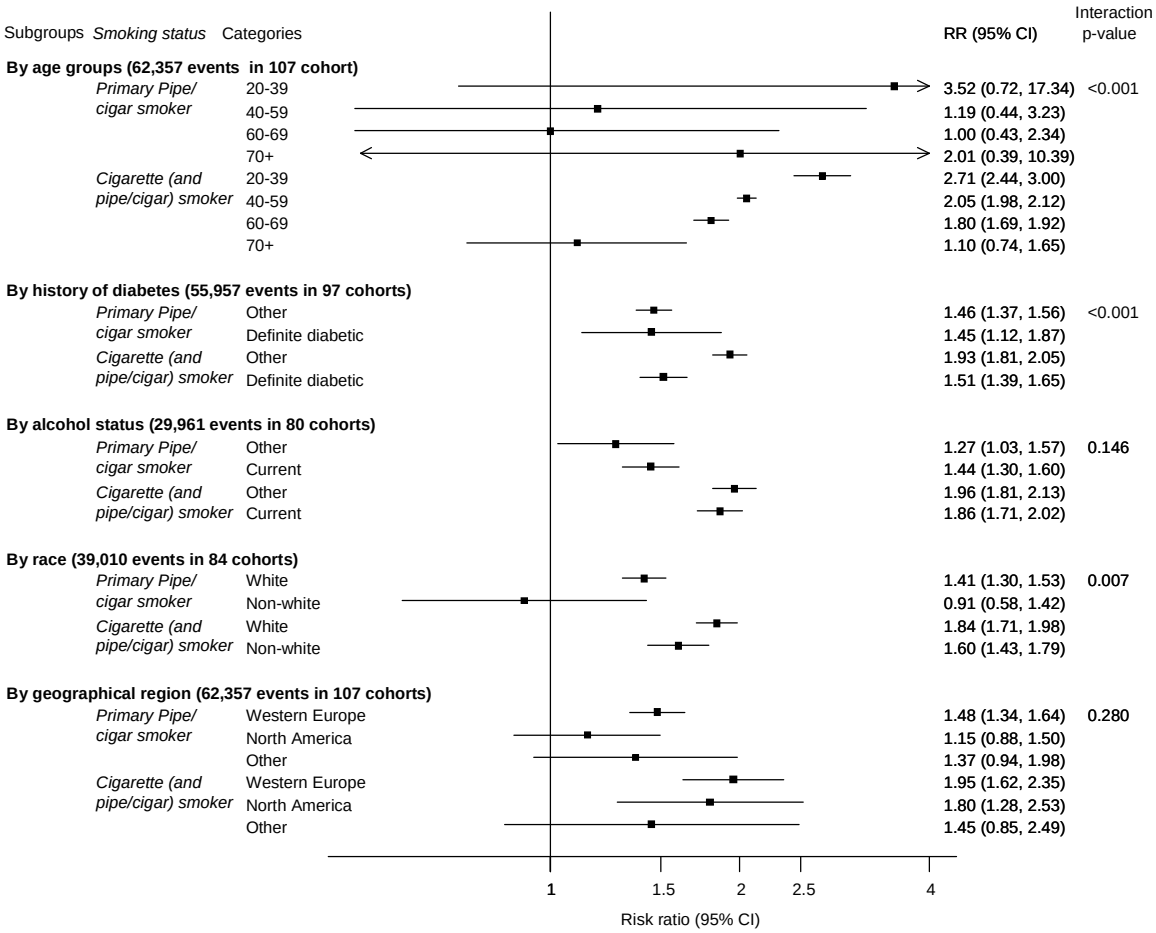
Figure 5.9 a-b: Subgroup analysis

a) Risk ratios for fatal and non-fatal myocardial infarction for current smokers compared to never or ex-smokers



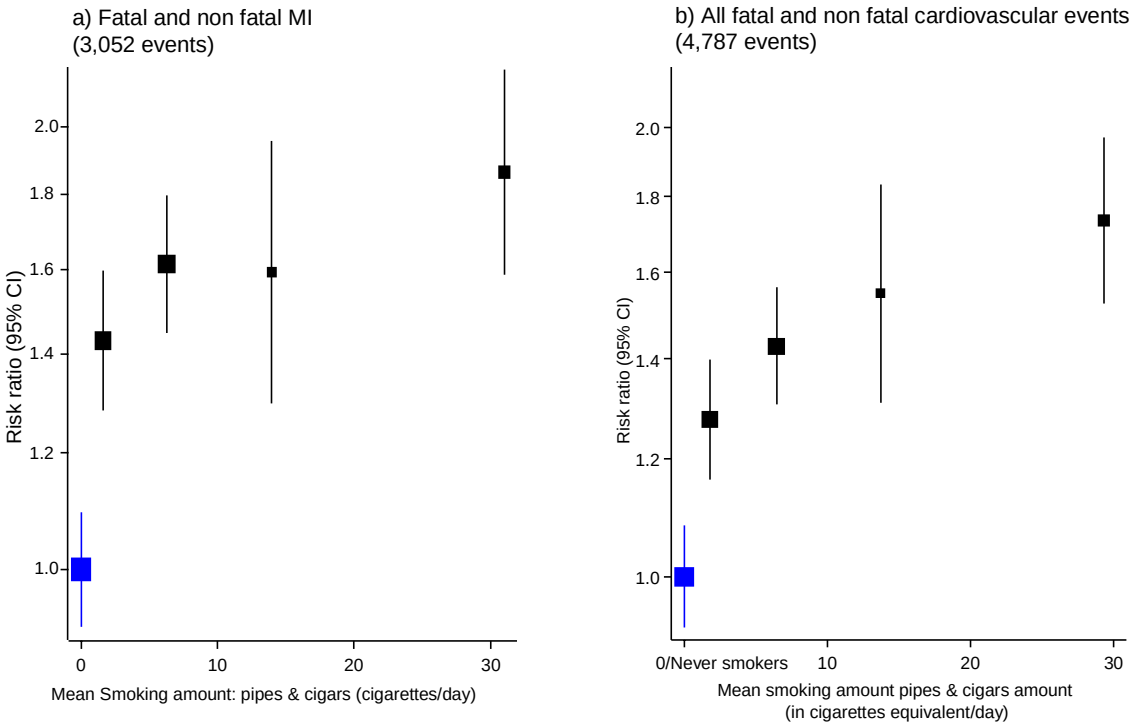
Study-specific adjusted risk ratios were stratified, where appropriate, by sex and trial arm and then pooled across studies by random effects meta-analysis. Studies with fewer than 5 events for a specific outcome were excluded from the analysis of that outcome. P-value of interaction comes from a formal t-test of interaction using age as a continuous variable and a Fisher test of significance of all interaction categories for the other categorical variables. A nominal p-value <0.001 was chosen as indicating presence of a significant effect modifier. For example, in this figure, RRs for current smokers significantly decreased with age and were significantly different according to diabetes status.

b) Risk of all cardiovascular events for current smokers compared to never or ex-smokers



Study-specific adjusted risk ratios were stratified, where appropriate, by sex and trial arm and then pooled across studies by random effects meta-analysis. Studies with fewer than 5 events for a specific outcome were excluded from the analysis of that outcome. P-value of interaction comes from a formal t-test of interaction using age as a continuous variable and a Fisher test of significance of all interaction categories for the other categorical variables. A nominal p-value <0.001 was chosen as indicating presence of a significant effect modifier.

Figure 5.10: Dose-response relationship of pipe and cigar smoking with risk of MI and CVD



Pipe and cigar smoking amount were converted into number of cigarettes equivalent per day (see **Chapter 1**) and categorized as "<5 cigarette equivalent per day", "5-10 cigarettes equivalent per day", "10-15 cigarettes equivalent per day" and ">15 cigarettes equivalent per day" amongst current pipe or cigar smokers who were not also current cigarette smokers. Never smokers (of neither cigarettes nor pies no cigars) were chosen as the reference group. Study-specific risk ratios adjusted for baseline age and Body Mass Index were stratified, where appropriate, by sex and trial arm and then pooled across studies by random effects meta-analysis. Studies with fewer than 5 events were not included in this graph.

Table 5.1: Cross-sectional correlates of pipe and cigar smoking

a) Demographic and lifestyle variables

		Percentage or mean (SD) according to smoking type				
		No of studies*	***No of subjects current smokers****	**Current cigarette only smokers**	**Current pipe only smokers**	**Current cigar only smokers**
Age (years)		*85*	*247,002*	52.89 (8.27)	53.83 (5.95)	51.39 (8.00)
Sex		*85*	*247,002*			
	Male		*144,494*	56%	99%	97%
	Female		*102,508*	44%	1%	7%
Race		*80*	*228,045*			
	White		*133,541*	88%	99%	97%
	Non white		*15,636*	12%	1%	3%
Alcohol status		*64*	*126,520*			
	Not current		*42,819*	35%	18%	11%
	Current		*83,701*	65%	82%	89%
History of diabetes		*78*	*232,102*			
	No		*226,526*	98%	98%	98%
	Yes		*5,576*	2%	2%	2%
Level of education reached		*42*	*97,354*			
	No schooling		*3,570*	4%	1%	1%
	Primary		*17,918*	18%	35%	21%
	Secondary		*51,906*	54%	43%	49%
	Vocational/University		*61,230*	25%	21%	29%
Occupation or job		*40*	*612,230*			
	Not working		*8,900*	15.84	2.36	1.93
	Manual		*18,620*	29.66	33.94	36.92
	Office		*6,012*	9.39	14.93	8.77
	Service		*17,527*	28.2	27.6	40.68
	Student		*77*	0.13	0.1	0.05
	Other		*10,094*	16.78	21.07	11.64

SD: standard deviation. Mean and standard deviation were calculated within studies and pooled across studies using random effects meta-analyses. *With information on cigarettes, cigars, pipe or pipe and cigar smoking.

b) Medical and biochemical variables

			Mean (SD) by smoking type		
	No of studies*	***No of subjects current smokers****	**Current cigarette only smokers**	**Current pipe only smokers**	**Current cigar only smokers**
Anthropometry					
Body Mass Index (kg/m2)	*85*	*247,002*	25.28 (3.77)	25.57 (3.06)	26.15 (3.51)
Waist to Hip Ratio	*35*	*34,490*	0.89 (0.08)	0.95 (0.06)	0.95 (0.09)
Blood pressure					
Systolic blood pressure	*82*	*200,387*	132.54 (17.99)	136.36 (18.1)	135.12 (18.09)
Diastolic blood pressure	*82*	*200,549*	79.73 (10.98)	83.72 (11.32)	83.72 (11.75)
Lipid markers					
Total cholesterol (mmoll/l)	*82*	*198,587*	5.80 (1.14)	5.88 (1.14)	6.03 (1.12)
Non HDL-C (mmol/l)	*68*	*106,894*	4.48 (1.17)	4.57 (1.15)	4.74 (1.18)
HDL-C (mmol/l)	*68*	*107,010*	1.29 (0.37)	1.25 (0.35)	1.25 (0.39)
Loge triglyceride (mmol/l)	*67*	*151,950*	0.37 (0.52)	0.48 (0.49)	0.49 (0.52)
Apo A1 (g/l)	*21*	*24,813*	1.41 (0.27)	1.49 (0.29)	1.46 (0.25)
Apo B (g/l)	*22*	*25,295*	1.12 (0.31)	1.05 (0.26)	1.14 (0.3)
Loge Lp(a)	*20*	*18,349*	2.20 (1.18)	2.45 (1.28)	2.48 (1.23)
Inflammatory markers					
Loge CRP (mg/l)	*34*	*18,870*	0.79 (1.12)	0.75 (1.14)	1.07 (1.12)
Fibrinogen (µmol/l)	*32*	*42,822*	9.63 (2.20)	8.64 (2.11)	8.63 (2.08)
Albumin (g/l)	*28*	*32,989*	43.14 (3.68)	44.04 (3.51)	44.15 (3.68)
Loge leukocyte count(x10^9/l)	*25*	*34,496*	1.98 (0.27)	1.87 (0.24)	1.88 (0.26)
Loge Interleukin 6 (ng/l)	*9*	*3,646*	0.57 (0.62)	0.36 (0.63)	0.21 (0.71)

SD: standard deviation. Mean and standard deviation were calculated within studies and pooled across studies using random effects meta-analyses. *With information on the type of tobacco used.

Table 5.2: Progressive adjustment of RRs of current pipe and cigar smoking compared to never smokers

a) Risk ratios for fatal and non-fatal MI

Progressive adjustment	Current pipe or cigars only smokers			Current cigarettes only			Current cigarettes & cigars/pipes		
	No of events	RR (95% CI)	I^2 (95%CI) %	*No of events*	RR (95% CI)	I^2 (95%CI) %	*No of events*	RR (95% CI)	I^2 (95%CI) %
Age, sex and BMI	240	1.58 (1.33, 1.87)	0 (0, 31)	3882	2.05 (1.84, 2.29)	40 (18, 56)	55	2.16 (1.62, 2.88)	0 (0, 31)
Plus systolic blood pressure	240	1.59 (1.33, 1.89)	0 (0, 31)	3882	2.09 (1.87, 2.34)	39 (17, 56)	55	2.20 (1.65, 2.93)	0 (0, 31)
Plus history of diabetes	240	1.51 (1.26, 1.80)	0 (0, 31)	3882	2.10 (1.91, 2.32)	36 (12, 54)	55	2.21 (1.65, 2.95)	0 (0, 31)
Plus non HDL-cholesterol, HDL-cholesterol and $\log_e$ triglycerides	240	1.43 (1.20, 1.71)	0 (0, 31)	3882	1.99 (1.83, 2.16)	21 (0, 44)	55	2.12 (1.58, 2.83)	0 (0, 31)
Additional adjustment									
Age, sex and BMI	373	1.43 (1.27, 1.61)	0 (0, 28)	6036	2.09 (1.93, 2.25)	46 (30, 59)	129	1.99 (1.65, 2.39)	0 (0, 28)
Plus alcohol consumption	373	1.44 (1.28, 1.62)	0 (0, 28)	6036	2.14 (1.98, 2.31)	47 (30, 59)	129	2.02 (1.67, 2.43)	0 (0, 28)
Age, sex and BMI	659	1.51 (1.33 1.70)	0 (0 52)	1439	2.00 (1.83 2.18)	0 (0 52)	227	1.98 (1.71 2.30)	0 (0 52)
Plus occupation	659	1.67 (0.45 6.24)	0 (0 52)	1439	2.01 (1.84 2.21)	0 (0 52)	227	2.02 (0.37 11.15)	0 (0 52)
Plus education attainment	659	1.70 (0.46 6.28)	0 (0 52)	1439	1.98 (1.77 2.22)	0 (0 52)	227	2.04 (0.37 11.19)	0 (0 52)

Study-specific adjusted risk ratios were stratified, where appropriate, by sex and trial arm and then pooled across studies by random effects meta-analysis. Studies with fewer than 10 events for a specific outcome were excluded from the analysis of that outcome. The dataset was restricted to studies with baseline information on age, sex, BMI and either 1) systolic blood pressure, history of diabetes, non HDL-cholesterol, HDL-cholesterol and $\log_e$ triglycerides; 2) alcohol consumption, or 3) occupation and education attainment. Restricting the data to individuals with baseline information on all additional covariates of adjustment at the same time (rather than in 3 subsets) would have resulted in too small a sample.

b) Risk ratios for all cardiovascular events

	Current pipe or cigars only smokers			Current cigarettes only			Current cigarettes & cigars/pipes		
Progressive adjustment	No of events	RR (95% CI)	I^2 (95%CI) %	*No of events*	RR (95% CI)	I^2 (95%CI) %	*No of events*	RR (95% CI)	I^2 (95%CI) %
Age, sex and BMI	340	1.36 (1.18, 1.56)	0 (0, 29)	6115	1.89 (1.71, 2.09)	49 (33, 62)	84	2.23 (1.76, 2.82)	0 (0, 29)
Plus systolic blood pressure	340	1.38 (1.20, 1.59)	0 (0, 29)	6115	1.91 (1.73, 2.10)	47 (29, 60)	84	2.25 (1.78, 2.85)	0 (0, 29)
Plus history of diabetes	340	1.34 (1.17, 1.54)	0 (0, 29)	6115	1.94 (1.77, 2.12)	45 (26, 59)	84	2.25 (1.78, 2.84)	0 (0, 29)
Plus non HDL-cholesterol, HDL-cholesterol and $\log_e$ triglycerides	340	1.32 (1.16, 1.51)	0 (0, 29)	6115	1.90 (1.76, 2.05)	33 (9, 51)	84	2.19 (1.73, 2.76)	0 (0, 29)
Additional adjustment									
Age, sex and BMI	562	1.44 (1.30, 1.59)	0 (0, 27)	9322	1.98 (1.84, 2.12)	62 (51, 70)	202	2.00 (1.72, 2.33)	0 (0, 27)
Plus alcohol consumption	562	1.44 (1.30, 1.59)	0 (0, 27)	9322	2.00 (1.86, 2.15)	62 (52, 70)	202	2.05 (1.76, 2.38)	0 (0, 27)
Age, sex and BMI	867	1.37 (1.19 1.59)	0 (0 50)	2367	1.90 (1.78 2.04)	0 (0 50)	303	1.92 (1.69 2.18)	0 (0 50)
Plus occupation	867	1.28 (0.47 3.47)	0 (0 50)	2367	1.82 (1.66 1.98)	0 (0 50)	303	2.41 (0.67 8.62)	0 (0 50)
Plus education attainment	867	1.33 (0.47 3.76)	0 (0 50)	2367	1.79 (1.62 1.97)	0 (0 50)	303	2.17 (0.62 7.61)	0 (0 50)

Study-specific adjusted risk ratios were stratified, where appropriate, by sex and trial arm and then pooled across studies by random effects meta-analysis. Studies with fewer than 10 events for a specific outcome were excluded from the analysis of that outcome. The dataset was restricted to studies with baseline information on age, sex, BMI and either 1) systolic blood pressure, history of diabetes, non HDL-cholesterol, HDL-cholesterol and $\log_e$ triglycerides; 2) alcohol consumption, or 3) occupation and education attainment. Restricting the data to individuals with baseline information on all additional covariates of adjustment at the same time (rather than in 3 subsets) would have resulted in too small a sample.

References

1. Office for National Statistics. Smoking and drinking among adults, 2007. General Household Survey. 2007.

2. Delnevo CD. Smokers' choice: what explains the steady growth of cigar use in the U.S.? *Public Health Rep.* 2006;121:116-119.

3. Jacobs EJ, Thun MJ, Apicella LF. Cigar smoking and death from coronary heart disease in a prospective study of US men. *Arch Intern Med.* 1999;159:2413-2418.

4. National Institute on drug abuse. NIDA InfoFacts: Cigarettes and other tobacco products. [http://www.nida.nih.gov/infofacts/tobacco.html#References], Accessed on 01/06/2010.

5. Baker F, Ainsworth SR, Dye JT et al. Health risks associated with cigar smoking. *Stat Med.* 2002;21:1539-1558.

6. O'Connor RJ, McNeill A, Borland R et al. Smokers' beliefs about the relative safety of other tobacco products: findings from the ITC collaboration. *Nicotine Tob Res.* 2007;9:1033-1042.

7. Doll R, HILL AB. The mortality of doctors in relation to their smoking habits: a preliminary report. *BMJ.* 1954;328:1529-1533.

8. Iribarren C, Tekawa IS, Sidney S, Friedman GD. Effect of cigar smoking on the risk of cardiovascular disease, chronic obstructive pulmonary disease, and cancer in men. *N Engl J Med.* 1999;340:1773-1780.

9. Jacobs EJ, Thun MJ, Apicella LF. Cigar smoking and death from coronary heart disease in a prospective study of US men. *Arch Intern Med.* 1999;159:2413-2418.

10. Shaper AG, Wannamethee SG, Walker M. Pipe and cigar smoking and major cardiovascular events, cancer incidence and all-cause mortality in middle-aged British men. *Int J Epidemiol.* 2003;32:802-808.

11. Tverdal A, Bjartveit K. Health consequences of pipe versus cigarette smoking. *Tob Control.* 2011;20:123-130.

12. Baker F, Ainsworth SR, Dye JT et al. Health risks associated with cigar smoking. *Stat Med.* 2002;21:1539-1558.

13. Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med.* 1991;10:1025-1035.

14. Jacobs EJ, Thun MJ, Apicella LF. Cigar smoking and death from coronary heart disease in a prospective study of US men. *Arch Intern Med.* 1999;159:2413-2418.

15. Iribarren C, Tekawa IS, Sidney S, Friedman GD. Effect of cigar smoking on the risk of cardiovascular disease, chronic obstructive pulmonary disease, and cancer in men. *N Engl J Med.* 1999;340:1773-1780.

16. Shaper AG, Wannamethee SG, Walker M. Pipe and cigar smoking and major cardiovascular events, cancer incidence and all-cause mortality in middle-aged British men. *Int J Epidemiol.* 2003;32:802-808.

17. Pechacek TF, Folsom AR, de GR et al. Smoke exposure in pipe and cigar smokers. Serum thiocyanate measures. *JAMA.* 1985;20;254:3330-3332.

18. Brunnemann KD, Hoffmann D. Chemical studies on tobacco smoke. XXIV. A quantitative method for carbon monoxide and carbon dioxide in cigarette and cigar smoke. *J Chromatogr Sci.* 1974;12:70-75.

19. Iribarren C, Tekawa IS, Sidney S, Friedman GD. Effect of cigar smoking on the risk of cardiovascular disease, chronic obstructive pulmonary disease, and cancer in men. *N Engl J Med.* 1999;340:1773-1780.

20. U.S.Department of Health and Human Services. *How tobacco smoke causes disease: the biology and behavioral basis of smoking-attributable disease: a report of the surgeon general*. 2010.

21. Doll R, Peto R. Mortality in relation to smoking: 20 years' observations on male British doctors. *Br Med J.* 1976;2:1525-1536.

22. Baker F, Ainsworth SR, Dye JT et al. Health risks associated with cigar smoking. *Stat Med.* 2002;21:1539-1558.

23. Tverdal A, Bjartveit K. Health consequences of pipe versus cigarette smoking. *Tob Control.* 2011;20:123-130.

24. Henley SJ, Thun MJ, Chao A, Calle EE. Association between exclusive pipe smoking and mortality from cancer and other diseases. *J Natl Cancer Inst.* 2004;96:853-861.

25. Iribarren C, Tekawa IS, Sidney S, Friedman GD. Effect of cigar smoking on the risk of cardiovascular disease, chronic obstructive pulmonary disease, and cancer in men. *N Engl J Med.* 1999;340:1773-1780.

26. Henley SJ, Thun MJ, Chao A, Calle EE. Association between exclusive pipe smoking and mortality from cancer and other diseases. *J Natl Cancer Inst.* 2004;96:853-861.

27. Baker F, Ainsworth SR, Dye JT et al. Health risks associated with cigar smoking. *Stat Med.* 2002;21:1539-1558.

28. Thun MJ, Apicella LF, Henley SJ. Smoking vs other risk factors as the cause of smoking-attributable deaths: confounding in the courtroom. *JAMA.* 2000;284:706-712.

Section B: Smoking and smokeless tobacco in relation to the risk of cardiovascular diseases in a developing country

Chapter 6: Description of the Pakistan Risk of Myocardial Infarction Study

Summary

The Pakistan Risk of Myocardial Infarction Study (PROMIS) is an ongoing case-control study collecting genetic, lifestyle and other determinants of myocardial infarction in South Asia. In particular, detailed questionnaire based information is available for smoking and smokeless forms of tobacco use. By March 2011, the PROMIS had enrolled 7,905 first ever MI patients and 7,458 age and sex frequency matched controls. After exclusion of individuals with missing information on covariates of adjustment, 6,871 controls and 6,050 cases were included in the analysis. Participants were recruited from six different urban hospitals located across Pakistan and sampled from all major Pakistani ethnic groups. PROMIS will provide a unique insight into the effect of smoking and smokeless tobacco on the risk of myocardial infarction in a developing country setting, bearing in mind the limitations inherent to the case-control study design.

6.1 Background of PROMIS

Over the next decade, the burden of CVD is projected to increase the most rapidly in South Asia. In a country like Pakistan, it threatens to counter major health gains achieved during the past half century, translated into the prolongation of life expectancy from 50 to 67 years between 1960 and 2010 [1]. CHD is currently responsible for 14% of all deaths in men and 12% in women in Pakistan (**Figure 6.1**), accounting for 150,000 deaths a year, representing a substantial burden for a middle low-income country of 185 million people [2]. In addition, several million people of Pakistani or other South Asian descent who live in developed countries tend to have striking excesses of vascular diseases [3] and vascular disease mortality [4,5] compared with natives of European ancestry. Despite the highlighted need for large scale epidemiological studies [6], previous studies on South Asians have been characterized by small sample sizes [7,8], conducted amongst South Asians living abroad [9] and have not taken into account ethnicity-specific differences within the South Asian subcontinent [10-12].

In this context, the objective of the PROMIS has been to establish an epidemiological and genetic resource to enable reliable study of a range of determinants of MI in South Asia, including genetic, lifestyle and other determinants [13] (http://www.cncdpk.com/projects/the-pakistan-risk-of-myocardial-infarction-study-promis.html). The PROMIS has established 6 centres of recruitment in five urban centres of Pakistan: Karachi, Hyderabad, Lahore, Multan and Faisalabad (**Figure 6.2**). Recruitment has been ongoing since 2005 and aims to achieve a total of 20,000 MI cases and 20,000 controls, at which stage it will represent the largest scale epidemiological resource for the study of CHD in South Asia. This will enable more detailed assessment than previously possible of the relationship between risk factors and MI risk in Pakistan. For some lifestyle habits such as tobacco uses, which are relatively similar across the South Asian subcontinent, findings from PROMIS may be generalizable to other South Asian populations. It may also be informative for Western populations with a South Asian origin who have retained some lifestyle characteristics and a similar genetic background to Pakistanis. Comparing findings of the PROMIS with those of studies conducted with Pakistani migrants settled in Western countries could also shed light on the relative importance of genetic versus behavioural and socio-economic risk factors, assuming Pakistani migrants are similar genetically to Pakistanis and only differ in their acquisition of Western lifestyle habits.

As recruitment centres are located across Pakistan, major ethnic groups are all represented within the PROMIS (**Table 6.1**). In Pakistan, major ethnic groups are Punjabi (45% of population), Pathan (15%), Sindhi (14%), Sariaki (8%), Muhajirs (also called Urdu) (8%), Balochi (4%) and other (6%) [14]. Punjabis originate from the Punjab region. Pathans are also called Pashtuns and are an indigenous group from the land located south of the Hindu Kush in Afghanistan and West of the Indus River in Pakistan. Sindhis are ancient people principally inhabiting the province of Sindh in Pakistan. Baluch originate from the Baluchistan province. Muhajirs, also referred as Urdu because of the language they speak, are a multi-ethnic group of Muslims who emigrated after the partition of Pakistan from India. The Saraikis maintain a separate language and culture, but are ethnically the same as the natives of Sindh and Punjab.

6.2 Methods

6.2.1 Inclusion criteria

Patients were eligible for inclusion as MI cases if they were between 20 and 80 years old and admitted to emergency rooms of participating hospitals fulfilling all the following criteria: (1) sustained symptoms suggestive of MI lasting longer than 20 minutes within the previous 24 hours; (2) ECG changes indicative of MI (i.e., new pathologic Q waves, at least 1mm ST elevation in any 2 or more contiguous limb leads, new left bundle branch block, new persistent ST-T wave changes diagnostic of non-Q wave MI); (3) positive troponin test; (4) no previous cardiovascular disease, defined as self-reported history of angina, MI, coronary revascularisation, transient ischaemic attack, stroke, or evidence of CHD on prior ECG or in other medical records; and (5) absence of cardiogenic shock.

Controls were individuals identified concurrently in the same hospitals as index cases frequency matched to cases by sex and by 5 year age bands, without a self-reported history of CVD and no ECG changes consistent with a previous MI. Controls were recruited in the following order of priority: (1) visitors of patients attending the outpatient department; (2) patients attending outpatient departments for routine non-cardiac complaints; (3) non-blood related visitors of index MI cases.

Participants were not recruited if they presented with (1) a history of viral or bacterial infection in the past two weeks; (2) a documented chronic condition such as malignancy, inflammatory disorder, hepatitis or renal failure; (3) pregnancy; (4) refused to participate. Institutional review boards provided approval and participants

gave informed consent for use of samples in genetic, biochemical and other analyses.

6.2.2 Measurement of lifestyle variables

Each PROMIS participant was administrated a questionnaire with detailed behavioural and lifestyle information questions, including smoking, diet, socio-economic status, ethnic group, medical and family history, physical activity and anthropometry (**List 6.1**). The questionnaire was locally relevant, recording local linguistic and ethnic groups (Urdu, Punjabi, Pathan, Balochi, Sindhi, Memon, Gujrati and others) and regional types of tobacco use (cigarettes, *bidis*, *huqqa*, *chilum*, *naswar*, *paan*, *supari*, *gutka*, *chalia*) (**List 6.2**). Individuals using smokeless products without adding tobacco ("non *tambako*" users) were considered as non users. For the purpose of the analysis, *paan*, *gutka* and *supari* users were grouped under the term "chew tobacco" and individuals using cigarettes, bidis, hookah or chilum as "smoke tobacco". The use of *naswar* was termed "snuff dipping". "Smokeless tobacco" includes both chewing and dipping tobacco. To assess dietary habits, participants completed a food-frequency questionnaire tailored to Pakistani diet, which had been developed with a local nutritionist after a pilot 24h recall study on a subset of 100 participants. For MI cases, questionnaires were administered after medical stabilization and related to habits and characteristics before the diagnosis of acute MI.

6.2.3 Measurement of biochemical information, weight and height

Using standardized procedures and equipment, research officers obtained measurements of height, weight, waist and hip circumference, systolic and diastolic blood pressure, and heart rate. Non-fasting serum, EDTA plasma, and whole blood samples were obtained while participants were recumbent at about 45^0, centrifuged within 45 minutes, and stored long-term at -80°C. In MI cases, blood samples were drawn within 24 hours of the onset of symptoms and prior to administration of any thrombolytic agents. Time since onset of symptoms and since last meal was recorded. Total cholesterol and high-density lipoprotein cholesterol were measured using a homogeneous enzymatic colorimetric method. Low-density lipoprotein cholesterol was measured using the same techniques or computed using the Friedewald formula.

6.3 The dataset

6.3.1 Descriptive characteristics

The PROMIS had recruited 7905 MI patients and 7458 controls by May 2011 from 6 different hospitals located in 5 urban centres of Pakistan: Karachi, Hyderabad, Lahore, Multan and Faisalabad. One centre, Faisalabad, had only started recruiting when this analysis was performed and so 11 cases and 24 controls from this centre were excluded from the analyses. For the purpose of the analysis, individuals with missing information on age, sex, recruitment centre, history of diabetes and hypertension, family history of MI, LDL-c and waist to hip ratio, were excluded, leaving 6,051 cases and 6,871 controls. All participants were recruited from large urban centres and the majority of participants came from Karachi or Lahore. All individuals reporting the use of *hookah* or *chilum* were also smokers of either cigarettes or *bidis*.

Cases were mostly men, accounting for 83% of all cases (**Table 6.2**). Mean age was 53.4 years old (SD: 10.3). The proportion of self-reported diabetics was 22%, self-reported hypertensive individuals: 49%, and individuals with a family history of MI: 21%. Average total cholesterol levels were: 5.07 (SD: 1.37), and average levels of LDL-C levels were: 3.3 (SD: 1.18). Waist to hip ratio had a mean of 0.959 (SD: 0.056). Average number of years of education was 8 years (IQR: 0-12) and average monthly income was 15,000 Pakistani rupees (IQR: 10,000-25,000). 27% of cases declared cooking with ghee only, 29% with a combination of ghee and oil, and the rest with oil only. The proportion of current tobacco users in cases was 18% of women and 59% of men; and the proportion of ex-tobacco users was 6%.

Frequency matching by sex was imperfect and men represented 78% of controls. Mean age of controls was 53.0 years old (SD: 9.5). Overall, 60% of controls had never used tobacco, 7% were ex-tobacco users and 32% current tobacco users. The main centre of recruitment was the National Institute of Cardiovascular Diseases located in Karachi, where 48% of controls and 39% of cases came from. As Karachi is located in an area predominantly populated by the Urdu ethnic-linguistic group, this ethnicity was the most represented in my data (39% of controls) (**Table 6.1**). Out of all controls, 14% declared having diabetes (975 individuals), 28% being hypertensive (1901 individuals), and 15% had a family history of MI (1045 individuals). Average total cholesterol levels were: 4.64 (SD: 1.37), and average levels of LDL-C levels: 2.87 (SD: 1.06). Waist to hip ratio had a mean of 0.942 (SD: 0.066). Average number

of years of education was 8 years (IQR: 0-12) and average monthly income was 12,000 Pakistani rupees (IQR: 8000-22,000). 23% of controls declared cooking with ghee only, 27% with a combination of ghee and oil, and the rest with oil only. The proportion of current tobacco users was 13% of women and 38% men, and the proportion of ex-tobacco users was 7% of controls.

6.3.2 Principal component analysis of diet and socio-economic status

The principal component analysis on education, monthly income, ownership, and employment status identified one main component labelled "socio-economic status" and explained 14% of the variance (**Appendix 2**). Two dietary patterns were identified by a principal component analysis of the food frequency questionnaire and explained nearly 15% of the variance (**Appendix 3**). The first pattern was labelled "vegetables and carbohydrate" based diet and the second pattern was labelled as "meat, fish and sweet" based diet.

6.4 Use of the dataset in later Chapters

The analysis is restricted to controls in **Chapter 7** for the investigation of the prevalence of tobacco uses and its relationship to other CVD risk factors in Pakistanis. In **Chapter 8**, cases and controls are included in the analysis which aims to quantify the risk of tobacco users compared to non-tobacco users, taking into account other relevant risk factors for CVD.

Figure 6.1: Estimated proportional mortality (%) in Pakistan, 2004

Source: Global Burden of disease [16]

Figure 6.2: Geographical location of PROMIS recruitment centres

City	Recruitment centre	Predominant ethnicity
Lahore	Punjab Institute of Cardiology	Punjabi
Faisalabad	Faisalabad Institute of Cardiology	Punjabi
Multan	Multan Institute of Cardiology	Punjabi
Hyderabad	Red Crescent Hospital	Urdu
Hyderabad	Deewan Mushtaq Institute of Cardiology	Punjabi
Karachi	National Institute of CVD	Urdu
	Karachi Institute of Heart Diseases	Urdu

Map source: U.S. Central Intelligence Agency: *The World Fact book- Pakistan* – accessed on 14/07/2011 [15]

Table 6.1: Number of a) cases and b) controls from the PROMIS per recruitment centre

a) Cases

	Ethnicity								
Recruitment centre	Urdu	Punjabi	Pathan	Balochi	Sindhi	Memon	Gujrati	Others	*Total*
DMIC	36	6	0	2	80	0	0	3	*127*
KIHD	779	99	35	12	26	19	16	39	*1,025*
MIC	112	188	4	2	2	0	0	290	*598*
NICVD	1,142	419	239	59	199	92	32	197	*2,379*
PIC	104	1,217	21	1	1	0	1	44	*1,389*
RCH	218	25	12	0	264	2	3	8	*532*
Total number of cases	*2,391*	*1,954*	*311*	*76*	*572*	*113*	*52*	*581*	*6,050*

b) Controls

	Ethnicity								
Recruitment centre	Urdu	Punjabi	Pathan	Balochi	Sindhi	Memon	Gujrati	Others	*Total*
DMIC	86	3	2	0	66	0	3	1	*161*
KIHD	905	104	54	9	32	8	15	38	*1,165*
MIC	74	148	6	2	0	0	0	185	*415*
NICVD	1,280	582	417	114	519	67	101	246	*3,326*
PIC	55	1,138	18	1	6	0	0	44	*1,262*
RCH	274	34	8	2	208	2	0	14	*542*
Total number of controls	*2,674*	*2,009*	*505*	*128*	*831*	*77*	*119*	*528*	*6,871*

DMIC: Deewan Mushtaq Institute of Cardiology, KIHD: Karachi Institute of Heart Diseases, MIC; Multan Institute of Heart Diseases, NICVD: National Institute of Cardiovascular Diseases, PIC: Punjab Institute of Cardiovascular Disease, RCH: Red Crescent Hospital. Note: Faisalabad Institute of Cardiology which had only recruited 35 individuals by the time of the analysis was excluded from the dataset.

Table 6.2: Descriptive characteristics of PROMIS cases and controls by gender

	Cases % or mean (SD)		**Controls** % or mean (SD)			
	Male (5,037 individuals)	Female (1,013 individuals)	Male (5,359 individuals)	Female (1,512 individuals)	P-value (1)	P-value (2)
Age	52.7 (10.2)	56.6 (9.9)	53.0 (9.7)	53.5 (8.6)	Matched[1]	
History of hypertension	45%	67%	27%	38%	<0.0001	<0.0001
History of diabetes	19%	37%	14%	17%	<0.0001	0.001
Family history of MI	21%	20%	13%	12%	<0.0001	0.531
Tobacco usage					<0.0001	<0.0001
None	41%	82%	62%	87%		
Current smoker	42%	4%	22%	2%		
Current smokeless user	10%	14%	12%	10%		
Current both	7%	0%	4%	1%		
Cooking medium used					<0.0001	0.001
Ghee only	28%	23%	22%	27%		
Oil only	44%	28%	51%	49%		
LDL-C (mmol/l)	3.29 (1.68)	3.38 (2.85)	2.84 (1.64)	2.96 (2.67)	<0.0001	0.015
Waist to Hip ratio	0.961 (0.090)	0.947 (0.153)	0.945 (0.088)	0.933 (0.143)	<0.0001	<0.0001

(1): P-value comparing cases versus controls obtained fitting a linear regression for continuous outcomes and a logistic (multinomial) for categorical outcomes adjusting for age, sex, ethnicity and centre. (2): P-value comparing men versus women in controls only obtained fitting a linear regression for continuous outcomes and a logistic (multinomial) for categorical outcomes adjusting for age, ethnicity and centre. Means for LDL-C, and WHR were adjusted for age equal 53 (the mean age in controls), ethnicity and centre

List 6.1: Summary of questionnaire-based information collected in the Pakistan Risk of Myocardial Infarction Study

Type of information	Detail recorded
Symptoms, arrival and management at the hospital	Date of MI, time of onset, time since last meal, review of symptoms, diagnosis and management at the hospital and any investigations ordered by the attending physician
Medication and medical history	Medication class and duration of use for each of anti-coagulant, blood pressure- and lipid-lowering, diabetes related, hormonal and, for women, contraceptive and HRT medications
Female reproductive history	Age at first menstrual period, age periods stopped, hormones for menopause treatment, oral contraceptive use
Family history	Approximate age of diagnosis/occurrence of hypertension, diabetes, angina, MI, stroke, cancer or sudden death, for each of mother, father, sister, brother, son and daughter
Ethnicity and other genetic related information	Place of birth, personal and parental ethnicity (e.g, Urdu, Punjabi, Pathan, Balooch, Sindhi, Memon, Gujrati), parental co-sanguinity
Socio-economic status	Occupation(s), monthly income, level and duration of education, marital status, no. of dependants, ownership/wealth
Physical activity	Type, frequency, duration and intensity of activity for each of occupational, work related commuting and leisure time
Tobacco consumption	See Table 2 for details
Psychosocial factors	Experience of traumatic events in the past year (e.g., loss of crop, family bereavement), perceived level of occupational, domestic and financial stress, perceived level of mental and physical health
Anthropometry	Height, weight, waist and hip circumference

HRT, hormone replacement therapy

List 6.2: Information collected on dietary intake and tobacco consumption

Type of food	Routinely recorded information	Additional information recorded to reflect local habits
Cooking medium	Oil (recording type), butter, margarine	Ghee (type of fat)
Breads	N/A	Naan, chapatti (recording type of flour used), paratha (butter-coated chapatti)
Fruits and vegetables	Separately for cooked and raw, green leafy vegetables, yellow vegetables, cruciferous vegetables, salad items, fruits	Method of cooking (grilled, curried and fried)
Meats	Chicken, beef, mutton, lamb and fish	Method of cooking (grilled, curried and fried)
Legumes	Nuts and seeds	Pulses (including different daals)
Sweets	Bakery items	Kheer (dairy based dessert), halwa, mithai, jalabe (unrefined sugars), nimko (fried desserts)
Brewed beverages	Tea, coffee, herbal tea	Qahwa (local form of herbal tea)
Other beverages	Milk, carbonated and non-carbonated soft drinks, fruit juices	Alcohol (including local varieties), lassi (yoghurt based drink, recorded separately for sweet and salty)
Tobacco consumption	Usage status, quantity, frequency and lifetime. Information on sources of passive smoking also recorded	For each of cigarettes, beedies (rolled tobacco), huqqa (water pipe), paan (chewing tobacco), naswar (snuff), gutka and supari (preparations of crushed betel nut, tobacco, and sweet or savory flavourings).

References

1. World Bank. Life expectancy at birth in Pakistan. Accessed online on 28-7-2011.

2. Mackay J, Mensah G, World Health Organization. *The Atlas of heart disease and stroke*. 2004.

3. Anand SS, Yusuf S, Vuksan V et al. Differences in risk factors, atherosclerosis, and cardiovascular disease between ethnic groups in Canada: the Study of Health Assessment and Risk in Ethnic groups (SHARE). *Lancet.* 2000;356:279-284.

4. Balarajan R. Ethnic differences in mortality from ischaemic heart disease and cerebrovascular disease in England and Wales. *BMJ.* 1991;302:560-564.

5. Forouhi NG, Sattar N, Tillin T, McKeigue PM, Chaturvedi N. Do known risk factors explain the higher coronary heart disease mortality in South Asian compared with European men? Prospective follow-up of the Southall and Brent studies, UK. *Diabetologia.* 2006;49:2580-2588.

6. Nishtar S. Prevention of coronary heart disease in south Asia. *Lancet.* 2002;360:1015-1018.

7. Ismail J, Jafar TH, Jafary FH, White F, Faruqui AM, Chaturvedi N. Risk factors for non-fatal myocardial infarction in young South Asian adults. *Heart.* 2004;90:259-263.

8. Joshi P, Islam S, Pais P et al. Risk factors for early myocardial infarction in South Asians compared with individuals in other countries. *JAMA.* 2007;297:286-294.

9. Forouhi NG, Sattar N, Tillin T, McKeigue PM, Chaturvedi N. Do known risk factors explain the higher coronary heart disease mortality in South Asian compared with European men? Prospective follow-up of the Southall and Brent studies, UK. *Diabetologia.* 2006;49:2580-2588.

10. Teo KK, Ounpuu S, Hawken S et al. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet.* 2006;368:647-658.

11. Iqbal R, Anand S, Ounpuu S et al. Dietary patterns and the risk of acute myocardial infarction in 52 countries: results of the INTERHEART study. *Circulation.* 2008;118:1929-1937.

12. Rosengren A, Subramanian SV, Islam S et al. Education and risk for acute myocardial infarction in 52 high, middle and low-income countries: INTERHEART case-control study. *Heart.* 2009;95:2014-2022.

13. Saleheen D, Zaidi M, Rasheed A et al. The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia. *Eur J Epidemiol.* 2009;24:329-338.

14. Central Intelligency Agency. The world factbook. 15-4-2012.

15. U.S.A Central Intelligence Agency. The World Factbook: Pakistan. *CIA publications*.

16. World Health Organization. Estimated proportional mortality (%), Pakistan, 2004. 17-8-2011.

Chapter 7: Correlates of tobacco use in a developing country

Summary

Whilst the prevalence of smoking has decreased in Western countries over the past few decades, the production and consumption of tobacco have been rapidly expanding in developing countries. In contrast to industrialized countries where the mass-produced cigarette is by far the most common form of tobacco use, cigarette smoking coexists in low and middle income countries with a range of other practices involving smoking, chewing or dipping tobacco. The prevalence of different types of tobacco use and how they correlate with traditional risk factors of CHD has been little studied in a developing country with a large population such as Pakistan.

The Pakistan Risk of Myocardial Infarction Study (PROMIS) recorded detailed lifestyle information on over 12,000 cases and controls based in urban centres across Pakistan. In controls, prevalence of tobacco use was higher in men, who mostly smoked cigarettes and *bidis*, than women, who generally chewed *paan*. Smoking forms of tobacco was correlated with a "protein and sweets" based diet, as well as with the use of *ghee* rather than oil for cooking, and was most common amongst the Punjabi ethnic group. Chewing tobacco was associated with a lower socio-economic status, the use of oil versus *ghee* as a cooking medium, and was more common amongst the Memon, the Gujrati and the Urdu ethnic groups. Tobacco users of any type were less likely to report a history of diabetes or hypertension. Waist to hip ratio and LDL-C did not differ significantly between users and non users of tobacco products. In a developing country where resources are scarce, these findings may help prioritize groups and design targets for public health prevention of CVD. Adjustment for relevant correlates when investigating the association with MI risk is also important and will be presented in **Chapter 8**.

7.1 Background

Of the total number of smokers worldwide, 84% live currently in developing countries, which amounts to approximately 1.1 billion people [1]. By 2030, more than 80% of tobacco related deaths are predicted to occur in lower and middle income countries [2]. However, evidence on the effect of tobacco use has been mainly derived from studies in developed countries and, as a result; cigarette smoking has been the main studied type of tobacco. In the developing world, and in particular in South Asia, several forms of tobacco use coexist and include various types of tobacco smoking, chewing and dipping [3-5]. In addition to the usual cigarettes, *bidis* (hand-rolled cigarettes where tobacco is wrapped in a dried temburini leaf and tied with a string), *hookah* (water-pipe) and *chilum* (straight conical clay pipe) [6,7] are popular smoking forms of tobacco. In Pakistan, smokeless tobacco is chewed as *paan* (a mixture of areca nut, betel leaf, lime, catechu and tobacco), *supari* (chewed as a digestive and made of processed areca nut, spices and flavourings), *gutka* (a powdered mixture of areca nut, tobacco, catechu, lime, spices and flavourings), or as oral snuff dipping called *naswar* (a moist mixture made from fresh tobacco leaves, calcium oxide and wood ash) [8-11].

Smoking and smokeless forms of tobacco use have been gaining in popularity amongst all sectors of the populations living in developing countries as a result of aggressive marketing techniques by tobacco companies looking for new markets with ineffective tobacco regulations [12], after restrictions on tobacco sale and advertisement have been imposed in the developed world [13]. There is evidence of a rise in both overall and per capita cigarette consumption in the Eastern Mediterranean and South Asian regions over the past decades [12,14]. In Pakistan itself, the production and consumption of cigarettes have doubled since 1970 [15]. Water-pipe, which was in decline worldwide in the 1980's, has been spreading in all parts of the world and mainly in the Eastern Mediterranean region over the past 10 years [16]. The current number of daily users of hookah and other water-pipes has been estimated at 100 million people, and they are especially popular amongst youths and women who perceive it as pleasant smooth smoke with a reduced harm compared to cigarettes. Increasing use of smokeless tobacco has been reported in India and several countries in South East Asia [17]. Smokeless tobacco users in India and Pakistan combined have been estimated to number 100 million individuals, with an over-representation of children, adolescents and women [18,19]. Chewable tobacco is sometimes believed in these countries to have medicinal value for improvement of toothache, headache and even stomach ache, and is often used as toothpaste amongst children[20,21].

As tobacco is a main preventable cause of disability and death, monitoring of its different forms of use in developing countries is necessary. Pakistan is a country with a large population of 165 million individuals where information on tobacco prevalence and correlates is quasi-inexistent [22]. A Pakistani based agency conducted a National Health Study of Pakistani (NHSP) households, but individuals were enrolled in 1990-1994 and there is no more recent information available [23]. The WHO has set up a Global Tobacco Surveillance System to evaluate tobacco control interventions and monitor key articles of the WHO's Framework Convention on Tobacco Control, which is a broad treaty for global tobacco control brought into force in 2005 [24]. With the aim of providing accurate documentation on prevalence of tobacco use, it recently conducted cross-sectional Global Adult Tobacco surveys (GATS) in representative countries. In its first phase it included a South Asian country, Bangladesh [25]. However, Pakistan is more advanced economically than Bangladesh, being classified as a lower middle income country rather than a lower income country [26], and has higher rates of urbanization, all factors which may influence prevalence and type of tobacco use. INTERHEART, a recent worldwide case-control study of MI with a South Asian component had only limited information on Pakistan itself and aggregated data from countries at different stages of economic development and with religious, lifestyle, ethnic and other differences which may reflect on tobacco use (India, Pakistan, Bangladesh, Sri Lanka and Nepal) [27].

In this context the PROMIS controls represent the largest and most recent effort to gather information on a range of lifestyle factors and biochemical factors relevant to CVD in urban Pakistanis, including information from more than 6,000 individuals on their smoking and smokeless habits. Controls were frequency matched to cases by sex and by 5 year age bands, and were recruited randomly and primarily from visitors of patients to the outpatient department. In this Chapter, I aim to better characterize recent trends of tobacco use in urban Pakistan using the control population of PROMIS participants, and investigate correlations of tobacco use with a range of traditional and locally relevant risk factors of CVD.

7.2 *Methods*

7.2.1 Participants

The analysis of prevalence and correlates of tobacco use was restricted to the control set of PROMIS participants. Details of the PROMIS data have been described in **Chapter 6**. In particular, controls were individuals without a self-reported history of CVD drawn from attendants of people visiting out-patients clinics

or non blood related attendants of cardiac patients [28]. Controls were matched to cases by hospital of recruitment and frequency matched by 5 year age bands and by sex. Major forms of tobacco products in Pakistan have been described in **Section 1.2**. Participants were requested to fill in a lifestyle and medical questionnaire with a detailed section on current and past use of the following tobacco products: cigarettes, *bidis*, *hookah*, *chilum*, *paan*, *supari*, *gutka*, *chalia* and *naswar*. No individuals declared smoking chalia. As all users of hookah or chilum also reported smoking manufactured cigarettes or *bidis*, I was not able to study them separately from cigarette smoking. For the assessments of correlates, tobacco use was grouped into "smoking" (cigarettes, *bidis*, *hookah*, or *chilum*), "chewing" (*paan*, *supari*, *chalia* or *gutka*) and "dipping".

Statistical methods

Mean and standard deviations (SD) were reported for approximately normally distributed variables. For skewed variables such as tobacco amounts, median and inter-quartile ranges were reported. To compare tobacco usages across levels of a categorical variable, χ^2 tests of independence were performed and "unadjusted p-values" were reported. To account for the large number of tests performed, p-values<0.01 (rather than p-values<0.05) were considered as significant. For continuous variables, a Student t-test of equalities of the means across tobacco usages was performed.

A multinomial logistic model adjusted for age, sex, ethnicity and centre was used to predict tobacco status categorized into "never, "ex", "current smoker only", "current chewer only", "current *naswar* dipper only", "current smoker and chewer/dipper" (there were no chewer and dippers who were not smokers) depending on lifestyle and medical characteristics. Multinomial logistic regression extends logistic regression by estimating the effect of one or more exposure variables on the probability that the outcome is in a particular category [29]. Let $X_1, X_2, \ldots, X_i$ denote the number that are in each of the *i* categories of tobacco use, in a sample of total size *n*, and the probabilities of belonging to these classes $p_1, p_2, \ldots, p_i$. Let z_j denote baseline characteristics *j*, $1 \le j \le k$. The random variables $X_1, X_2, \ldots, X_i$ have a multivariate distribution given by:

$$P(X_1 = n_1, X_2 = n_2, ..., X_i = n_i) = \frac{n!}{n_1! n_2! ... ! n_i!} p_1^{n_1} p_2^{n_2} ... p_i^{n_4}$$

This likelihood represents an extension of the binomial distribution to the multivariate case and is maximised under the constraints

(i) $\sum_{l=1}^{i} n_l = n$,

(ii) $\sum_{l=1}^{i} p_l = 1$ and

(iii) for $2 \leq j \leq i$, $\ln(\frac{p_l}{p_1}) = \beta_0 + \sum_{j=1}^{k} \beta_j z_j$,

The first category of tobacco use was chosen as the reference group. Models were adjusted for at least age, sex, recruitment centre and ethnicity.

In order to investigate the correlations between tobacco and diet, a principal component analysis (PCA) on the 43 food items recorded in the questionnaire was conducted on the model of the INTERHEART study [30] (see **Appendix 2**). Another PCA was used to define "socio-economic status" and used the following variables: monthly income (<10,000 Pakistani rupees, 10-20,000 Pakistani rupees, >20,000 Pakistani rupees), education (no formal education, 1-10 years, >10 years), asset ownership (mobile phone, television, motorcycle, bicycle, home, computer, car, radio); and occupation (unemployed, retired, housewife, unskilled labour, skilled labour, clerical, business, professional, farmer, self-employed other) when available (see **Appendix 3**). All analyses were run using STATA version 10 (StataCorp, Texas).

7.3 Results

Descriptive characteristics of controls have been detailed in the previous Chapter. There were 6,871 controls included in this analysis. Controls were recruited from three sources: primarily as visitors of patients to the outpatient department of the hospital where cases were recruited (71%), and otherwise as patients of the outpatient department (14%) or as non blood related of MI patients (15%). Nearly half of controls were recruited at the National Institute of CVD based in Karachi (48%), and the rest of controls came mainly from the Punjab Institute of CVD (18%) and the Karachi Institute of Heart Diseases (17%). Within each centre, the proportion of women, mean age and tobacco uses was similar across different sources of controls (**Table 7.1**). There were, however, some ethnic differences between centres, which were sometimes located far apart. Unsurprisingly, the population in the catchment area of the Punjab Institute of CVD was mainly Punjabi (90% of

controls) whilst centres located in Karachi mainly recruited individuals belonging to the Urdu ethno-linguistic group (53% of Deewan Mushtaq Institute of CVD and 77% of individuals from the Karachi Institute of Heart Diseases).

7.3.1 Tobacco use and sex

Current tobacco use was three times more common amongst men than women: 38% versus 12% (**Table 7.2**). In men, the proportion of never users was 54% and of ex-users 9%. In women, the proportion of never users was 86% and of ex-users 2%.

The most popular products amongst men were cigarettes or *bidis* : 26% prevalence amongst male controls; followed by *paan* :8% of controls; and naswar 7% of controls. The least popular products amongst men were *gutka:* 1% of controls; and *supari* : 0.4% of controls (**Figure 7.1**). 4% of men reported the use of several types of tobacco products. Amongst men who chewed, 33% also reported smoking; and amongst men who dipped *naswar*, 16% also reported smoking. Male tobacco users consumed a median of 10 cigarettes or *bidis*, 5 *paan*, 5 *gutka*, 2 *supari*, and 1 dose of *naswar* a day. Men who combined several types of tobacco use in general reported lower amounts of each tobacco use than single users. For example men who smoked and chewed reported an average of 4 *paan* per day and 7 cigarettes or *bidis* per day; versus 5 *paan* per day for chewers only and 10 cigarettes or *bidis* per day for smokers only.

The most popular product amongst women was *paan*, consumed by 8.1% of women. *Naswar* was used by 2% of women, cigarettes or *bidis* by 2%, *gutka* by 0.4% and *supari* by 0.3%. Multi-usage was rare: only 5% of women chewing also reported smoking and/or dipping. Female current tobacco users reported an average of 5 cigarettes or *bidis*, 5 *paan*, 3 *supari*, 2 *gutka* and 1 dose of *naswar* a day.

In a multinomial model adjusted for age, ethnicity and centre of recruitment, compared to never or ex tobacco users, male sex was significantly associated with the probability of smoking (OR for women versus men: 0.05; 95% CI: 0.04-0.08) and with dipping *naswar* (0.19; 0.12-0.29), but not with chewing (1.14; 0.91-1.42) (**Table 7.3**).

7.3.2 Tobacco use and age

Ex-users of tobacco were older on average than never or current tobacco users (**Table 7.2**). The mean age was 56.3 years old (SD: 9.7) in ex-users, 52.3 (9.2) in never users, and 53.1 (9.9) in current users (p-value<0.001). Younger individuals were more likely to smoke than older individuals, but equally likely to use smokeless products: the proportion of smokers only amongst controls aged ≤45 years old was 21%, smokeless users only 11% and smoking and smokeless users 4%. By comparison, amongst individuals aged >55 years old, the proportion of smokers only was 16%, smokeless users only 13% and smoking and smokeless users 2%. Investigating in more detail consumption of smokeless tobacco, prevalence of chewing was similar across ages while prevalence of *naswar* use was greater in older age groups: 5% of individuals ≤45 years old were dipping naswar, compared to 7% of individuals aged >55 years old. After adjustment for conventional risk factors, individuals >55 years old had 52% higher probability of dipping *naswar* than individuals ≤45 years old (p-value: 0.01), 28% lower probability to chew (p-value: 0.01) and 18% lower probability to smoke (p-value: 0.02) (**Table 7.3**).

7.3.3 Tobacco use and ethnicity

Tobacco users varied according to ethnic groups. Tobacco use was most common amongst Pathan and least common amongst Punjabi (**Table 7.2**). Amongst men, around 1 in 3 Punjabi (32%) smoked, compared to 1 in 4 Urdu and Sindh (24% and 25% respectively) and fewer than 1 in 6 Pathan (15%) (**Figure 7.2**) Naswar was most popular amongst Pathan (36% of men) and minority ethnic groups (11% of men), whilst it was relatively uncommon in other major ethnic groups: <5% in Punjabi, Urdu and Sindh men. Chewable tobacco was favoured by Urdu (14% of men current users or *paan*, *supari* or *gutka*) with low prevalence in Punjabi, Pathan and Sindh (3%, 3% and 6% respectively). In minority ethnic groups, 10% reported being current chewers. Prevalence of all forms of tobacco use was much lower in women, independently of ethnic background. Women from the region of Sindh had the highest prevalence of smoking: 8%, whilst smoking was not practiced by women from other groups. Naswar was used by 5% of women from Sindh, 4% of women from minority ethnicities and 2% belonging to the Pathan ethnic group. Regarding chewable tobacco, nearly 1 in 5 Urdu women reported using chewable products (17%), while the proportion was 10% of women from minorities, 6% of Sindhi women, 3% of Pathan women and 2% of Punjabi women. In a multinomial model, after adjustment for age, sex and centre of recruitment, Pathan ethnicity was

positively associated with *naswar* use, and Urdu ethnicity with the use of chewable forms of tobacco. Both Pathan and Urdu ethnicities had a reduced probability of smoking, while Punjabi were the most likely to smoke (**Table 7.3**).

There was no significant correlation between centre of recruitment and type of tobacco use, apart from the Multan Institute of Heart Diseases (MIC) where the probability of chewing was lower compared to other centres, even after adjusting for sex, age and ethnic group (**Table 7.3**). The MIC is the only recruitment centre located in the geographical centre of the country which recruited individuals with an ethnic composition different from that of the other centres (45% of controls and 49% of cases were classified as "Others" ethnic groups, probably Multani ethnicity).

7.3.4 Tobacco use and conventional risk factors of MI

There was a higher proportion of never or past tobacco users amongst individuals diagnosed with diabetes or hypertension than in the general population of the PROMIS controls: 73% of diabetics were non tobacco users versus 67% of non diabetics, and the contrast was 74% versus 66% of controls for hypertension (**Table 7.4**). After adjustment for age, ethnicity, sex and centre of recruitment, individuals with self reported diabetes or hypertension remained less likely to declare themselves as smokers (P-value: 0.002 and <0.0001 respectively), but were equally likely to chew tobacco or to dip *naswar* (p-values>0.01) than to be non users of tobacco (**Table 7.5**). Compared to never users of tobacco, individuals with a history of diabetes were marginally more likely to have stopped smoking than non-diabetics (adjusted multinomial OR 1.28; 95%CI: 1.00-1.63; p-value: 0.048); and individuals with hypertension to have stopped chewing than non-hypertensives (adjusted multinomial OR 1.48; 95%CI: 1.02-2.15; p-value 0.034). Tobacco use was not significantly influenced by family history of MI. Sensitivity analyses further adjusting for socio-economic status and diet did not modify these results. Mean levels of lipids did not significantly differ across different types of tobacco use and across status (current, ex or never user) (**Figure 7.3**). There was also no evidence of a relationship between waist to hip ratio and the use of tobacco (**Figure 7.4**).

7.3.5 Tobacco use and socio-economic status

Low socio-economic status was associated with higher tobacco consumption. Amongst individuals earning ≥20,000 Pakistani rupees a month (~230 US dollars), the proportion of current tobacco users was 25%, while it was 36% amongst

individuals earning <10,000 Pakistani rupees a month (~115 US dollars) (**Table 7.6**). Similar differences were observed between individuals reporting >10 years of formal education, compared to no formal education (25% versus 36% of current users). The proportion of current users of tobacco in the top third of the socio-economic gradient was 25% versus 38% in the lowest third. In particular naswar dippers and chewers were over-represented in lower socio-economic groups: 70% of naswar and 47% of chewers were in the bottom third of socio-economic status. Onaverage, smokers in the lowest third of socio-economic gradient reported smoking 10 cigarettes or bidis per day (IQR: 5-20), compared to 12 (5-20) in the top third of socio-economic status (p-value: 0.034).

After adjustment for age, sex, recruitment centre and ethnicity, earning ≥20,000 Pakistani rupees a month was associated with a 37% lower likelihood to smoke, 55% lower likelihood to chew and 70% lower probability to dip *naswar* compared to individuals earning <10,000 Pakistani rupees a month (**Table 7.7**). There was a 53% reduction in the probability of smoking, 61% reduction in the probability of chewing and 90% reduction in the probability of dipping naswar for individuals reporting >10 years of formal education compared to individuals with no formal education. For socio-economic status a reduction of 42%, a 63% and an 85% in the likelihood of smoking, chewing and dipping was observed when comparing top versus bottom third of the socio-economic gradient.

7.3.6 Tobacco use and diet

Individuals who cooked with *ghee* only were more likely to report current tobacco use than individuals who cooked with oil (p-value<0.001) (**Table 7.8**). Amongst never or ex-users of tobacco the prevalence of *ghee* users, exclusively or in combination with oil, was 49%, while it was 42% amongst chewers, 53% amongst current smokers (or multi-users) and 62% amongst *naswar* dippers. In a multinomial adjusted model, the relationship between tobacco and *ghee* use remained significant: *ghee* only users had twice the probability to be also naswar dippers and 57% higher probability to smoke than oil only users (p-values<0.001) (**Table 7.9**). The association with tobacco chewing was non significant.

Individuals in the top fifth of the distribution of the “vegetables and carbohydrate” based diet were more likely to use tobacco than individuals consuming “vegetables and carbohydrates” less frequently. The proportion of current tobacco users was lowest in the middle categories of the “protein and sweets” based diet, whilst extremes reported a higher prevalence of smokers. Amongst individuals scoring in

the top fifth for the “high meat, fish and sweet diet” dietary pattern, the proportion of smokers was 26%, chewers 7% and *naswar* dippers 3%; compared to 19% smokers, 4% chewers and 4% naswar dippers in the middle quintile. However, in a multinomial model adjusted for demographic variables, there was no significant evidence of a correlation between tobacco use and dietary patterns (**Table 7.9**).

7.4 Discussion

7.4.1 Tobacco use and gender

The high prevalence of tobacco use amongst PROMIS participants is comparable to other published findings. Smoking is most prevalent in men from developing countries and South Asia is the region of the world with the highest prevalence of smokeless tobacco users [1]. Tobacco in its smoking, chewing or snuff forms was used by nearly 2 in every 5 men and more than 1 in every 10 women of PROMIS control participants. The proportion of current smokers was 26% men and 2% women. The use of smokeless types of tobacco attained 9% of chewers and 7% of *naswar* dippers in men; versus 9% and 2% in women. Rates of tobacco cessation were low: only 9% of men and 2% of women declared being ex-tobacco users. The patterns of tobacco use differed between sexes in PROMIS. The majority of male tobacco users smoked while women were more likely to chew *paan*. Combining smoking, chewing and dipping, tobacco use was more common amongst men than women. After adjustment for other demographic characteristics, smoking and dipping *naswar* were predominant amongst men, while chewing was not associated with a specific gender.

These findings are in broad agreement with previous studies in South Asian populations. The National Health Study of Pakistani (NHSP) conducted a decade ago in Pakistan reported that 34% of men and 12.5% of women in Pakistan use tobacco regularly [31]. In the GATS survey of Bangladeshis established by WHO, higher rates of both smoking and smokeless tobacco use were observed compared to PROMIS. Bangladeshi men aged ≥15 years old reported a 47% current smoking rate, including 25% daily manufactured cigarette smoking, and a 26% current smokeless tobacco rate [33]. In women, rates were 1.5% current smoking, including 0.2% daily cigarette smoking and 28% current smokeless use. Regarding cessation rates, rates in PROMIS were lower than in INTERHEART [34] (proportions of male ex-smokers were 13% and 22% among young and old men respectively) and than in

GATS (17% in men and 41% in women). As INTERHEART and GATS were partially or wholly set in South Asian countries other than Pakistan, these differences may reflect differences in countries' prevalence and reinforce the need for country specific statistics in order to effectively monitor tobacco control in South Asia. Lower rates of tobacco cessation in PROMIS compared to the rest of South Asia also raise the alarm of a lack of effectiveness of public health campaigns intended to discourage tobacco use in Pakistan relative to the rest of South Asia [35,36].

7.4.2 Tobacco use and age and socio-economic characteristics

Smoking was more likely amongst younger age groups, a pattern also observed by INTERHEART [37]. This raises concerns that younger age groups in Pakistan and more generally in developing countries see smoking as part of the adoption of a modern and westernized lifestyle which accompanies poor diet, low levels of physical activity and is currently held responsible for an epidemic of obesity [38]. Punjabi and Sindhi favoured smoking, Pathan reported mainly using *naswar*, and Urdu both smoked and chewed tobacco. In the NHSP, the highest prevalence of smoking was also seen amongst urban Punjabi (30% amongst men) and Sindhi (24% amongst men) compared to other ethnic groups and rural populations [39]. Tobacco use was more common in my study amongst individuals in lower socio-economic groups characterized by short education, low income, poor asset ownership, over-representation in jobs such as unskilled labour, and had a relatively large proportion of unemployed individuals. In the NHSP study, illiteracy and low levels of education were also significantly associated with tobacco use, but there was no significant association with monthly income [40].

7.4.3 Tobacco use and conventional risk factors of MI

Regarding medical and biochemical risk factors for CVD, hypertension and diabetes were highly prevalent amongst controls. Diabetes and hypertension are spreading with urbanization in Pakistan. In the NHSP, 21% of urban compared to 16% of rural participants reported being hypertensive [41]. In the Pakistan National Diabetes Study, the prevalence of diabetes was 10.8% in rural areas compared to 11.9% in urban areas [42]. In comparison, in PROMIS, 28% of controls had self reported hypertension and 14% had diabetes, among which a quarter used tobacco products. There was a slightly higher proportion of never or past tobacco users amongst diabetics and hypertensive than in the general population of the PROMIS controls, maybe reflecting a better awareness of the dangers of tobacco use in this group of people.

There was no significant difference in LDL-C and total cholesterol levels in current versus ex or never tobacco users. This is in contrast to results from a large scale meta-analysis of Western prospective cohort studies presented in **Chapter 3** which found a modest association with lipid levels. Similarly, no correlation was observed between waist to hip ratio and tobacco use in PROMIS, which could be due to measurement error in either waist to hip ratio or the recording of tobacco use, confounded by factors not adjusted for in the linear regression of waist to hip ratio on tobacco use such as physical activity, or a real lack of association between anthropometric measurements and tobacco use in South Asian populations. These findings would need to be replicated using a prospective design in order to confirm the direction of causality.

7.4.4 Tobacco use and diet

Tobacco users were more likely to use *ghee*. Traditional South Asian diet is known to be rich in fried food and, rather than being boiled, vegetables and meat are usually fried in either *ghee*, which is clarified butter and is a saturated fat containing cholesterol oxides, or vegetable oil [43]. *Ghee* use has been associated with higher adipose tissue levels of trans-fatty acids [44] than oil for cooking. The lifestyle in Pakistan has been rapidly changing over the past 30 years, as urbanization has been accompanied by higher consumption of meat, sweets and fat-rich food and a preference for cigarettes over chewable forms of tobacco [45,46] However, the association of tobacco use with the two main dietary patterns, “high vegetables and carbohydrate diet” and “high meat, fish and sweet diet”, was not significant after adjustment for demographic factors and *ghee* use.

7.4.5 Strengths and limitations

My study contains several strengths and limitations. Recruitment for the PROMIS was performed in 6 centres located in 5 urban centres across Pakistan, enabling investigation of tobacco use in all major ethnic groups of Pakistan. The PROMIS represents the largest dataset in Pakistan with detailed information on several smoking and smokeless forms of tobacco use, as well as other lifestyle and biochemical factors for CVD. In comparison, INTERHEART included ~2,000 controls from South Asian countries, and less than 1,000 specifically from Pakistan [47]. The NHSP included 9442 individuals ≥15 years old [48]; however its recruitment started more than a decade ago [49].

The PROMIS controls were not recruited to form a representative sample of the population of Pakistan, but rather to be age and sex frequency matched to cases without a self-reported history of CVD, and mainly identified as visitors of patients. Nevertheless, prevalence of tobacco use was in broad agreement with previous estimates [50]. To remedy confounding by the demographic structure of my data, such as an over-representation of men and Urdu in PROMIS compared to the overall population of Pakistan, multinomial logistic regression were fitted adjusting at least for age, sex, ethnicity and recruitment centre. Socio-economic status was assessed, taking into account several social and economic dimensions including monthly income, education, employment status, asset ownership of nine types of common household possessions, and type of job categorized into six categories. Assessment of diet included a question on the use of ghee, as well as a 43 food item questionnaire on meals and food locally relevant to Pakistanis. Information was sparse on amount amongst chewers and naswar dippers and was not collected separately for cigarettes and for *bidis*. Data on specific types of tobacco such as *huqqa*, *chilum*, *paan*, *supari* and *gutka* were sparse and these types could not be investigated on their own but had to be grouped into broader categories such as "smoking tobacco" and "chewing tobacco". Epidemiological studies with designs favouring enrolment of users of smokeless tobacco, for example set up in an Urdu population, would be of interest in this respect, allowing the analysis of correlates of specific subtypes of smokeless tobacco.

7.5 Conclusion

Tobacco use was widespread in the control set of PROMIS participants, with high prevalence in all ethnic groups. Younger age groups were more likely to smoke cigarettes or chew tobacco, whilst older age groups had a higher probability of dipping *naswar*. Men had increased likelihood of smoking cigarettes or *bidis*, often in combination with chewing or dipping tobacco; and women had increased likelihood of chewing *paan*. Stopping smoking was rare, even amongst populations at high risk of MI such as diabetics and hypertensive individuals. Current tobacco use was more common amongst lower socio-economic groups and was correlated with a poor diet characterized by the use of *ghee* rather than vegetable oil for cooking. These findings emphasize the need for accurate monitoring of tobacco trends in Pakistan, in order to better implement tobacco control.

Figure 7.1: Current tobacco use amongst 5,359 men and 1512 women controls

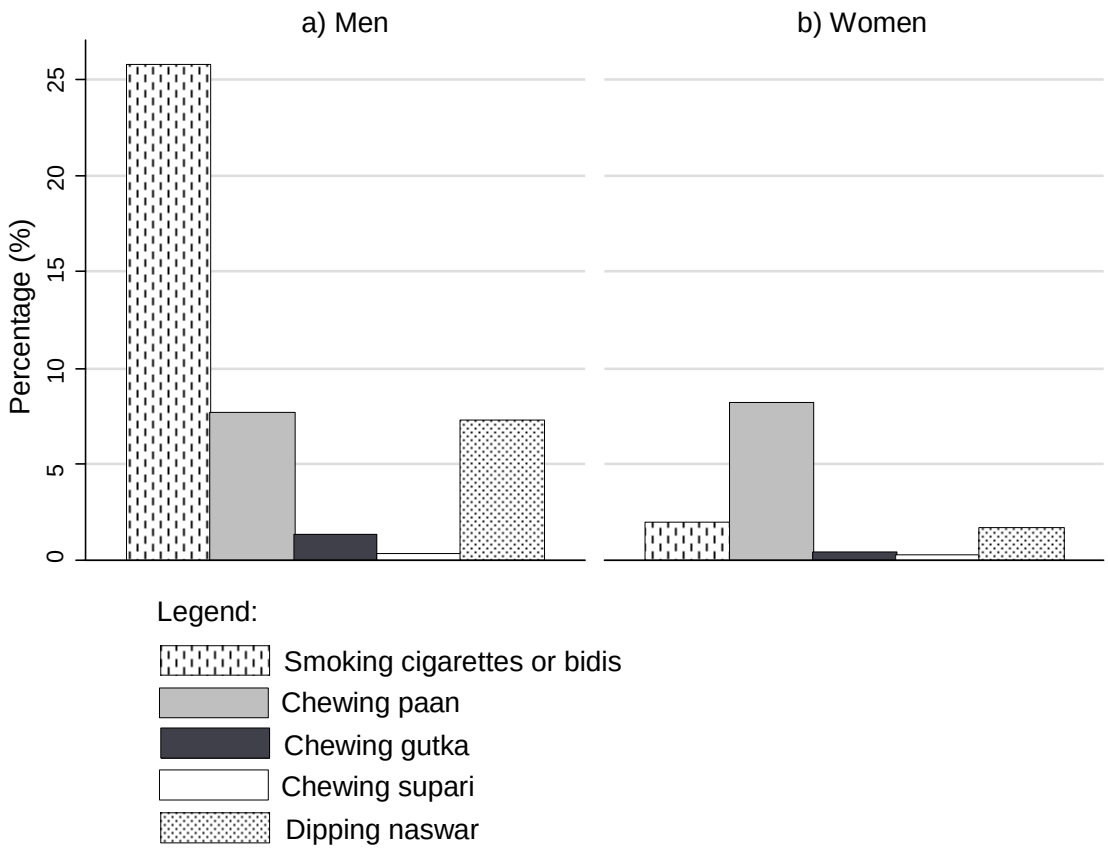


Figure 7.2: Current tobacco use amongst controls by ethnic group amongst 5,359 men and 1512 women controls

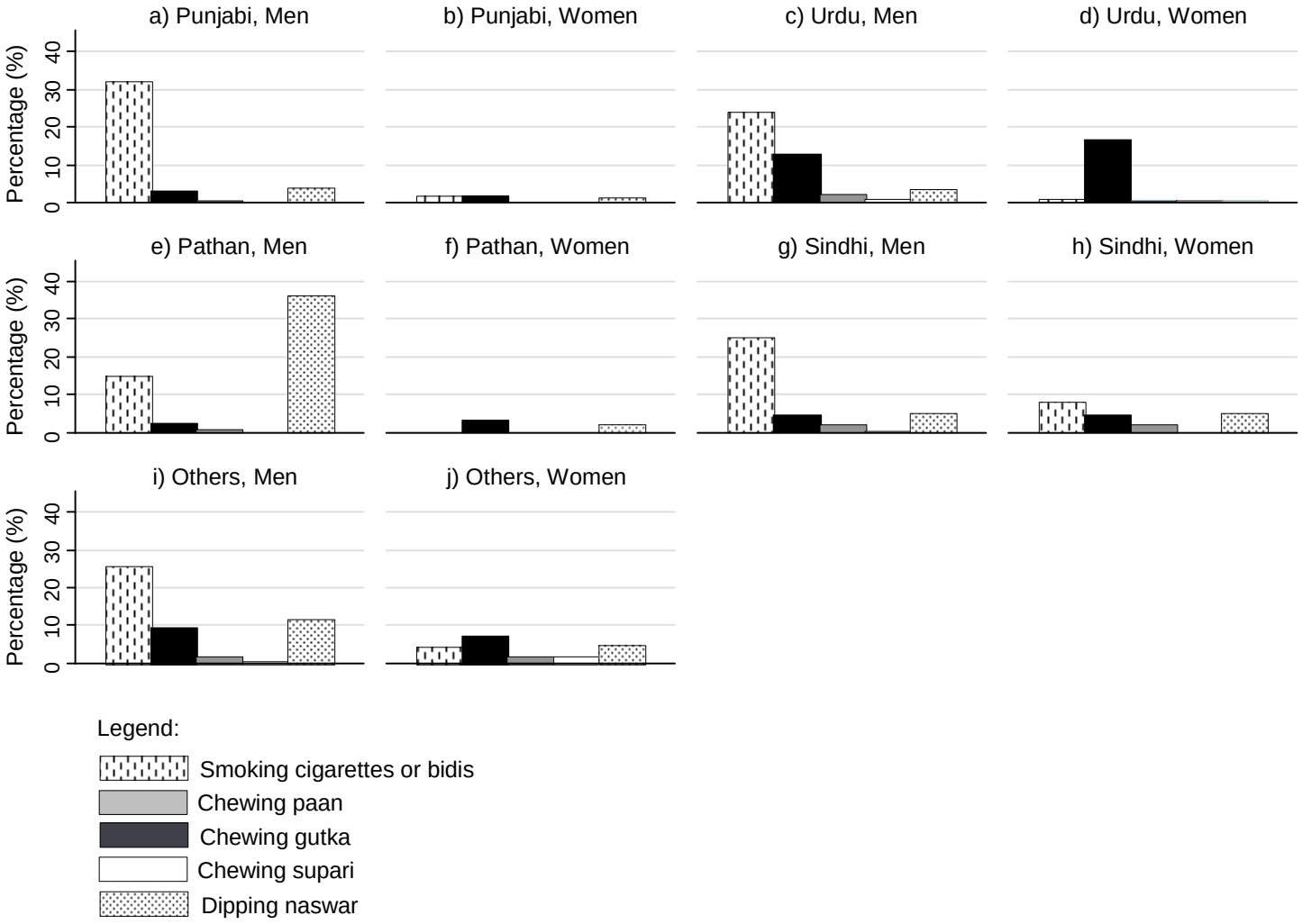
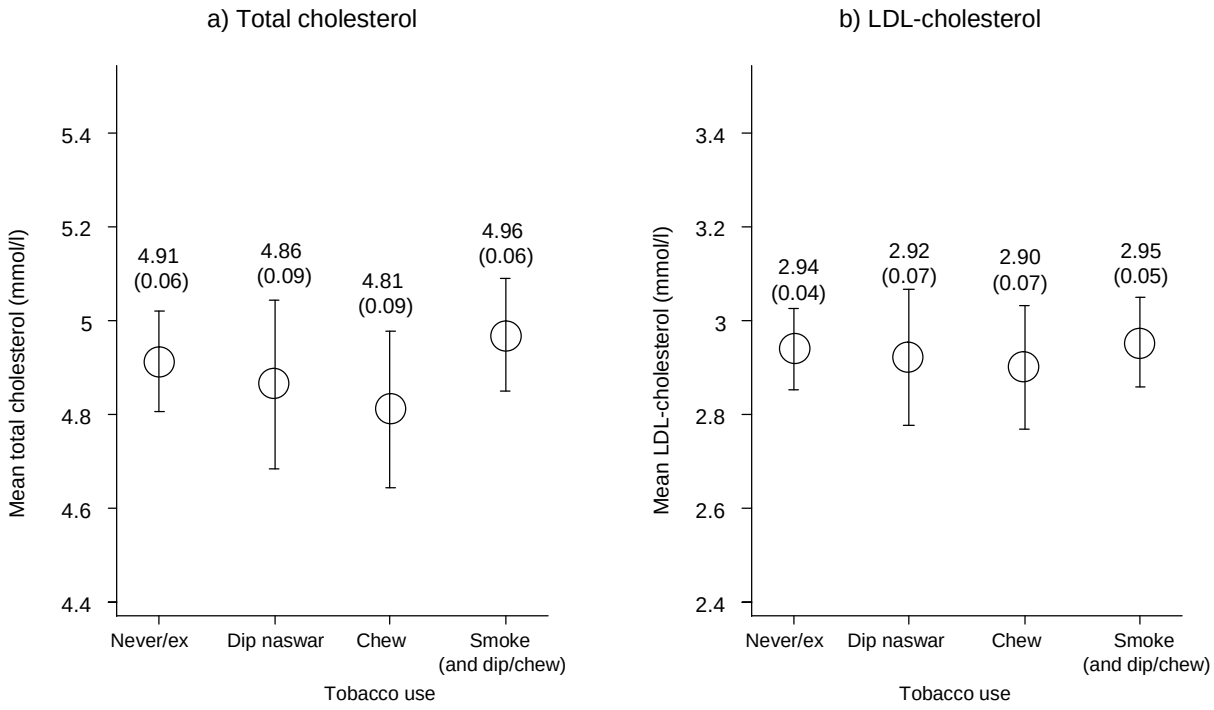
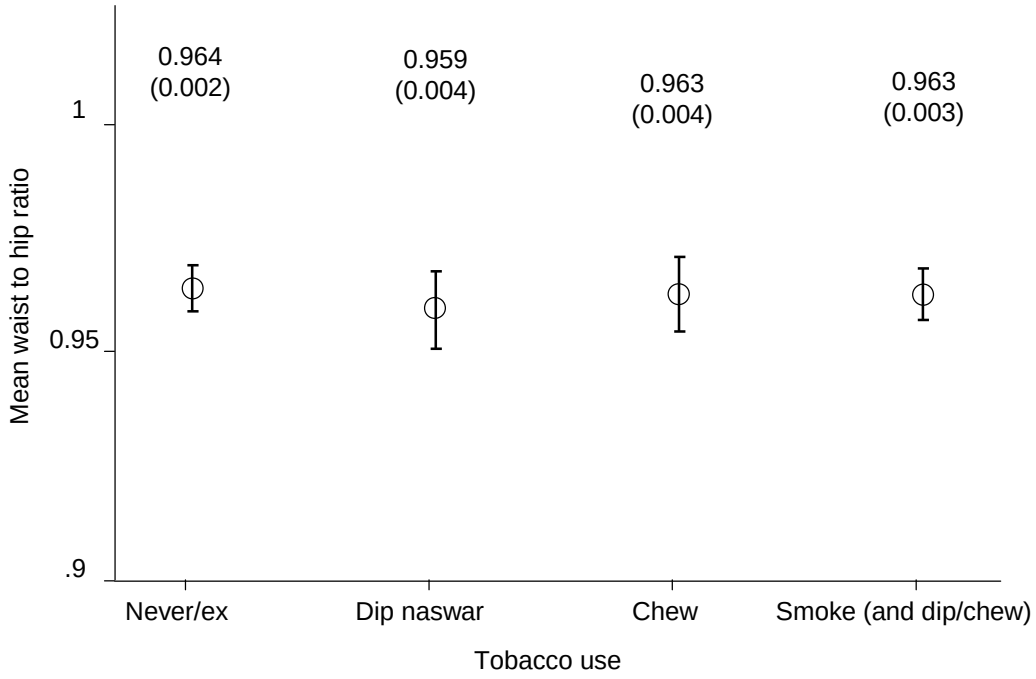


Figure 7.3: Adjusted lipid levels and 95%confidence intervals by tobacco use after multiple adjustment



Means and 95% CI were computed by fitting a linear regression of lipid levels over categories of tobacco use, adjusting for age, sex, centre of recruitment and ethnicity. Adjusted means were obtained as predicted coefficients for age 50 years old, Male, Urdu and averaging the effects of centre. Adjusted means (standard deviations) are given for each category.

Figure 7.4: Levels of waist to hip ratio by tobacco use after multiple adjustment



Means and 95% CI were computed by fitting a linear regression of lipid levels over categories of tobacco use, adjusting for age, sex, centre of recruitment and ethnicity. Adjusted means were obtained as predicted coefficients for age 50 years old, Male sex, Urdu and averaging the effects of centre. Adjusted means (standard deviations) are given for each category

Table 7.1: Prevalence of demographic factors by source of controls and by recruitment centre

Variable	DMIC centre (n=161)	KIHD centre (n=1,165)		MIC centre (n=415)	NICVD centre (n=3,326)			PIC centre (n=1,2363)		RCH centre (n=542)	
		Visitors of MI patients (n=137)	Visitors of OPD patients (n=1027)		Visitors of MI patients (n=158)	OPD patients (n=869)	Visitors of OPD patients (n=1826)	Visitors of MI patients (n=307)	Visitors of OPD patients (n=956)	Visitors of MI patients (n=352)	Visitors of OPD patients (n=188)
Age (years)	50.3 (9.2)	54.2 (7.5)	55.2 (9.1)		50.7 (7.9)	58.1 (8.3)	52.7 (9.9)	48.3 (9.6)	50.4 (8.9)	52.5 (3.5)	51.6 (98.3)
Female sex	18%	14%	19%	15%	25%	29%	18%	25%	30%	21%	27%
Major sub ethnicities											
Punjabi	2%	8%	9%	36%	13%	17%	20%	88%	91%	5%	10%
Urdu	53%	80%	77%	18%	41%	42%	34%	7%	4%	53%	46%
Pathan	1%	3%	5%	1%	11%	14%	13%	2%	1%	1%	2%
Sindhi	41%	3%	3%	0%	20%	18%	17%	0%	1%	39%	37%
Others	2%	6%	6%	45%	14%	8%	16%	4%	3%	2%	6%
Tobacco use											
Never user	53%	49%	52%	67%	61%	63%	54%	65%	71%	75%	76%
Ex-user	11%	18%	13%	7%	6%	5%	8%	5%	6%	1%	1%
Dipping only	1%	1%	4%	1%	11%	9%	8%	3%	1%	1%	2%
Chewing only	8%	12%	12%	0%	5%	5%	8%	1%	1%	3%	4%
Smoking only	21%	16%	16%	21%	14%	16%	18%	25%	20%	18%	17%
Smoking & dipping/chewing	6%	4%	3%	2%	4%	1%	4%	2%	1%	1%	2%

N: Number of individuals; SD: standard deviation; DMIC: Deewan Mushtaq Institute of Cardiology; FIC: Faisalabad Institute of Cardiology; KIHD: Karachi Institute of Heart Diseases; MIC: Multan Institute of Heart Diseases; NICVD: National Institute of Cardiovascular Diseases; PIC: Punjab Institute of Cardiovascular Disease; RCH: Red Crescent Hospital. OPD: Outpatient Department. Information on source of controls was missing from 473 controls from NICVD and 1 control from KIHD. 1 individual from MI and 2 individuals from RCH who were OPD patients were excluded from this Table. Columns are column percentages

Table 7.2: Description of tobacco use by demographic characteristics

	Tobacco use					Median number cigarettes / bidis per day
	Never/Ex-user (n=4675)	Dip naswar (n=348)	Chew (n=450)	Smoke (and dip/chew) (n=1416)	P-value	
Sex					*<0.0001*	
Male	62%	6%	6%	26%		10
Female	88%	2%	9%	2%		5
Age, mean (SD)	53.0 (9.4)	55.5 (9.9)	53.1 (9.9)	52.2 (9.8)	*<0.0001*	
≤45 years old	66%	4%	6%	24%	<0.0001	10
46-55 years old	69%	5%	6%	20%		12
>55 years old	68%	7%	7%	19%		10
Major ethnic groups					*<0.0001*	
Punjabi	72%	3%	1%	24%		12
Urdu	67%	2%	12%	19%		10
Pathan	60%	26%	2%	12%		6
Sindhi	70%	4%	4%	22%		10
Other	63%	8%	7%	22%		10
Centre of recruitment					*<0.0001*	
DMIC	64%	1%	8%	27%		10
FIC	79%	0%	0%	21%		5
KIHD	65%	4%	12%	19%		10
MIC	74%	1%	0%	24%		10
NICVD	64%	8%	8%	20%		10
PIC	75%	2%	1%	22%		12
RCH	77%	1%	3%	19%		10

N: Number of individuals; SD: standard deviation; DMIC: Deewan Mushtaq Institute of Cardiology; FIC: Faisalabad Institute of Cardiology; KIHD: Karachi Institute of Heart Diseases; MIC: Multan Institute of Heart Diseases; NICVD: National Institute of Cardiovascular Diseases; PIC: Punjab Institute of Cardiovascular Disease; RCH: Red Crescent Hospital. Percentages correspond to row percentages. P-value from a χ^2 test of independence between row and column variables. The median number of cigarettes or bidis per day is reported amongst current controls who smoke.

Table 7.3: Demographic determinants of tobacco use after multiple adjustment

	Dip naswar		Chew		Smoke (and dip/chew)	
	Basic adjustment		Basic adjustment		Basic adjustment	
	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value
Sex						
Male	1	-	1	-	1	-
Female	0.19 (0.12;0.29)	<00001	1.14 (0.91;1.42)	0.3	0.05 (0.04;0.08)	<0.001
Age, mean (SD)	1.02 (1.00;1.03)	0.01	0.99 (0.98;1)	0.01	0.99 (0.99;1)	0.09
≤45 years old	1	-	1	-	1	-
46-55 years old	1.28 (0.92; 1.78)	0.14	0.75 (0.58; 0.97)	0.03	0.88 (0.75; 1.03)	0.11
>55 years old	1.52 (1.10; 2.10)	0.01	0.72 (0.55; 0.93)	0.01	0.82 (0.70; 0.97)	0.02
Major ethnic groups						
Punjabi	1	-	1	-	1	-
Urdu	0.55 (0.36;0.85)	0.007	4.22 (2.76;6.46)	<0.001	0.69 (0.57;0.83)	<0.001
Pathan	6.27 (4.23;9.3)	<0.001	0.67 (0.31;1.46)	0.3	0.49 (0.35;0.67)	<0.001
Sindhi	0.99 (0.60;1.62)	1	1.57 (0.92;2.69)	0.1	0.73 (0.57;0.93)	0.01
Other	2.45 (1.61;3.71)	<0.001	3.42 (2.11;5.54)	<0.001	0.84 (0.67;1.06)	0.1
Centre of recruitment						
DMIC	1	-	1	-	1	-
KIHD	2.66 (0.62;11.34)	0.2	1.3 (0.7;2.4)	0.4	0.71 (0.47;1.07)	0.1
MIC	0.48 (0.09;2.49)	0.4	0.06 (0.01;0.27)	<0.001	0.59 (0.38;0.94)	0.03
NICVD	3.28 (0.79;13.61)	0.1	1.17 (0.64;2.14)	0.6	0.78 (0.53;1.14)	0.2
PIC	1.00 (0.22;4.52)	1	0.2 (0.08;0.5)	0.001	0.62 (0.4;0.95)	0.03
RCH	0.94 (0.2;4.54)	0.9	0.31 (0.14;0.67)	0.003	0.62 (0.41;0.96)	0.03

OR: Odds ratio; 95% CI: 95% confidence interval. Odds ratios were computed using a multinomial logistic regression with never or ex tobacco users chosen as the reference group. Models were adjusted for age as a continuous variable (except when looking at the effect of categories of age, when age as a continuous covariate was omitted from the model), sex, major ethnic groups and centre of recruitment. Coefficients for Faisalabad Institute of Cardiology were not estimable because of low numbers.

Table 7.4: Conventional CVD risk factors and tobacco use

	Tobacco use				
	Never/Ex-user (n=4675)	Dip naswar (n=348)	Chew (n=450)	Smoke (and dip/chew) (n=1416)	P-value
Self reported diabetes					*0.0004*
No	67%	5%	7%	21%	
Yes	73%	4%	6%	16%	
Self reported hypertension					*<0.0001*
No	66%	5%	6%	23%	
Yes	74%	4%	7%	15%	
Family history of CAD					*0.009*
No	68%	5%	6%	21%	
Yes	69%	3%	7%	20%	
Biochemical information					
Total cholesterol (mmol/l), mean (SD))	4.66 (1.35)	4.37 (1.26)	4.37 (1.41)	4.69 (1.42)	
LDL-C (mmol/l), mean (SD)	2.88 (1.05)	2.77 (1.08)	2.78 (1.08)	2.87 (1.11)	
Waist to hip ratio, mean (SD)	0.943 (.0671)	0.937 (.0616)	0.936 (.0665)	0.946 (.0631)	

N: Number of individuals; SD: standard deviation; cig: cigarettes. Percentages correspond to row percentages. P-value from a χ^2 test of independence between row and column variables. The median number of cigarettes or bidis per day is reported amongst current controls who smoke.

Table 7.5: Conventional CVD risk factors and tobacco use after multiple adjustment

	Dip naswar only (n=348)		Chew only (n=450)		Smoke (and dip/chew) (n=1416)	
	Basic adjustment OR (95% CI)	P-value	Basic adjustment OR (95% CI)	P-value	Basic adjustment OR (95% CI)	P-value
Diabetes	0.75 (0.53;1.07)	0.115	0.83 (0.62;1.11)	0.203	0.74 (0.61;0.89)	0.002
Hypertension	0.75 (0.57;0.99)	0.043	0.9 (0.72;1.12)	0.351	0.69 (0.59;0.8)	<0.0001
Family history of MI	0.94 (0.63;1.41)	0.773	1.23 (0.93;1.64)	0.147	0.92 (0.77;1.11)	0.387

Multinomial logistic regression adjusted for age, sex, major ethnic groups and centre of recruitment. Never/ex tobacco users were chosen as the reference group.

Table 7.6: Association of socio-economic status with tobacco use

	Tobacco use					
	Never/Ex-user (n=4675)	Dip naswar only (n=348)	Chew only (n=450)	Smoke (and dip/chew) (n=1416)	P-value	Median number of cig./bidis per day
Income					*<0.0001*	
Low (<10,000 Pakistani rupees/month)	64%	7%	7%	21%		10
Middle (10-20,000 Pakistani rupees/month)	68%	5%	7%	20%		10
High (≥20,000 Pakistani rupees/month)	75%	2%	3%	20%		12
Education					*<0.0001*	
No formal education	64%	9%	8%	19%		10
1-10 years	65%	5%	7%	23%		10
>10 years	75%	1%	4%	20%		10
Socio-economic gradient					*<0.0001*	
Lower third	62%	9%	7%	21%		10
Middle third	67%	5%	8%	20%		10
Top third	75%	1%	3%	21%		12

N: Number of individuals; SD: standard deviation; Percentages correspond to row percentages; cigs: cigarettes. P-value from a χ^2 test of independence between row and column variables. The median number of cigarettes or bidis per day is reported amongst current controls who smoke.

Table 7.7: Socio-economic status and tobacco use after multiple adjustment

	Dip naswar only (n=348)		Chew only (n=450)		Smoke (and dip/chew) (n=1416)	
	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value
Income						
Low (<10,000 Pakistani rupees/month)	1	-	1	-	1	-
Middle (10-20,000 Pakistani rupees/month)	0.79 (0.61; 1.04)	0.091	0.91 (0.72; 1.16)	0.474	0.88 (0.76; 1.02)	0.101
High (≥20,000 Pakistani rupees/month)	0.32 (0.21; 0.48)	0	0.45 (0.33; 0.63)	0	0.73 (0.62; 086)	0
Education						
No formal education	1	-	1	-	1	-
1-10 years	0.53 (0.40; 0.69)	0	0.93 (0.73; 1.19)	0.556	0.75 (0.64; 0.88	0.001
>10 years	0.10 (0.06; 0.17)	0	0.39 (0.29; 0.52)	0	0.47 (0.40; 0.56)	0
Socio-economic gradient						
Lower third	1	-	1	-	1	-
Middle third	0.5 (0.38;0.66)	0	0.93 (0.73;1.18)	0.56	0.78 (0.67; 0.91)	0.003
Top third	0.16 (0.1;0.24)	0	0.37 (0.27; 0.51)	0	0.58 (0.50; 0.69)	0

Multinomial models were adjusted for age, sex, ethnicity and recruitment centre.

Table 7.8: Diet and tobacco use

	Tobacco use					Median number of cig./bidis per day
	Never/Ex-user (n=4675)	Dip naswar only (n=348)	Chew only (n=450)	Smoke (and dip/chew) (n=1416)	P-value	
Cooking fat					*<0.0001*	
Oil	70%	4%	7%	20%		10
Oil & ghee	67%	5%	7%	21%		10
Ghee only	65%	7%	3%	25%		10
Dietary pattern 1 "High vegetables and carbohydrates"					*0.0009*	
Quintile 1	70%	4%	4%	22%		10
Quintile 2	69%	4%	6%	21%		12
Quintile 3	70%	4%	6%	20%		10
Quintile 4	71%	5%	6%	18%		10
Quintile 5	65%	6%	7%	23%		10
Dietary pattern 2 "High mean, fish and sweets"					*<0.0001*	
Quintile 1	69%	6%	6%	19%		10
Quintile 2	69%	5%	6%	20%		10
Quintile 3	72%	4%	5%	19%		10
Quintile 4	69%	4%	7%	20%		10
Quintile 5	64%	3%	7%	26%		10

N: Number of individuals; SD: standard deviation; Percentages correspond to row percentages; cigs: cigarettes. P-value from a χ^2 test of independence between row and column variables. The median number of cigarettes or bidis per day is reported amongst current controls who smoke.

Table 7.9: Diet and tobacco use after multiple adjustment

	Dip naswar only (n=348)		Chew only (n=450)		Smoke (and dip/chew) (n=1416)	
	Basic adjustment OR (95% CI)	P-value	Basic adjustment OR (95% CI)	P-value	Basic adjustment OR (95% CI)	P-value
Cooking fat						
Oil	1	-	1	-	1	-
Oil & ghee	1.48 (1.06;2.06)	0.021	1.36 (1.05;1.76)	0.02	1.09 (0.93;1.29)	0.286
Ghee only	*2.11 (1.52;2.93)*	0	0.88 (0.59;1.32)	0.542	1.57 (1.31;1.89)	0
Dietary pattern 1 "High vegetables and carbohydrates"						
Quintile 1	1	-	1	-	1	-
Quintile 2	*1.14 (0.72;1.82)*	0.576	1.62 (1.08;2.45)	0.021	1.05 (0.85;1.32)	0.638
Quintile 3	*1.26 (0.79;2.03)*	0.336	1.72 (1.13;2.62)	0.011	1.08 (0.86;1.36)	0.519
Quintile 4	*1.37 (0.9;2.08)*	0.143	1.77 (1.19;2.62)	0.005	0.95 (0.77;1.19)	0.68
Quintile 5	*1.64 (1.09;2.46)*	0.017	1.59 (1.07;2.34)	0.021	1.21 (0.98;1.5)	0.073
Dietary pattern 2 "High mean, fish and sweets"						
Quintile 1	1	-	1	-	1	-
Quintile 2	0.94 (0.65;1.36)	0.749	0.95 (0.67;1.34)	0.767	1.16 (0.94;1.43)	0.159
Quintile 3	0.79 (0.53;1.18)	0.246	0.72 (0.49;1.06)	0.095	0.98 (0.79;1.22)	0.849
Quintile 4	0.79 (0.52;1.21)	0.281	1.06 (0.74;1.53)	0.734	1.02 (0.81;1.28)	0.868
Quintile 5	0.54 (0.34;0.88)	0.012	1.05 (0.73;1.53)	0.78	1.41 (1.13;1.76)	0.003

n: Number of individuals. Multinomial models adjusted for age, sex, ethnicity, recruitment centre, type of ghee used for cooking and dietary patterns.

References

1. World Health Organization. *Tobacco: deadly in any form or disguise*. 2006.

2. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med.* 2006;3:e442.

3. Changrani J, Gany FM, Cruz G, Kerr R, Katz R. Paan and Gutka Use in the United States: A Pilot Study in Bangladeshi and Indian-Gujarati Immigrants in New York City. *J Immigr Refug Stud.* 2006;4:99-110.

4. Dobe M, Sinha DN, Rahman K. Smokeless tobacco use and its implications in WHO South East Asia Region. *Indian J Public Health.* 2006;50:70-75.

5. Imam SZ, Nawaz H, Sepah YJ, Pabaney AH, Ilyas M, Ghaffar S. Use of smokeless tobacco among groups of Pakistani medical students - a cross sectional study. *BMC Public Health.* 2007;7:231.:231.

6. Nisar N, Qadri MH, Fatima K, Perveen S. A community based study about knowledge and practices regarding tobacco consumption and passive smoking in Gadap Town, Karachi. *J Pak Med Assoc.* 2007;57:186-188.

7. Anjum Q, Ahmed F, Ashfaq T. Knowledge, attitude and perception of water pipe smoking (Shisha) among adolescents aged 14-19 years. *J Pak Med Assoc.* 2008;58:312-317.

8. Ali NS, Khuwaja AK, Ali T, Hameed R. Smokeless tobacco use among adult patients who visited family practice clinics in Karachi, Pakistan. *J Oral Pathol Med.* 2009;38:416-421.

9. Akhtar S, Sheikh AA, Qureshi HU. Chewing areca nut, betel quid, oral snuff, cigarette smoking and the risk of oesophageal squamous-cell carcinoma in South Asians: A multicentre case-control study. *Eur J Cancer.* 2011.

10. Zakiullah, Saeed M, Muhammad N et al. Assessment of potential toxicity of a smokeless tobacco product (naswar) available on the Pakistani market. *Tob Control.* 2011.

11. Mazahir S, Malik R, Maqsood M et al. Socio-demographic correlates of betel, areca and smokeless tobacco use as a high risk behavior for head and neck cancers in a squatter settlement of Karachi, Pakistan. *Subst Abuse Treat Prev Policy.* 2006;1:10.:10.

12. World Health Organization. WHO report on the global tobacco epidemic; the MPOWER package. World Health Organization . 2008.

13. Tong EK, Glantz SA. ARTIST (Asian regional tobacco industry scientist team): Philip Morris' attempt to exert a scientific and regulatory agenda on Asia. *Tob Control.* 2004;13 Suppl 2:ii118-24.:ii118-ii124.

14. Shafey O, Dolwick S, Guindon EG, World Health Organization, American Cancer Society, International Union Against Cancer. *Tobacco control country profile*. 2nd ed. 2003.

15. World Health Organization. Country profiles, Pakistan. 25-8-2011.

16. Maziak W. The global epidemic of waterpipe smoking. *Addict Behav.* 2011;36:1-5.

17. Dobe M, Sinha DN, Rahman K. Smokeless tobacco use and its implications in WHO South East Asia Region. *Indian J Public Health.* 2006;50:70-75.

18. Imam SZ, Nawaz H, Sepah YJ, Pabaney AH, Ilyas M, Ghaffar S. Use of smokeless tobacco among groups of Pakistani medical students - a cross sectional study. *BMC Public Health.* 2007;7:231.:231.

19. Gupta PC, Ray CS. Smokeless tobacco and health in India and South Asia. *Respirology.* 2003;8:419-431.

20. Imam SZ, Nawaz H, Sepah YJ, Pabaney AH, Ilyas M, Ghaffar S. Use of smokeless tobacco among groups of Pakistani medical students - a cross sectional study. *BMC Public Health.* 2007;7:231.:231.

21. Gupta PC, Ray CS. Smokeless tobacco and health in India and South Asia. *Respirology.* 2003;8:419-431.

22. The World Bank: Pakistan. http://data.worldbank.org/country/pakistan accessed 04/07/2012..

23. Ahmad K, Jafary F, Jehan I et al. Prevalence and predictors of smoking in Pakistan: results of the National Health Survey of Pakistan. *Eur J Cardiovasc Prev Rehabil.* 2005;12:203-208.

24. www.who.int/fctc/en. WHO Framework Convention on Tobacco Control. Accessed 24/08/2012.

25. Giovino GA, Mirza SA, Samet JM et al. Tobacco use in 3 billion individuals from 16 countries: an analysis of nationally representative cross-sectional household surveys. *Lancet.* 2012;380:668-679.

26. Chartsbin (http://chartsbin.com/view/2438) and The World Bank. Country Income Groups (World Bank Classificatin). Accessed 21/08/2012.

27. Joshi P, Islam S, Pais P et al. Risk factors for early myocardial infarction in South Asians compared with individuals in other countries. *JAMA.* 2007;297:286-294.

28. Saleheen D, Zaidi M, Rasheed A et al. The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia. *Eur J Epidemiol.* 2009;24:329-338.

29. Kirkwood B, Sterne J. Logistic regression: controlling for confounding and other extensions. In: Blackwell Science, *Medical statistics*. 2 ed. Massachusetts, USA: 2003:212-3.

30. Iqbal R, Anand S, Ounpuu S et al. Dietary patterns and the risk of acute myocardial infarction in 52 countries: results of the INTERHEART study. *Circulation.* 2008;118:1929-1937.

31. Nanan, D. and White, F. PROCOR: 7/25/99: The National Health Survey of Pakistan: Review and discussion of report findings pertaining to selected risk factors for cardiovascular disease.

32. Ahmad K, Jafary F, Jehan I et al. Prevalence and predictors of smoking in Pakistan: results of the National Health Survey of Pakistan. *Eur J Cardiovasc Prev Rehabil.* 2005;12:203-208.

33. Giovino GA, Mirza SA, Samet JM et al. Tobacco use in 3 billion individuals from 16 countries: an analysis of nationally representative cross-sectional household surveys. *Lancet.* 2012;380:668-679.

34. Teo KK, Ounpuu S, Hawken S et al. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet.* 2006;368:647-658.

35. IRIN. Pakistan: anti-smoking law have little impact. *IRIN Humanitarian news and analysis.* 2004.

36. Ali NS, Khuwaja AK, Ali T, Hameed R. Smokeless tobacco use among adult patients who visited family practice clinics in Karachi, Pakistan. *J Oral Pathol Med.* 2009;38:416-421.

37. Teo KK, Ounpuu S, Hawken S et al. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet.* 2006; 368:647-658.

38. Raymond SU, Leeder S, Greenberg HM. Obesity and cardiovascular disease in developing countries: a growing problem and an economic threat. *Curr Opin Clin Nutr Metab Care.* 2006;9:111-116.

39. Nasir K, Rehan N. Epidemiology of cigarette smoking in Pakistan. *Addiction.* 2001;96:1847-1854.

40. Alam AY, Iqbal A, Mohamud KB, Laporte RE, Ahmed A, Nishtar S. Investigating socio-economic-demographic determinants of tobacco use in Rawalpindi, Pakistan. *BMC Public Health.* 2008;8:50.:50.

41. Nanan, D. and White, F. PROCOR: 7/25/99: The National Health Survey of Pakistan: Review and discussion of report findings pertaining to selected risk factors for cardiovascular disease. 25-7-1999.

42. Shera AS, Rafique G, Khawaja IA, Baqai S, King H. Pakistan National Diabetes Survey: prevalence of glucose intolerance and associated factors in Baluchistan province. *Diabetes Res Clin Pract.* 1999;44:49-58.

43. Lip GY, Malik I, Luscombe C, McCarry M, Beevers G. Dietary fat purchasing habits in whites, blacks and Asian peoples in England--implications for heart disease prevention. *Int J Cardiol.* 1995;48:287-293.

44. Misra A, Khurana L, Isharwal S, Bhardwaj S. South Asian diets and insulin resistance. *Br J Nutr.* 2009;101:465-473.

45. Singh RB, Niaz MA, Ghosh S et al. Association of trans fatty acids (vegetable ghee) and clarified butter (Indian ghee) intake with higher risk of coronary artery disease in rural and urban populations with low fat consumption. *Int J Cardiol.* 1996;56:289-298.

46. Hakeem R, Thomas J, Badruddin SH. Food habits and nutrient density of diets of Pakistani children living in different urban and rural settings. *J Health Popul Nutr.* 2002;20:255-263.

47. Yusuf S, Hawken S, Ounpuu S et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet.* 2004;364:937-952.

48. Ahmad K, Jafary F, Jehan I et al. Prevalence and predictors of smoking in Pakistan: results of the National Health Survey of Pakistan. *Eur J Cardiovasc Prev Rehabil.* 2005;12:203-208.

49. Nanan, D. and White, F. PROCOR: 7/25/99: The National Health Survey of Pakistan: Review and discussion of report findings pertaining to selected risk factors for cardiovascular disease. 25-7-1999.

50. Nanan, D. and White, F. PROCOR: 7/25/99: The National Health Survey of Pakistan: Review and discussion of report findings pertaining to selected risk factors for cardiovascular disease. 25-7-1999.

Chapter 8: Smoking and smokeless tobacco use and the risk of myocardial infarction in Pakistan

Summary

South Asians are particularly susceptible to premature CHD in their home countries and compared to their host population, when they migrate to Western industrialised countries. The epidemic of CHD amongst South Asians has been attributed to the increasing prevalence of known risk factors such as cigarette smoking, and some have hypothesized the importance of locally relevant factors such as the use of smokeless forms of tobacco. However, epidemiological studies on cigarette smoking have rarely been conducted in South Asian settings, and the relationship between alternative forms of tobacco use and MI risk remains uncertain. PROMIS represents the largest epidemiological resource for the study of indigenous forms of tobacco use and the risk of CHD, with over 12,000 first ever MI cases and age and sex frequency matched controls recruited in 5 urban centres across Pakistan. Odds ratios for non-fatal MI were 3.37 (3.05; 3.71) with current smoking only, 1.71 (1.46; 2.00) with current chewing only, and 1.46 (1.21; 1.78) with current snuff dipping only, when compared to never tobacco users. Combining several types of tobacco use conferred a risk of 3.91 (3.21; 4.76) with respect to never tobacco users. There was evidence of a dose-response relationship between both smoking and smokeless tobacco with the risk of MI, with the sharpest increase in risk experienced below 5 doses a day. These effects were not modified upon adjustment for local diet, use of *ghee* for cooking and socio-economic status. My analysis represents the first robust epidemiological evidence of a harmful effect of smoking and smokeless tobacco in the South Asian context. These findings reinforce the fact that urgency is needed to tackle the tobacco epidemic in Pakistan, and discourage the use of both smoking and smokeless forms of tobacco in all segments of the population, in order to alleviate the exploding burden of CHD in the region.

8.1 Background

Low and middle income countries now shoulder 80% of the burden of chronic diseases; and the South Asian continent, which hosts a quarter of the world's population, carries a substantial share of this burden [1,2]. In Pakistan, coronary heart diseases represent about 25% of all deaths and account for 150,000 deaths a year [1]. They threaten to counter major health gains achieved in the country over the past decades which have led to a prolongation of the life expectancy from 50 to 67 years between 1960 and 2009 [3]. They also represent a substantial economic cost for a middle low-income country of 165 million people [4].

The reasons for the disproportionately high burden of CHD in developing countries are multiple and complex. Compared to their Western hosts, South Asian migrants have been reported to shoulder a 40% to 60% higher risk of CVD [5] and experience higher vascular mortality [6,7] at younger ages : 5 to 10 years earlier than their Western counterparts on average [8]. Higher levels of conventional risk factors of CHD in the Pakistani population is likely to play a role [9]. South Asian migrants living in Western industrialised countries have a higher prevalence of diabetes and lower HDL-C levels than their host populations [9]. In Pakistan itself, it is estimated that a third of Pakistanis more than 45 years old have hypertension and the country ranks in the top 10 world nations for the prevalence of diabetes [10].

Smoking is another one of the "Western" risk factors which have spread to developing countries over the past decades and contributed to the epidemic of chronic disease. Already in 1983, Crofton was writing that "many countries are well on the way to adding a formidable *epidemic* of *smoking* related diseases to their already overwhelming health problems... All of us should feel a responsibility towards helping to prevent the tragedy." [11]. In South Asia, the per-capita consumption of manufactured cigarettes has been continuously increasing since the 1970s [12]. Smoking has traditionally been done using hand-rolled cigarettes called *bidis*, and water pipes such as *hookah* and *chilum*, and these local types of tobacco remain extremely popular [13]. In addition, over the past decades, tobacco companies have developed a large panel of manufactured tobacco products suited to South Asian taste such as *gutka*, by adding tobacco to traditional chewing and snuffing mixtures of betel and areca nut called *paan* [14] (see **Chapter 1 Section 1.2**).

Despite the need for large scale epidemiological studies investigating the relationship between these multiple forms of tobacco use and MI risk [2], previous studies on

South Asians have been characterized by small sample sizes, generally less than 1,000 cases and controls[15,16-18], conducted amongst South Asians living abroad [19] or have not taken into account ethnic specific differences within the South Asian subcontinent [20-22]. Most studies have focused on manufactured rather than hand-rolled cigarettes (*bidis*) which are popular in South Asian countries. Smokeless tobacco has been associated with CVD risk in Western countries [23,24], where practices and compositions of tobacco products are likely to differ from the types which are used in Pakistan. Most of the studies were based in Sweden or the USA where the main type of smokeless tobacco used is a type of snuff called *snus* with a specific heating procedure of tobacco leaves destined to limit its carcinogenic effects. In comparison, chewing is more popular amongst Pakistanis than snuff dipping and the type of snuff used differs in composition and is called *naswar* [25].

In this context, there is a need to investigate the association between different smoking and smokeless types of tobacco use with the risk of cardiovascular diseases in South Asia. PROMIS represents the largest case-control study of first-onset acute MI conducted in South Asia with a detailed lifestyle questionnaire enquiring about types of tobacco use in Pakistan as well as biochemical information. The present analysis aims to better assess than the published literature the effect of cigarette and indigenous modes of tobacco use in the context of traditional and locally relevant biochemical, medical and lifestyle factors on the risk of MI in urban Pakistanis. As tobacco usages are relatively similar across South Asia [26], findings from this analysis could have important public health implications for prevention of CVD in the whole region.

8.2 Method

8.2.1 Participants

Details of PROMIS participants have been described in **Chapter 6**. The analysis included 6,051 first ever MI cases and 6,871 age and sex frequency matched controls enrolled by March 2011 who were recruited from 6 centres located in 5 cities in Pakistan. Tobacco use was categorized as “never user”, “ex-user”, “current user of dipping tobacco only (*naswar*)”, “current user of chewing tobacco only (*paan*, *gutka* or *supari*)”, “current smoker only (manufactured cigarettes or *bidis*) and “current smoker and smokeless user (chewing or dipping tobacco)”. To investigate the dose-response relationship, number of manufactured or hand rolled cigarettes per day was categorised as 0 “non smoker”, <5, 5-10, 10-15, 15-20 and ≥20 per day for current

smokers. Smokeless tobacco amount was categorised as “non smokeless user”, <5, 5-10, 10-15 and ≥15 doses per day for current smokeless users.

8.2.2 Statistical methods

Odds ratios (ORs) were calculated using unconditional logistic regression as cases and controls were frequency matched by sex and 5 year age bands rather than individually matched. The general form of a logistic regression is:

$$\ln(\frac{p}{(1-p)}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k$$

Where *p* represents the probability of experiencing an MI, $\frac{p}{(1-p)}$ represents the odds of experiencing an MI, $x_1...x_k$ are *k* exposure variables and $\beta_1....\beta_k$ are regression coefficients associated with the *k* exposure variables. In this model, β_0 corresponds to the odds of disease in the baseline group and $\exp(\beta_i)$ corresponds to the odds ratio for exposure *i (i>0)*.

Odds ratios were adjusted at least for matching and demographic variables (age, sex, recruitment centre and ethnicity); and when indicated, also adjusted for conventional risk factors of MI: history of diabetes or hypertension, family history of MI, waist to hip ratio and LDL-C. ORs investigating the relationship between smoking amount (respectively smokeless amount) and MI risk were also adjusted for smokeless (respectively smoking) tobacco use categorized as “current”, “past” or “never” user. Further analyses investigated the effect of progressive adjustment for diet (high protein and sweet diet; and high vegetable and carbohydrate diet), the use of *ghee* – clarified butter used in traditional cooking and high in trans-fatty acids [27] - versus oil for cooking; and socio-economic status categorized into low, middle and high social class. Dietary patterns and socio-economic status were identified by performing principal component analyses (see **Appendices 2 & 3**). The dietary analysis was done on 43 questions from a locally validated food frequency questionnaire capturing dietary habits of Pakistanis; and the analysis of socio-economic status was done using monthly income, ownership of household items and formal education.

For ORs of tobacco uses with MI risk, never tobacco users were chosen as the reference group. For the computation of ORs with smoking/smokeless amount, never smokers/smokeless users were chosen as reference groups and past tobacco users were excluded.

To characterize shapes of associations, ORs were calculated within pre-defined categories of smoking and smokeless amount and were plotted against mean values within each quintile or category. Graphs representing odds ratios used log-linear scale to allow graphical assessment of the linearity of dose-response relationships. To enable graphical comparison of any two ORs and not only with the reference group, 95% confidence intervals were drawn using "floating absolute" variances, which are extracted after transforming the matrix of variance-covariance of the coefficients under the constraint of nullity of the transformed covariances [28].

Effect-modification was investigated graphically and analytically for age group (≤50 years old compared to >50 years old), sex, ethnicity, socio-economic status, dietary patterns, medical history and metabolic factors. Forest plots displayed ORs for current and ex versus never tobacco users by subgroups. Never tobacco users who were also either ≤50 years old, Men, Punjabi, in the bottom third of LDL-C, in the bottom third of WHR, in the bottom third of socio-economic status, below median for the consumption of vegetables and carbohydrate diet, or below median for the consumption of meat and protein diet, were chosen as reference groups. Formal tests of effect modification were performed by adding interaction terms to logistic regression models and then tested for significance. To maximize power, continuous variables in their original forms (rather than divided around the median or into thirds) were used for the tests of interaction. To take into account the large number of tests, p-values of interaction <0.001 were emphasized. Because there were relatively few women smoking cigarettes or *bidis* (41 cases and 30 controls), it was not possible to investigate the effect of smoking (only or in combination with smokeless products) by sex.

Sensitivity analyses explored heterogeneity across centres of recruitment and by source of controls. In a first analysis, ORs were computed within each centre and pooled by meta-analysis rather than adjusted for centre in a single model. A second analysis was performed when controls in a specific centre were recruited from different sources (>100 controls for a specific source). In that case, ORs comparing a specific type of controls to cases in each were computed, to allow comparison of ORs depending on the source of controls. Heterogeneity levels between estimates for different centres or sources were assessed by reporting I^2 values and confidence intervals [29]. All analyses were conducted using STATA v10 (StataCorp, Texas) and used code purposely developed.

8.3 Results

Cases and controls have been described in **Chapter 6**; and correlations between tobacco use and other lifestyle and medical risk factors have been described in **Chapter 7**. The average age of cases was 53.3 years old (SD: 10.2) and 83% were men. Amongst cases, 41% reported being never tobacco users, 6% ex-users, 4% snuff dippers only, 6% current chewers only, 36% current smokers only; and 6% combined smoking and smokeless use.

8.3.1 Association of smoking with MI risk

Current cigarette or bidi smoking was most strongly associated with MI risk. Current cigarette and bidi smokers versus never tobacco users had an OR for non-fatal MI of 3.37 (3.05; 3.70), after adjustment for demographic and conventional risk factors (**Figure 8.1a**). Associations were unchanged upon further adjustment for diet and for socio-economic status (**Figure 8.1b**). There was a non linear positive association between the number of cigarettes or *bidis* smoked per day and MI risk which was not altered upon adjustment (**Figure 8.2a & 8.3a**). Compared to never users, individuals smoking <5 cigarettes or *bidis* per day had an OR (95% CI) of 1.76 (1.45; 2.13), smokers of 5-10 cigarettes or *bidis* per day an OR of 2.51 (2.03; 3.10), smokers of 10-15 cigarettes or *bidis* per day an OR of 2.63 (2.24; 3.07), smokers of 15-20 cigarettes or *bidis* per day an OR of 2.75 (2.05; 69) and smokers of ≥20 cigarettes or *bidis* per day an OR of 4.33 (3.84; 4.87).

The effect of current smoking of cigarettes or bidis was higher in younger age groups (≤50 years old) compared to older age groups (>50 years old): ORS of 3.93 (3.46; 4.46) versus 2.88 (2.58; 3.22), compared to never tobacco users (**Figure 8.4**). The association did not differ according to sex or ethnic group (p-value >0.001). Smoking was associated with MI risk in individuals independently of the presence or absence of a history of hypertension (**Figure 8.5**). Individuals in the top third of the distribution of LDL-C who smoked had an OR of 6.08 (5.24; 7.05) compared to never smokers in the bottom third of LDL-C, and individuals in the top 3rd of WHR had an excess risk of 5.41 (4.62; 6.34) compared to never smokers in the bottom 3rd of the distribution of WHR (**Figure 8.6**). Smoking was most harmful for individuals with low socio-economic status compared to individuals with high socio-economic status: 4.31 (3.78; 4.91) versus 3.11 (2.72; 3.57) (**Figure 8.7**). There was no evidence of effect modification by diet.

There was no significant heterogeneity in ORs between recruitment centre and depending on the source of controls (p-value of heterogeneity: 0.64) (**Figure 8.8**). ORs were comprised of values between 2.8 and 4.6 with overlapping confidence intervals and I^2 was equal to 0. Pooled RRs from fixed effects and random effects meta-analyses of centre specific ORs were similar and did not differ from ORs when using a one-step rather than a two-step approach (**Figure 8.9**).

8.3.2 Association of smokeless tobacco with MI risk

Ninety percent of the individuals who chewed tobacco declared using *paan* rather than other chewing products (*supari* and *gutka*). Only one type of snuff, called *naswar*, was recorded in PROMIS. Compared to never-consumers of tobacco, ORs for non-fatal MI were 1.46 (1.20; 1.77) with current snuff dipping and 1.71 (1.46; 2.00) with current chewing after adjusting for demographic and conventional risk factors (**Figure 8.1a**). Individuals who currently used smokeless tobacco on top of smoking had a greater risk than individuals who currently only chewed or dipped and their risk was slightly higher than that of current smokers alone: ORs were 3.91 (3.21; 4.76) for individuals using smoking and smokeless tobacco, versus 3.36 (3.05 to 3.71) for smokers alone. The associations were independent of known lifestyle and medical risk factors of MI, and were minimally affected by adjustment (**Figure 8.1b**).

Looking at the shape of association in relation to smokeless amount, there was a positive association up to 10 smokeless products a day and a non significant decrease in risk afterwards. ORs were 1.32 (1.16; 1.52) for users of <5 smokeless products a day, 1.77 (1.43; 2.20) for 5-10 products a day, 1.74 (1.26; 2.41) for 10-15 products a day and 1.49 (1.01; 2.19) for ≥15 products a day (**Figure 8.2 & 8.3**).

ORs for chewing and dipping tobacco did not significantly differ across age groups, by sex, and by ethnic group (p-values>0.001) (**Figure 8.4**). Individuals who dipped tobacco and were hypertensive had an OR of 3.12 (2.30; 4.24) compared to normotensive and never users of tobacco (**Figure 8.5**). There was evidence suggestive of a stronger increase in MI risk when using smokeless tobacco and belonging to the top third of socio-economic status compared to the bottom third (**Figure 8.7**).

There was evidence of heterogeneity across centre and source of controls for the effect of smokeless tobacco use (chewing and dipping tobacco combined) versus never users: I^2 was equal to 54.8% and ORs for non-fatal MI ranged from 1.05 to

5.32 (**Figure 8.8**). However, random and fixed effects meta-analyses produced similar ORs, which were significant and only slightly lower than the 1-step estimate: pooled OR by random effect meta-analysis was 1.53 (1.14; 2.05) and by fixed effects meta-analysis it was 3.38 (3.04; 3.76) (**Figure 8.9**). Level of heterogeneity was low for the ORs comparing smoking and smokeless tobacco users versus never users: I^2: 37.5% and p-values: 0.099.

8.3.3 Ex-tobacco users

Ex-users of all forms of tobacco experienced a slightly increased risk compared to never users, which was significantly lower than current users of any type of tobacco. The OR for past versus never users was 1.22 (1.04; 1.42) (**Figure 8.1**). There was no significant effect modification according to age, sex or ethnicity (**Figure 8.4**). The increase in risk due to past tobacco use was significant in individuals who otherwise had a relatively low risk of CVD: non diabetics, non hypertensive, individuals without a history of MI; whilst it did not carry extra-risk in individuals who were already at high risk because they were diabetics, hypertensive or reported a family history of MI (**Figure 8.5**). The level of heterogeneity was below 50% for ex versus never users (**Figure 8.8**). In a sensitivity analysis, pooling centre-specific ORs gave a pooled estimate of 1.31 (1.08; 1.58) by random effects meta-analysis, and 1.30 (1.10; 1.52) by fixed effects meta-analysis (**Figure 8.9**).

8.4 Discussion

My analysis represents the first robust epidemiological evidence of a harmful effect of smoking and smokeless tobacco on the risk of CVD in the South Asian context. The findings contained in this Chapter should alert South Asian governments to the urgency needed to tackle both the smoking and smokeless tobacco epidemics happening in their countries. They may also be generalisable to individuals of South Asian origins who currently live in Western countries such as the USA and the UK and who have retained the tobacco habits of their countries of origin.

8.4.1 Strengths of the analysis

I studied smoking in relation to MI risk in a large Pakistani urban population using a case-control study design. This analysis included 6,050 cases and 6,871 controls from Pakistan, making it more than 3 times as large as the recent INTERHEART case-control study on risk factors for MI risk in South Asians [8]. INTERHEART also did not investigate the effect of dipping tobacco, only that of chewing tobacco. Studies on dipping tobacco have all been conducted in Northern Europe and North

America, where products and toxic contents differ from South Asian products [23,24]. PROMIS individuals had detailed lifestyle and biochemical information on a range of risk factors for CVD, including locally relevant factors such as dietary habits and *ghee* consumption. Analyses were adjusted for a large number of potential confounders and I performed tests of effect modifications of both smoking and smokeless tobacco with other cardiovascular risk factors. Alcohol was not included in the list of potential confounders. This is because alcohol consumption is religiously prohibited to Muslims and as 95% of PROMIS individuals declared being Muslim alcohol consumption may be subject to misreporting.

Individuals were drawn from all major ethnic groups of Pakistan, enabling comparison across ethnic groups, and making my results applicable to neighbouring South Asian countries where these ethnicities are represented. PROMIS was conducted in 5 urban centres across Pakistan. Currently 37% of the Pakistani population and 30% of the South Asian population live in urban areas [30,31]. Urban Pakistanis tend to favour smoking products, in particular manufactured cigarettes, while Pakistanis living in rural areas have especially high rates of smokeless tobacco and overall higher rates of tobacco prevalence than urban areas [32,33]. Urbanization is happening at a fast rate in Pakistan and in South Asia in general and therefore smoking habits and the use of manufactured cigarettes are likely to keep expanding with economic development, unless drastic political measures are taken to curb the trend [34].

On the other side, smokeless tobacco has a long history and is unlikely to regress in the immediate future in South Asia [13,35]. It is reportedly more acceptable than smoking amongst children, teenagers and women [35]. Low socio-economic status groups favour smokeless tobacco, and especially snuff dipping, which tends to be cheaper than manufactured cigarettes. The commercial production and marketing of smokeless tobacco products has also promoted a rapid increase in sale over the past decades. In India, per capita smokeless tobacco consumption has increased among the poor since the 1960s in both rural and urban areas and the total number of smokeless tobacco users in India and Pakistan combined has been estimated to number 100 million individuals [18,19]. The rate of growth of *gutka* , a manufactured product with long shelf life, has overtaken that of smoking forms of tobacco in India [26]. In this context, the findings of this study have wider implications for the whole of the South Asian region where similar smoking and smokeless products coexist.

8.4.2 Smoking tobacco and MI risk

First, I have confirmed an association between cigarette smoking and the risk of MI in a South Asian population. Current smokers experienced an OR of 3.36 (3.05; 3.70) compared to never tobacco users. The strength of association was similar to that of previous reports from case-control studies [20], and was relatively higher than the summary estimate of a doubling in risk in Western populations I obtained analysing the ERFC (see **Chapter 4**). Several factors may account for this observed difference. ORs were estimated here because of the retrospective nature of the PROMIS study, whilst HRs were estimated in the ERFC dataset which uses a prospective design. Cases enrolled in PROMIS may be subject to recall bias and may be more likely to report their smoking history than controls, inflating the ORs; whilst the prospective nature of the ERFC ensures information on exposure is collected before individuals experience their first cardiovascular event. This difference could also be due to the pooling of cigarettes with bidi smokers in PROMIS, whilst only cigarettes smokers were considered in the ERFC. This difference could finally reflect real differences in susceptibility to CVD in Pakistanis compared to Western European population when smoking.

There was evidence of a dose-response relationship between the number of cigarettes or *bidis* smoked per day and MI risk. The relationship was non linear, which is in agreement with findings from the ERFC dataset, with the highest relative increase experienced for smokers of 1-5 cigarettes per day versus never smokers corresponding to ~80% higher risk. By comparison, INTERHEART reported an approximately linear increase of 6% higher risk per additional cigarette [20]. The association was independent of conventional risk factors as well as economic status and diet, showing that clustering of poor health records and unhealthy diet and low socio-economic status cannot account for this contrast.

ORs were higher in younger individuals. This is in agreement with published findings that younger age groups are at higher risk when they smoke [37] and with my findings in developed populations (see **Chapter 3**). There was similarly no effect modification according to sex, in contrast to a literature based meta-analysis which found a ~25% greater increase in risk for women versus men [5]; and in agreement with ERFC results and findings from INTERHEART.

8.4.3 Smokeless tobacco use and MI risk

Secondly, these results provide a novel insight into the epidemiology of smokeless tobacco in relation to CHD in a South Asian context. Smokeless tobacco is known to be responsible for cancer and tobacco use has been estimated to account for about

50% of oral cancers in India [25]. With the growth of *gutka*, the incidence of oral submucous fibrosis has reached epidemic proportions in India among individuals below 35 years old. Because tobacco is not burned and therefore there is no production of carbon monoxide, smokeless tobacco is often thought of as having no cardiovascular effect, and it is commonly used as a breath freshener or as toothpaste [38,39].

In my analysis, both chewing and dipping tobacco were strongly associated with MI risk. The magnitude of the estimate for chewing tobacco is in agreement with that of INTERHEART but with a narrower confidence interval in my analysis which may be attributable to greater homogeneity of chewing products used in Pakistan compared to the rest of the world [20]. Indeed, INTERHEART aggregated data from participants located in 52 countries located in all continents around the world and chewing tobacco is sold in various forms and shapes which are likely to differ in toxicity. A recent meta-analysis of chewing and dipping products combined, where most of the evidence came from Northern American studies of chewable tobacco, also demonstrated a significant association with smokeless tobacco of a relatively lower magnitude: the RR for MI was 1.13 (1.03; 1.21) for all smokeless products[23].

An association between snuff dipping and MI risk is being debated in the literature and most of the evidence comes from studies conducted in northern Europe [40]. A recent meta-analysis selecting only Swedish data found a non significant increase in MI risk for snuff users (HR of 1.04; 95% CI 0.93 to 1.17) [41]. Possible explanations for the presence of a significant association in Pakistan and non significant association in Western countries with regard to MI risk include true heterogeneity of results by geographic area due to differences in the composition of the products used in the USA, Northern Europe and South Asia. The major components of tobacco snuff are alkaloids, with nicotine as the main compound (85-95% of total alkaloids) [25]. During product manufacturing, tobacco leaves, stems and other ingredients are blended to achieve a specific nicotine content, pH, taste, flavour and aroma. The pH strongly affects the concentration of bioavailable nicotine, whereas the nitrite content affects nitrosamine concentrations in the product. The major form of smokeless tobacco used in Sweden and the US is *snus* which is made of air-cured and fire-cured tobacco, flavoured and powdered into fine particles, containing 20-55% moisture by weight. Products used in the USA have higher nitrosamine content than those in Northern Europe and, as a result, an increased risk of oral cancer for use of smokeless tobacco has been reported in the USA and not in Nordic countries [25]. The Pakistani form of tobacco dipping, *naswar,* contains tobacco which undergoes a

different curing process than *snus*, and is typically less moist and has higher levels of nitrosamines [26].

The magnitude of excess risk when chewing or dipping tobacco was relatively strong in my dataset, close to 50% increased risk for naswar dipping and over 70% increased risk for chewing tobacco. This was smaller than the OR associated with smoking products, but still highly significant. Individuals who combined the use of smokeless products with that of smoking products did not experience any benefit compared to using smoking products alone in terms of MI risk. A study conducted in the US found that the use of smokeless tobacco may lead to subsequent cigarette smoking. Young males who were not smokers but regularly used smokeless tobacco were more than three times as likely as never users to become current smokers within the next four years [42]. In PROMIS, individuals belonging to the control group who were current chewers were also more likely to smoke than never chewers (24% versus 20%), and individuals who had stopped chewing were also more likely to have stopped smoking (30% of ex-chewers were ex-smokers, compared to 5% of current chewers) (p-value χ^2 of independence between smoking and chewing status <0.001). Education on the dangers of all forms of tobacco use should encourage individuals to stop smoking and using smokeless forms of tobacco, rather than (partially) switch from one form to another, without any real benefit on cardiovascular health.

These findings also suggest a role for toxins that are intrinsic to tobacco itself, and not just confined to the smoked form. Animal experiments and studies in human have found short term effects of smokeless tobacco on the hemodynamic system. Nicotine, which is present in all forms of tobacco use, has been shown in animal experiments to produce arrhythmias and ventricular fibrillations (see **Chapter 1 Section 1.4.1**) and may be one mediator of smokeless tobacco effect on MI risk. Other mechanisms include acute elevation of blood pressure, chronic hypertension and acute activation of the sympathetic nervous system [24]. Higher blood pressure levels have been found in smokeless tobacco users as well as a higher prevalence of hypertensive individuals, in a Swedish cross-sectional study including more than 30,000 individuals [43]. Smokeless tobacco products also have high amounts of sodium chloride, which could contribute to inflammation and tumour promotion [25]. In my dataset, individuals who were diabetics or hypertensive showed a non significant increase in risk of MI when they used smokeless tobacco products. Some of this attenuation could be due to recall bias in non diabetics and normotensive, especially as these variables were self-reported.

8.4.4 Past use of tobacco and MI risk

Ex-users experienced a slightly increased risk, which was non significant amongst younger age groups. The magnitude of the OR for ex-users was below the OR for current users of any tobacco type. Stopping the use of tobacco needs to be more actively encouraged in Pakistan and more largely in South Asia. In PROMIS, ex-smokers represented between 5% and 10% of controls depending on ethnic group, three or four times less than the proportion of current tobacco users in these groups. In women controls, <2% indicated having stopped tobacco use, and the proportion in men was below 10% (9%). Implementing tobacco regulations has not been really successful so far in the South Asian subcontinent. In India, legislation mandating pictorial warnings on smokeless and smoking tobacco packaging is not effective because of inappropriate pictures chosen and irrelevant text messages [44]. In Pakistan, anti-smoking laws which ban consumption in public areas and storage near educational institutions were passed in 2009 [45], but are not being enforced [46]. This analysis provides evidence supporting public health campaigns to encourage cessation of all forms of tobacco use.

8.4.5 Limitations

Despite its strengths, this analysis contains several limitations. Controls came from 3 different sources, with nearly ¾ of controls recruited amongst visitors of patients from the outpatient department. However, estimates were relatively homogeneous across different types of controls within the same centre and across different centres. Heterogeneity levels were relatively low, and estimates using a one step approach were in broad agreement with estimates obtained from a two-step approach combining centre-specific ORs by random and fixed effects meta-analyses. Case-control studies are prone to recall bias due to MI cases being more likely to overestimate their past smoking and smokeless tobacco use, and MI controls being more likely to under-estimate it, leading to an over-estimation of ORs. As a result, case-controls are unable to prove causality and a prospective design would be needed to strengthen the evidence of a causal association between smokeless tobacco and MI risk.

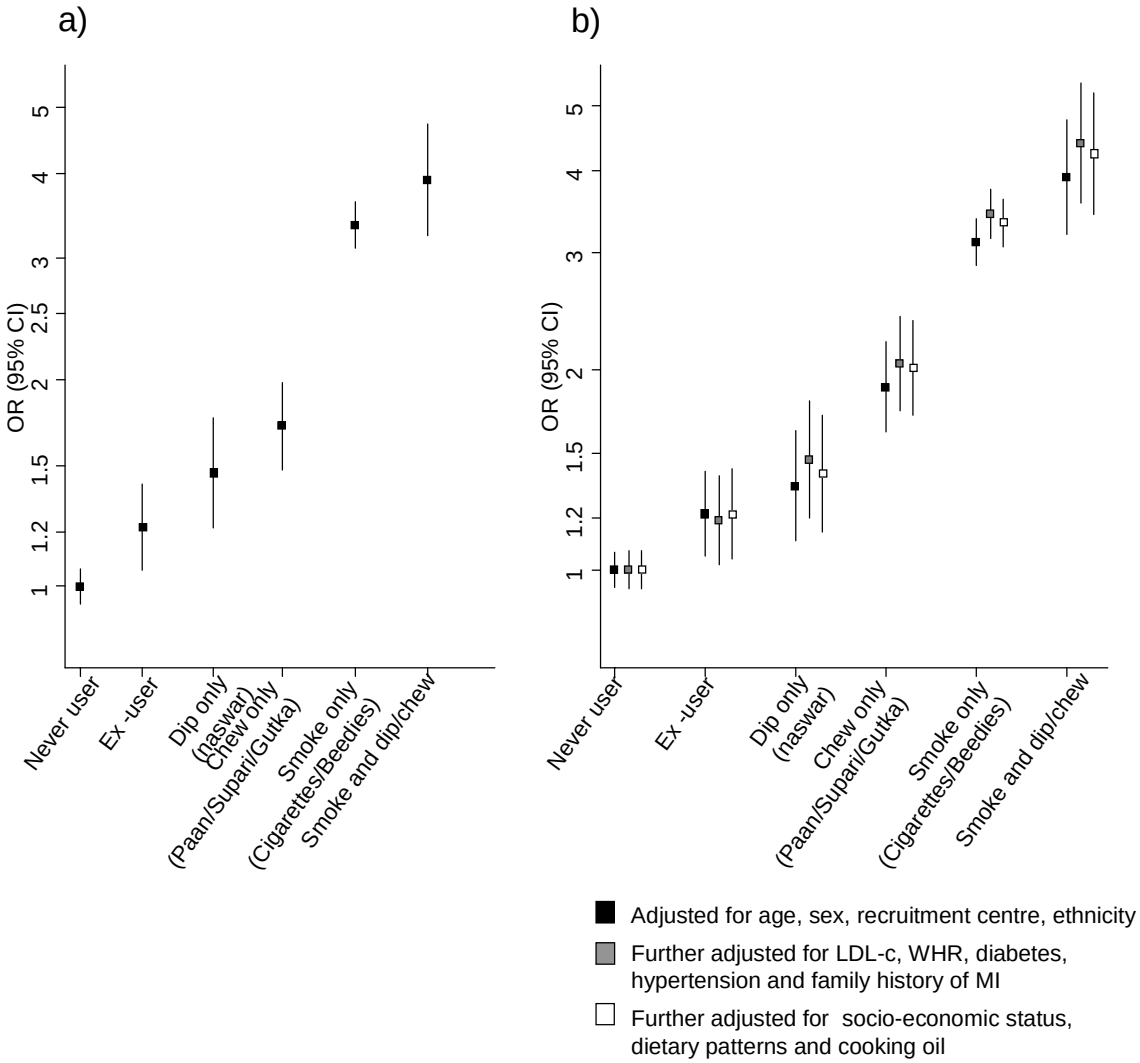
Information was not available on the type of tobacco used by past users. The brevity of the questionnaire on tobacco use meant previous information was not asked such as duration of tobacco use, age starting and, in ex-users, age stopped and number of years since stopped. I was also not able to investigate cigarettes and *bidies*

separately, as amount was not available separately for manufactured and hand rolled cigarettes. Smokeless products can be homemade or sold in pouches and quantities of tobacco are not standardized across products and brands. Therefore, it was not possible to accurately investigate the relationship between tobacco dose of smokeless tobacco and risk of MI. I had to group all chewing types of tobacco (*paan*, *supari* and *gutka*) together into a single category "chewing tobacco" to maximize power. Finally, analyses were adjusted for a large number of conventional and locally relevant risk factors. In particular, I adjusted analyses for local dietary pattern, *ghee* consumption and several variables representing socio-economic status. However, some risk factors were not available for these analyses. Blood pressure was for example measured but considered as unreliable in cases that had just been administrated stabilizing drugs after experiencing an MI. Self reported history of hypertension was used instead.

8.5 Conclusion

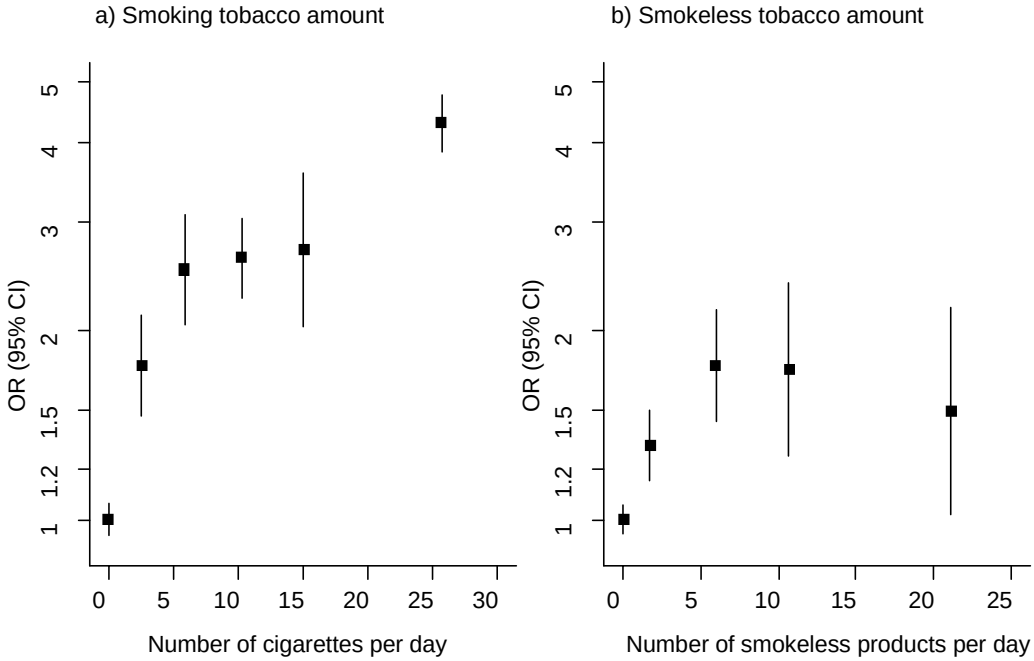
In Pakistan, all forms of tobacco use are hazardous to cardiovascular health. The increase in risk was independent of conventional risk factors, diet and socio-economic markers. The effect of chewing and dipping tobacco was intermediate between that of non tobacco users and that of smoking. These findings should help devise strategies to address the increasing burden of tobacco related CVD in South Asia.

Figure 8.1: Odds ratios for myocardial infarction with tobacco usage a) with adjustment for conventional factors, and b) showing progressive adjustment



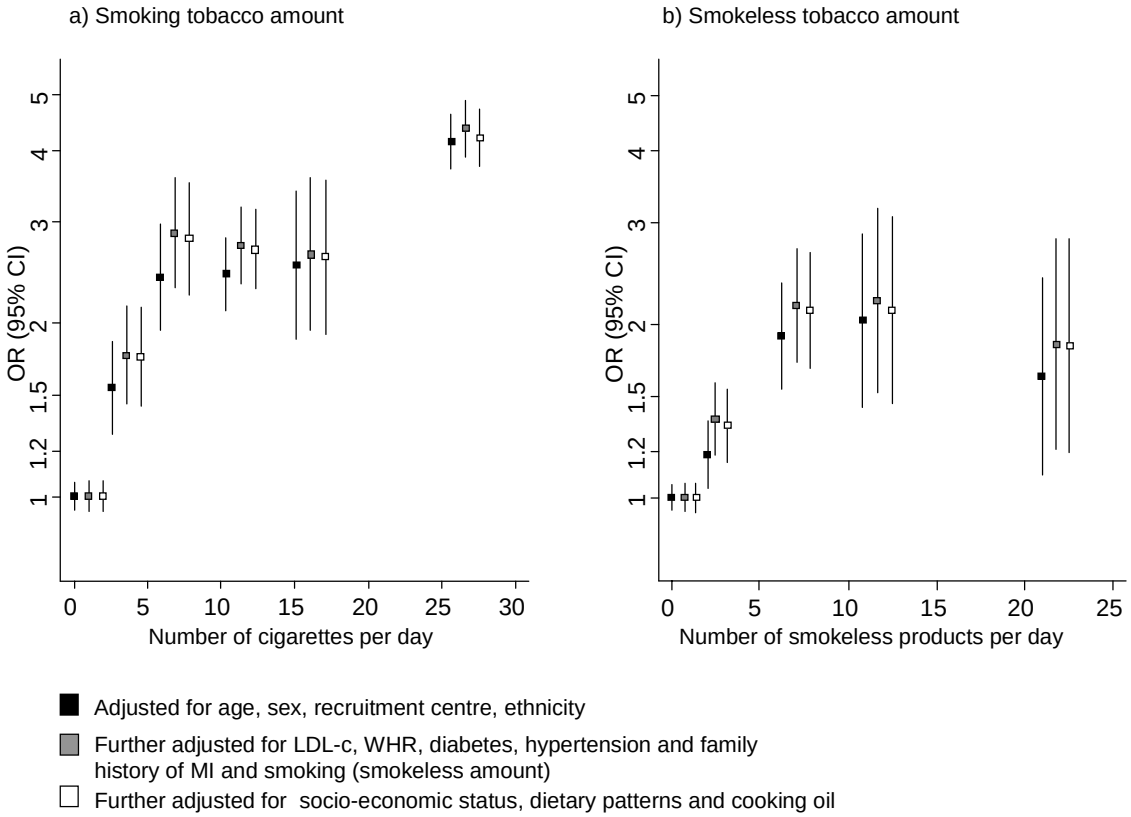
OR: Odds ratio; 95% CI: 95% Confidence interval plotted using "floating absolute variances". "Conventional risk factors" of adjustment include age, sex, ethnicity, recruitment centre; self reported diabetes or hypertension, LDL-C levels, WHR and family history of MI. In figure a), ORs (95% CI) compared to never users were 1.22 (1.04; 1.42) for ex-users, 1.46 (1.20; 1.77) for dipper only, 1.71 (1.46; 2.00) for chewers only, 3.36 (3.05; 3.71) for smokers only and 3.91 (3.21; 4.76) for individuals who dipped or chewed tobacco and also smoked. For figure b), the dataset was restricted to 5,365 cases and 5,557 controls with information on all the covariates of adjustment.

Figure 8.2: Odds ratios for myocardial infarction with a) smoking and b) smokeless tobacco amount



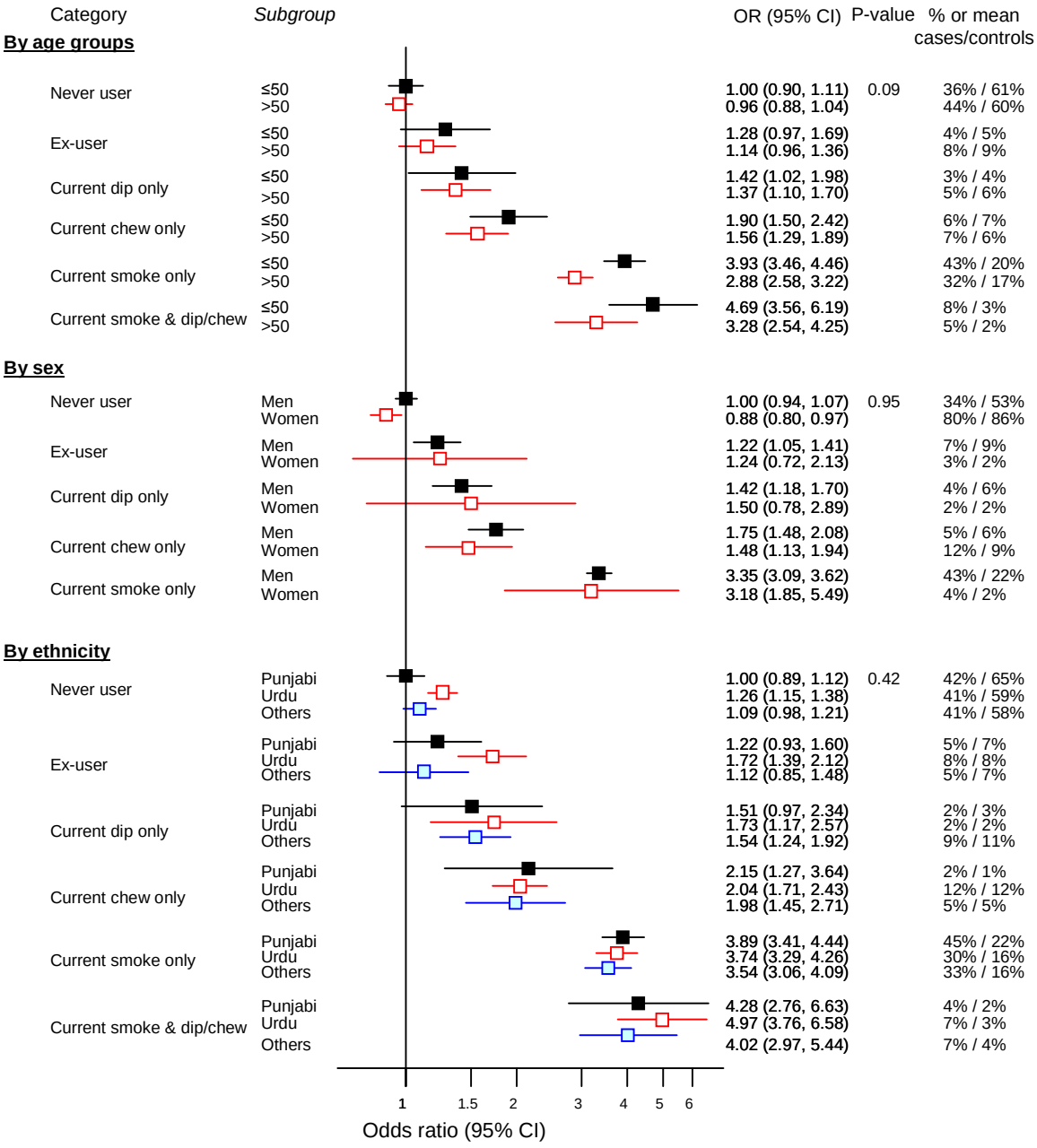
OR: Odds ratio; 95% CI: 95% Confidence interval plotted using "floating absolute variances". Models were adjusted for age, sex, ethnicity, recruitment centre; self reported diabetes or hypertension, LDL-C levels, WHR, family history of MI and a) chewing and dipping status or b) smoking status. ORs were 1.32 (1.16; 1.52) for users of <5 smokeless products a day, 1.77 (1.43; 2.20) for 5-10 products a day, 1.74 (1.26; 2.41) for 10-15 products a day and 1.49 (1.01; 2.19) for ≥15 products a day.

Figure 8.3: Progressive adjustment of the associations of a) smoking and b) smokeless tobacco amount with risk of myocardial infarction



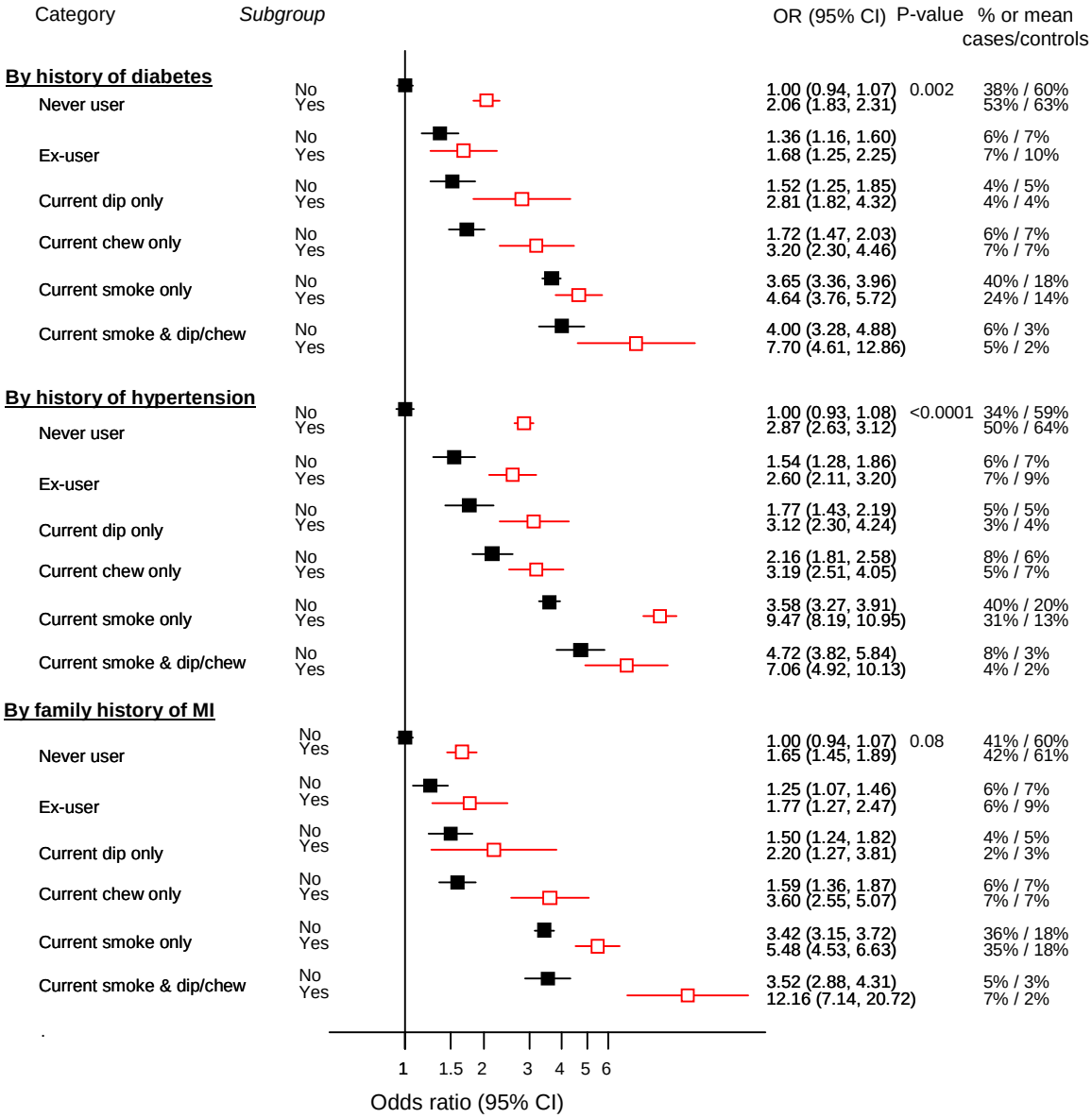
OR: Odds ratios and 95% confidence interval. Dataset restricted to 5,365 cases and 5,557 controls with information on all the covariates of adjustment. Odds ratios were computed within categories of number of cigarettes or bidis per day smoked by current smokers (0: non smokers, <5, 5-10, 10-15, 15-20 and ≥20) or categories of smokeless tobacco (0: non smokers, <5, 5-10, 10-15, and ≥15)) and plotted against the arithmetic mean within each category.

Figure 8.4: Association of tobacco use with MI risk by socio-demographic characteristics



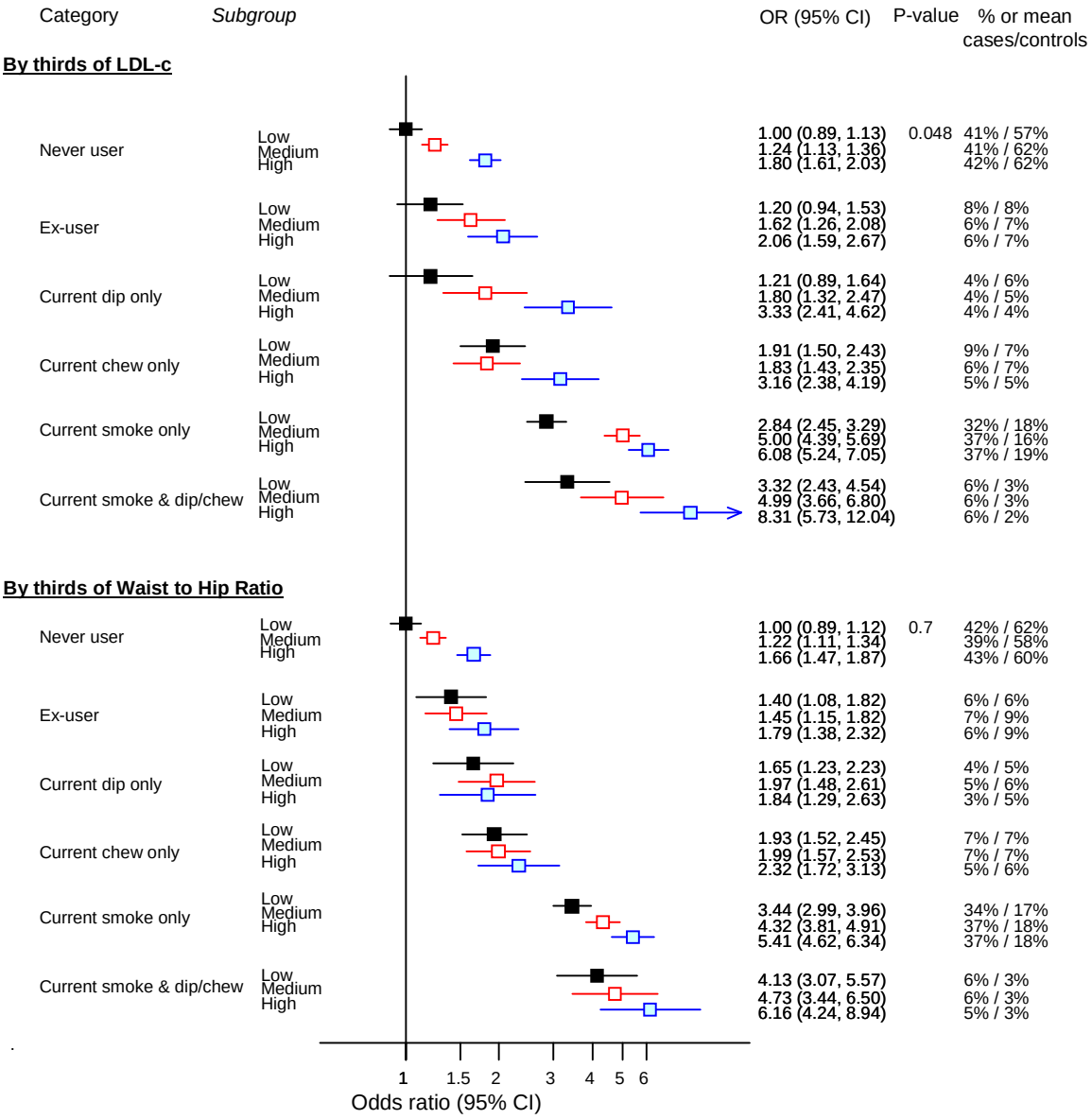
OR: Odds ratios; 95% CI: 95% confidence interval, P-value: p-value of interaction. Confidence intervals were plotted using "floating absolute" variances for tobacco use. Models were adjusted for conventional risk factors: age, sex, ethnicity, centre of recruitment, LDL-C, WHR, history of diabetes or hypertension and family history of MI. P-values of interaction are derived from tests of significance of the interaction terms added to the model. The proportion of women smoking and dipping or chewing was too small to enable investigation of potential interactions with sex on MI risk. Black boxes correspond to the reference group (age 50 and below, Men, Punjabi); whilst blue and red boxes indicate other groups.

Figure 8.5: Association of tobacco use with MI risk according to medical history



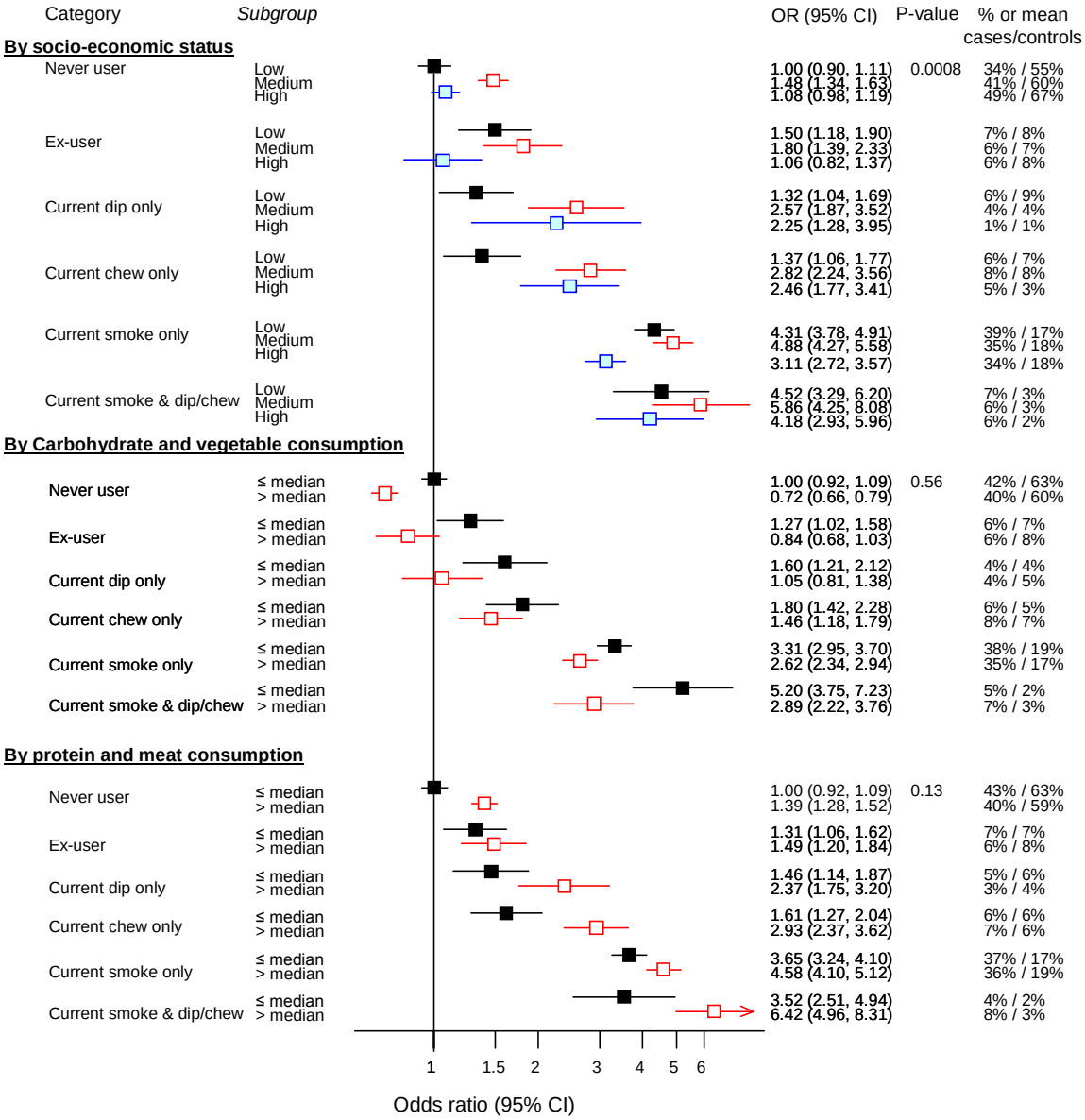
OR: Odds ratios; 95% CI: 95% confidence interval. Confidence intervals were plotted using "floating absolute" variances for tobacco use. Models were adjusted for conventional risk factors: age, sex, ethnicity, centre of recruitment, LDL-C, WHR, history of diabetes or hypertension and family history of MI. P-values of interaction are derived from tests of significance of the interaction terms added to the model. Black boxes correspond to the reference group (individuals without a medical history of either diabetes, hypertension of family history of MI); whilst red boxes indicate other groups (individuals with a history of diabetes, a history of hypertension or a history of MI).

Figure 8.6: Association of tobacco use with MI risk according to levels of LDL-C and Waist to Hip Ratio



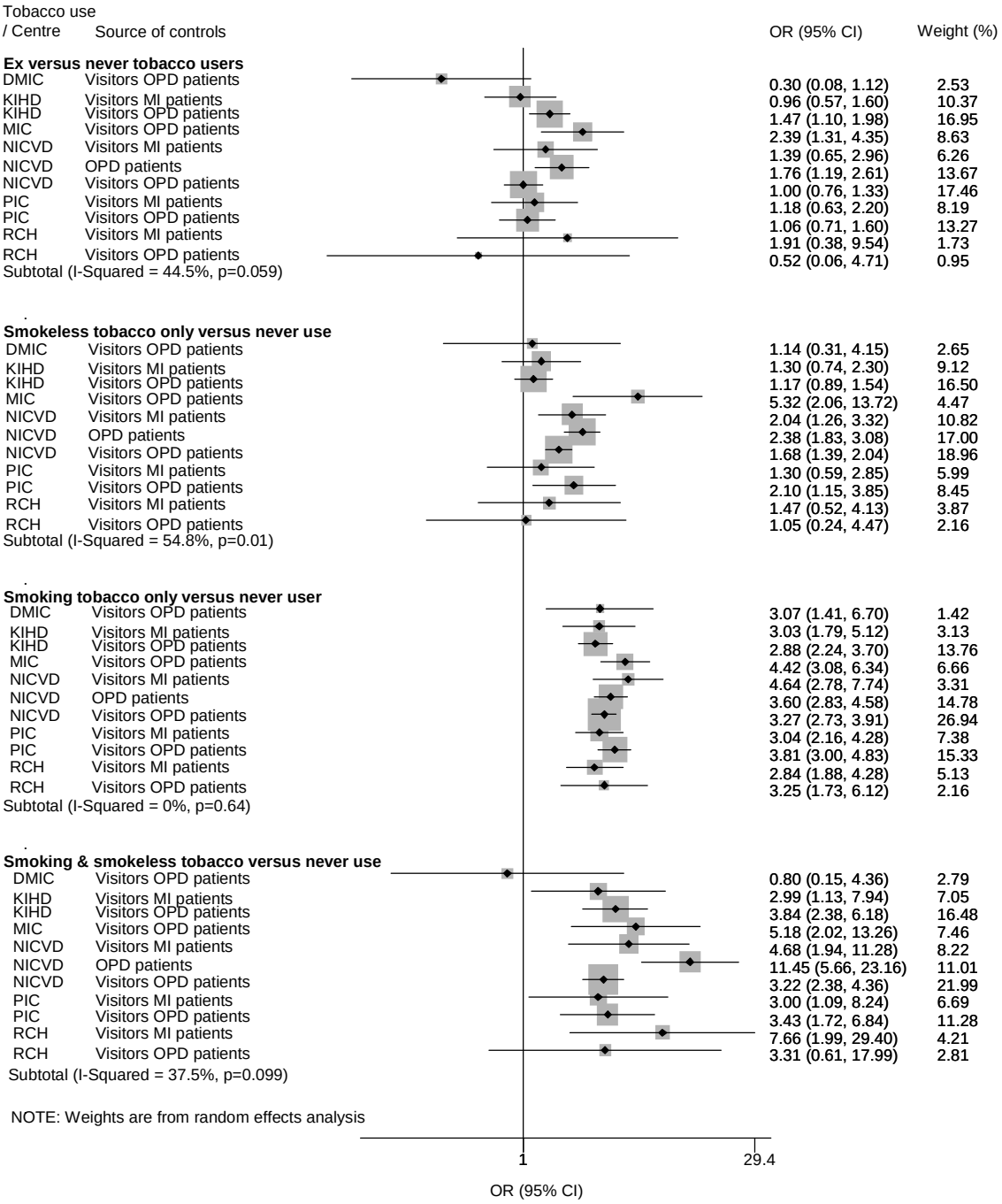
OR: Odds ratios; 95% CI: 95% confidence interval. Confidence intervals were plotted using "floating absolute" variances for tobacco use. Models were adjusted for conventional risk factors: age, sex, ethnicity, centre of recruitment, LDL-C, WHR, history of diabetes or hypertension and family history of MI. P-values of interaction are derived from tests of significance of the interaction terms added to the model. Black boxes correspond to the reference group (lowest third of LDL-C, lowest third of WHR); whilst blue and red boxes indicate other groups (middle and top thirds of LDL-C and WHR).

Figure 8.7: Association of tobacco use with MI risk by socio-economic status and diet



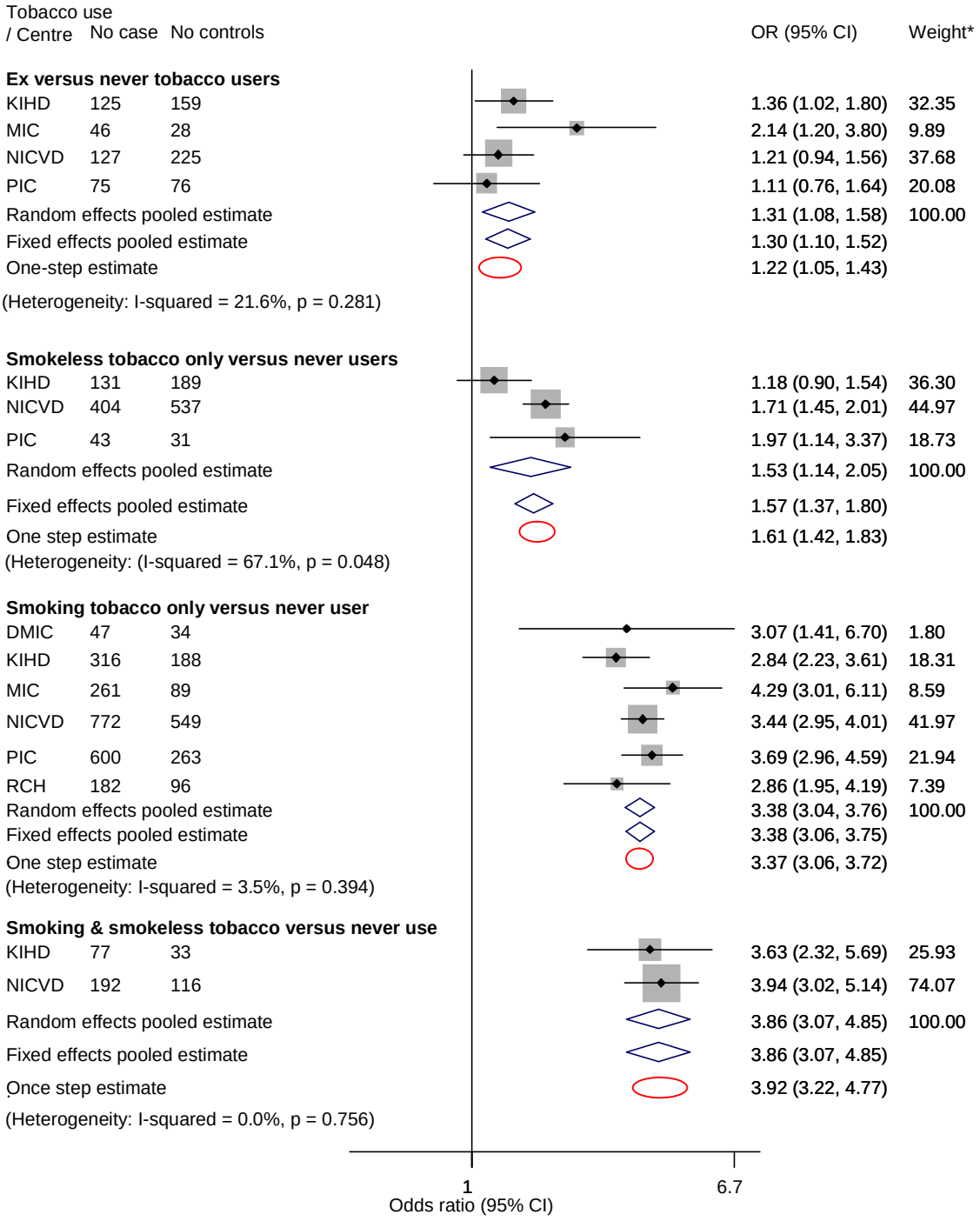
OR: Odds ratios; 95% CI: 95% confidence interval, P-value: p-value of interaction. Confidence intervals were plotted using “floating absolute” variances for tobacco use. Models were adjusted for conventional risk factors: age, sex, ethnicity, centre of recruitment, LDL-C, WHR, history of diabetes or hypertension and family history of MI. P-values of interaction are derived from tests of significance of the interaction terms added to the model. To investigate diet, dietary patterns were created using principal component analysis of 43 food items from a food frequency questionnaire. The two principal components were chosen after orthogonal rotation of the matrix of components and labelled respectively “Carbohydrate and vegetable diet” and “protein and sweet diet”, because of their high loadings for questions rich in protein, sweet, carbohydrate or vegetable. These patterns were then dichotomised at the median. Black boxes correspond to the reference group (lowest third of socio-economic status, below median consumption of a carbohydrate and vegetable diet, below median of a protein and meat diet); whilst blue and red boxes indicate other groups.

Figure 8.8: Association of tobacco use with MI risk by recruitment centre and by source of controls



OPD: Outpatients Department. Models adjusted for age, sex, ethnicity, LDL-C, Waist to hip ratio, history of diabetes or hypertension and family history of MI. ORswere computed within centres and for each type of controls versus all cases of the centre separately, when at least 100 controls for a specific type were available. DMIC: Deewan Mushtaq Institute of Cardiology; FIC: Faisalabad Institute of Cardiology; KIHD: Karachi Institute of Heart Diseases; MIC: Multan Institute of Heart Diseases; NICVD: National Institute of Cardiovascular Diseases; PIC: Punjab Institute of Cardiovascular Disease; RCH: Red Crescent Hospital; MI: Myocardial infarction.

Figure 8.9: Association of tobacco use with MI risk by recruitment centre with fixed and random effect meta-analyses



*Weights are from random effects meta-analysis. Models adjusted for age, sex, ethnicity, LDL-C, Waist to hip ratio, history of diabetes or hypertension and family history of MI. ORswere computed within centres, when at least 100 controls for a specific type were available. Random effects pooled estimate was obtained using an estimate of heterogeneity being taken from the inverse-variance fixed-effect model. The one-step estimate corresponds to ORs adjusting for conventional risk factors and centre of recruitment.

References

1. Mackay J, Mensah G, World Health Organization. *The Atlas of heart disease and stroke*. 2004.

2. Nishtar S. Prevention of coronary heart disease in south Asia. *Lancet.* 2002;360:1015-1018.

3. World Bank. Life expectancy at birth in Pakistan. 28-7-2011.

4. U.S.Department of State. Background Note: Pakistan. 6-10-2011.

5. Anand SS, Yusuf S, Vuksan V et al. Differences in risk factors, atherosclerosis, and cardiovascular disease between ethnic groups in Canada: the Study of Health Assessment and Risk in Ethnic groups (SHARE). *Lancet.* 2000;356:279-284.

6. Balarajan R. Ethnic differences in mortality from ischaemic heart disease and cerebrovascular disease in England and Wales. *BMJ.* 1991;302:560-564.

7. Forouhi NG, Sattar N, Tillin T, McKeigue PM, Chaturvedi N. Do known risk factors explain the higher coronary heart disease mortality in South Asian compared with European men? Prospective follow-up of the Southall and Brent studies, UK. *Diabetologia.* 2006;49:2580-2588.

8. Joshi P, Islam S, Pais P et al. Risk factors for early myocardial infarction in South Asians compared with individuals in other countries. *JAMA.* 2007;297:286-294.

9. Bhopal R. Epidemic of cardiovascular disease in South Asians. *BMJ.* 2002;324:625-626.

10. Ghaffar A, Reddy KS, Singhi M. Burden of non-communicable diseases in South Asia. *BMJ.* 2004;328:807-810.

11. Crofton J. The gathering smoke clouds: a worldwide challenge. *Int J Epidemiol.* 1984;13:269-270.

12. Guindon EG, Boisclair D. Past, current and future trends in tobacco use. *Economics of Tobacco Control.* 2003;1-45.

13. Maziak W. The global epidemic of waterpipe smoking. *Addict Behav.* 2011;36:1-5.

14. Mazahir S, Malik R, Maqsood M et al. Socio-demographic correlates of betel, areca and smokeless tobacco use as a high risk behavior for head and neck cancers in a squatter settlement of Karachi, Pakistan. *Subst Abuse Treat Prev Policy.* 2006;1:10.

15. Ismail J, Jafar TH, Jafary FH, White F, Faruqui AM, Chaturvedi N. Risk factors for non-fatal myocardial infarction in young South Asian adults. *Heart.* 2004;90:259-263.

16. Rahman MA, Zaman MM. Smoking and smokeless tobacco consumption: possible risk factors for coronary heart disease among young patients attending a tertiary care cardiac hospital in Bangladesh. *Public Health.* 2008;122:1331-1338.

17. Pais P, Fay MP, Yusuf S. Increased risk of acute myocardial infarction associated with beedi and cigarette smoking in Indians: final report on tobacco risks from a case-control study. *Indian Heart J.* 2001;53:731-735.

18. Pais P, Pogue J, Gerstein H et al. Risk factors for acute myocardial infarction in Indians: a case-control study. *Lancet.* 1996;348:358-363.

19. Forouhi NG, Sattar N, Tillin T, McKeigue PM, Chaturvedi N. Do known risk factors explain the higher coronary heart disease mortality in South Asian compared with European men? Prospective follow-up of the Southall and Brent studies, UK. *Diabetologia.* 2006;49:2580-2588.

20. Teo KK, Ounpuu S, Hawken S et al. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet.* 2006;19;368:647-658.

21. Iqbal R, Anand S, Ounpuu S et al. Dietary patterns and the risk of acute myocardial infarction in 52 countries: results of the INTERHEART study. *Circulation.* 2008;118:1929-1937.

22. Rosengren A, Subramanian SV, Islam S et al. Education and risk for acute myocardial infarction in 52 high, middle and low-income countries: INTERHEART case-control study. *Heart.* 2009;95:2014-2022.

23. Boffetta P, Straif K. Use of smokeless tobacco and risk of myocardial infarction and stroke: systematic review with meta-analysis. *BMJ.* 2009;339.

24. Yatsuya H, Folsom AR. Risk of incident cardiovascular disease among users of smokeless tobacco in the Atherosclerosis Risk in Communities (ARIC) study. *Am J Epidemiol.* 2010;172:600-605.

25. Boffetta P, Hecht S, Gray N, Gupta P, Straif K. Smokeless tobacco and cancer. *Lancet Oncol.* 2008;9:667-675.

26. World Health Organization. *Tobacco: deadly in any form or disguise*. 2006.

27. Singh RB, Niaz MA, Ghosh S et al. Association of trans fatty acids (vegetable ghee) and clarified butter (Indian ghee) intake with higher risk of coronary artery disease in rural and urban populations with low fat consumption. *Int J Cardiol.* 1996;56:289-298.

28. Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med.* 1991;10:1025-1035.

29. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21:1539-1558.

30. The World Bank. The World Bank: Pakistan. http://data.worldbank.org/country/pakistan accessed 04/07/2012 . 2012.

31. The World Bank. Regional Data: South Asia. 2010.

32. Alam AY, Iqbal A, Mohamud KB, Laporte RE, Ahmed A, Nishtar S. Investigating socio-economic-demographic determinants of tobacco use in Rawalpindi, Pakistan. *BMC Public Health.* 2008;8:50.:50.

33. Nasir K, Rehan N. Epidemiology of cigarette smoking in Pakistan. *Addiction.* 2001;96:1847-1854.

34. World Health Organization. WHO report on the global tobacco epidemic; the MPOWER package. 2008.

35. Dobe M, Sinha DN, Rahman K. Smokeless tobacco use and its implications in WHO South East Asia Region. *Indian J Public Health.* 2006;50:70-75.

36. Dhirendra SN. *Report on oral tobacco use and its implications in South East Asia*. Patna, India: 2004.

37. Teo KK, Ounpuu S, Hawken S et al. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet.* 2006;%19;368:647-658.

38. Khawaja MR, Mazahir S, Majeed A et al. Chewing of betel, areca and tobacco: perceptions and knowledge regarding their role in head and neck cancers in an urban squatter settlement in Pakistan. *Asian Pac J Cancer Prev.* 2006;7:95-100.

39. Gupta PC, Ray CS. Smokeless tobacco and health in India and South Asia. *Respirology.* 2003;8:419-431.

40. Foulds J, Ramstrom L, Burke M, Fagerstrom K. Effect of smokeless tobacco (snus) on smoking and public health in Sweden. *Tob Control.* 2003;12:349-359.

41. Hansson J, Galanti MR, Hergens MP et al. Use of snus and acute myocardial infarction: pooled analysis of eight prospective observational studies. *Eur J Epidemiol.* 2012.

42. Tomar SL. Is use of smokeless tobacco a risk factor for cigarette smoking? The U.S. experience. *Nicotine Tob Res.* 2003;5:561-569.

43. Bolinder GM, Ahlborg BO, Lindell JH. Use of smokeless tobacco: blood pressure elevation and other health hazards found in a large-scale population survey. *J Intern Med.* 1992;232:327-334.

44. Oswal KC, Raute LJ, Pednekar MS, Gupta PC. Are current tobacco pictorial warnings in India effective? *Asian Pac J Cancer Prev.* 2011;12:121-124.

45. Government of Pakistan and Ministry of Health. Statutory Notifications containing Rules and Orders issued by all Ministries and Divisions of the Government of Pakistan and their Attached and Subordinate Offices and the Supreme Court of Pakistan. M-302. 23-10-2009.

46. IRIN. Pakistan: anti-smoking law have little impact. *IRIN Humanitarian news and analysis.* 2004.

Chapter 9: Discussion

Summary

Tobacco use remains the main preventable cause of death and accounts for 1 in 10 CVD events. Worldwide, smoking prevalence is highest in developed countries at around one fourth of the population, but, when considering only men, developing countries are at the top of the list in terms of both smoking and smokeless tobacco rates. In the ERFC, prevalence of tobacco use was higher than current estimates in the developed world and there was evidence of a gap between men and women. In PROMIS controls, tobacco prevalence was relatively low compared to previous reports. All forms of tobacco use investigated in this thesis, namely cigarettes, pipes, cigars, and South Asian chewing and dipping tobacco, were associated with a relatively strong increase in CVD risk. These findings lend more weight to public health campaigns which address all forms of tobacco rather than cigarette smoking alone, and limit the scope for safe alternatives to cigarette smoking. Future work includes harvesting genetic information already available in PROMIS to investigate genetic determinants of tobacco dependence and genetic susceptibility to CVD amongst tobacco users. In the longer term, additional studies on smoking and smokeless tobacco using a prospective design would help strengthen the evidence of a causal relationship between all forms of tobacco use and CVD and help disentangle the source of the additional CVD risk experienced by tobacco users.

9.1 Introduction

CVDs account for 30% of all deaths and 20% of global Disability Adjusted Life Years (DALYs) in individuals aged 30 years and older, being the leading single cause of death and disability worldwide; and this dominance is set to increase as populations are ageing worldwide [1]. Tobacco use remains the main preventable cause of CVD and death. By 2030, if current trends continue, smoking will be killing 9 million people annually [2]. Most tobacco users now live in developing countries and South Asia accounts for a large proportion of them. India itself is thought to contain more than 10% of worldwide smokers [2]. At the recent high-level meeting of the United Nations on non-communicable diseases, in September 2011, delegates asserted that "the increasing global crisis in non-communicable diseases is a barrier to development goals including poverty reduction, health equity, economic stability, and human security … The most urgent and immediate priority is tobacco control." [3]

It has been estimated that reducing death rates from chronic diseases (which include CVD, cancers, chronic respiratory diseases and diabetes) by 2% from 2006 to 2015 would avert 36 million deaths, of which 28 million lives would be saved in low-income and middle-income countries [4]. Almost half of these averted deaths would be in men and women younger than 70 years old. The experience of high income countries shows what can be achieved with sustained interventions, in particular sensitizations to the dangers of smoking. Over the past three decades, the chronic diseases death rate has been reduced by between 1% and 3% every year in the developed world, translating into a fall of up to 70% in the death rate of CHD in some countries (Australia, Canada, Japan, the UK, and the USA). During this period, the number of CVD deaths averted has been estimated to be 14 million in the USA, 8 million in Japan and 3 million in the UK. Reasons for these decreases remain partially understood but have been largely attributed to the halving in smoking prevalence which happened during the same period in these countries.

In this thesis, I have investigated in more detail than previously possible the relationship between tobacco use and the risk of CVD. In particular, I have strengthened the evidence concerning pipe and cigar smoking in relation to MI and stroke using a prospective design, and showed a relatively strong association between South Asian chewing and dipping tobacco with MI risk using a case-control design. I have also investigated in more detail than previous large scale studies or meta-analyses have been able to regarding the relationship between cigarette

smoking and CVD risk. In this **Discussion**, I review my main findings, their public health relevance and envision future work.

9.2 Tobacco prevalence in developed and developing countries

9.2.1 Smoking tobacco products

The developed world has been successfully implementing strategies to reduce smoking since the 1960's, and this reflects in prospective studies included in the ERFC. Early studies started in the 1960's and 1970's reported a higher proportion of baseline current smokers, especially amongst men, than later studies which began enrolment in the 1990's and 2000's and had prevalence compatible with estimates of 20-25% in both men and women. In developing countries, and particularly in South Asia, the picture is vastly different. Firstly, the gap between men and women in smoking prevalence remains present with a much higher prevalence in men. Secondly, the prevalence of smoking is still high and rising with time. The share of cigarette production and consumption of developing was 40% of worldwide production in 1970, whilst it is over 70% nowadays [2]. In Pakistan, the production and consumption of cigarettes has doubled since 1970 [5]. Thirdly, cigarettes coexist with other popular smoking and smokeless products. In the PROMIS controls, nearly a quarter of controls smoked, and more than 1 in 10 used smokeless products. Smoking was associated with the use of *ghee*, a traditional cooking medium which is cheaper than oil and has been shown to be associated with increased risk of CVD, demonstrating a clustering of harmful habits in tobacco users [6]. All forms of tobacco use were inversely associated with socio-economic status, indicating that it is mainly the poorest section of the population that is attempting to emulate the bad habits of the West.

9.2.2 Smokeless tobacco products

The use of traditional smokeless tobacco products appears stable in the USA at around 3% of the population, representing 8.1 million individuals [7, 8]. Smokeless tobacco had fallen out of fashion in the Western world during the 1st half of the 20th century but there has been a resurgence since the 1970's as smokeless tobacco was rebranded as a new, cheap and trendy alternative to cigarettes [9]. An example includes dipping snuff which does not require spitting: It is called *snus* and is already popular in Northern Europe, especially Sweden and Norway. It has variants in the US with higher nitrosamine and carcinogenic content than the Swedish product; while it remains forbidden in other countries of the European Union [10]. With globalization, there is a risk of developed countries importing products already

popular in developing countries. Migrations of population from developing to developed countries, the internet, and tourism make these products more accessible and more visible [11]. As *snus* has becomes the dominant form of tobacco used by Swedish men [12], this popularity may serve to promote other dipping products such as the South Asian version called *naswar*. By contrast, betel quid chewing and dipping have been part of local South Asian cultures for centuries, even millennia, and tobacco has been added to chewing and dipping mixtures in the 17th and 18th century at the time of colonization. Smokeless tobacco was traditionally hand prepared, but over the past half century, industrialized pouches such as *gutka* in Pakistan have been developed and have been so successful that they have overridden traditional products (such as *paan* and *supari*). In India, there is evidence that per capita smokeless tobacco consumption has increased among the poor since the 1960s in both rural and urban areas, and the total number of smokeless tobacco users has been estimated at around 100 million people [13, 14]. In the PROMIS controls, smokeless tobacco use was correlated with traditional Pakistani diet high in carbohydrates and vegetables and was especially favoured by women. Dipping tobacco, called *naswar*, was mainly favoured by older individuals and lower socio-economic groups.

9.3 Association of tobacco with cardiovascular risk

9.3.1 Smoking tobacco use and CVD risk

Cigarette smoking caused an increased risk of CVD even at low intensities. The risk doubled from 0 to 5 cigarettes per day; and, to double again, a smoker had to increase their intake from 5 to more than 25 cigarettes per day. Inversely, stopping smoking was rapidly associated with a reduced risk of CVD, and ex-smokers had CVD risk generally below 15%. This would indicate that most of the effect of smoking on CVD is relatively short-term and reversible, promoting plaque rupture and thrombosis. In that respect, cigarette smoking has been shown to increase myocardial workload, reduce oxygen-carrying capacity of the blood, cause coronary vasoconstriction, increase catecholamine release and induce a hypercoagulable state [9]. However, risk for ex-smokers did not go back to that of never smokers until 20 years after cessation, which means that some components of smoking durably affect vessel walls, and are probably involved in the early processes of atherosclerosis and in plaque building by promoting degradation of the vessel wall and lipid uptake by macrophages.

Smoking cigarettes was associated with a doubling in risk of MI in the ERFC and a tripling in risk of MI in the PROMIS participants. It has been suggested that South Asians are more susceptible to traditional CVD risk factors and that this could account for the fact that they experience CVD on average 10 years earlier than Western individuals [18]. However, the difference I observed in risk ratios for smoking could also be an artefact resulting from the difference in study designs. Indeed, the ERFC is a collection of prospective studies and allowed estimation of hazard ratios, whilst PROMIS used a case-control design and only allowed estimation of odds ratios, which are prone to over-estimate risk ratios, and are subject to recall bias [19]. Studies included in the ERFC mostly started enrolment before the 1980's, when smoking prevalence was high and therefore rates of passive smoking in never smokers were probably high as well, with few legislations at the time to limit its impact. Smoke-free policies, even if they are becoming more common [2], are still rarely enforced in developing countries. In Pakistan, legislation on tobacco use is not implemented [20] and rates of passive smoking in non-tobacco users are likely to be elevated as a result. Passive smoking has been shown to increase the risk in non-smokers by up to a factor of 3.7 [21]. Therefore, if passive smoking was to exert its effect mainly in never smokers as has been shown previously, ERFC and PROMIS estimates could both be conservative and the true increase in risk caused by smoking and smokeless tobacco use may be even higher.

The association between pipe or cigar smoking with CVD was strong and intermediate between that of never smokers and that of cigarette smokers, in agreement with previous findings [15-17]. RRs for all CVD events were 1.31 (1.19; 1.44) for cigar only smokers, 1.68 (1.56; 1.81) for pipe only smokers; and 1.97 (for cigarette only users). At high intensities (≥15 cigarettes equivalent per day), pipes or cigars experienced close to a doubling in risk: RR was 1.86 (1.54; 2.25). By contrast, cigarette smokers smoking 15-20 cigarettes per day had a RR of 2.50 (2.34; 2.66), and >25 cigarettes per day a RR of 2.93 (2.63; 3.26). Smoke from pipes and cigars contains the same toxic substance as cigarette smoke and individuals who switched from cigarettes to pipes or cigars are more likely to inhale tobacco than never cigarettes users [15]. In this context, the reason that cigarette smoking caused a greater increase in the risk of CVD compared to pipes and cigars at similar intensities remains unclear and would need to be further investigated in studies where full smoking history is available for all smoking types, including doses and duration.

The investigation of interplay between smoking and other risk factors on CVD risk showed that individuals with lower absolute risk, such as younger age groups, women, non-diabetics and non-hypertensives had a higher increase in risk when they smoked than older age groups, men, diabetics and hypertensive individuals. These statistically significant synergies remain difficult to interpret. They could be attributed to a relative absence of competing risks in individuals with lower baseline risk. It is a clear public health message for women, non-diabetics, non-hypertensive and younger age groups that the added impact of smoking may counterbalance their relatively better CVD prospect than groups with higher baseline risk such as men, older individuals, hypertensive and diabetics. Nevertheless, this interaction on a relative scale did not translate into a substantially higher burden of smoking related CVD in these low risk groups compared to the higher risk groups. In absolute terms, more CVD events were still caused by smoking amongst older age groups, men, diabetics and hypertensive. The interactions with body mass index and total cholesterol were non-significant.

9.3.2 Smokeless tobacco use and CVD risk

The effect of South Asian smokeless tobacco was significant and intermediate between that of a never tobacco user and that of a current smoker of cigarettes or *bidis* in PROMIS participants. These estimates were more precise and broadly comparable to the published literature [22]. ORs for myocardial infarction were 1.31 (1.39-1.88) with current chewing of tobacco, and 1.35 (1.12-1.63) with current snuff dipping, compared to never tobacco users. Individuals who chewed or dipped tobacco on top of smoking seemed to be at higher risk than smokers alone (3.91 versus 3.36), even if the difference was not statistically significant.

My results regarding chewable tobacco were in broad agreement with a worldwide case-control study [22] as well as with recent meta-analyses of the effect of smokeless tobacco in Western population where most of the evidence came from a Northern American study on chewable rather than snuff tobacco [23]. Regarding tobacco dipping, my results differ from a recent meta-analysis selecting only Swedish data which found a non-significant increase in MI risk for snuff users [24] and has led some to argue that Swedish snuff called *snus* is an acceptable alternative to smoking [25, 26]. However, Pakistani snuff (*naswar)* and Swedish *snus* have different processes of fabrication and Swedish *snus* is especially low in nitrosamines [26]. Therefore caution should remain when generalizing results from Swedish studies to other populations and promoting liberalization of snuff dipping outside of Northern Europe.

As pointed out by Last in his Dictionary of Epidemiology, the presence of a dose-response relationship is another argument in favour of a causal relationship between smokeless tobacco and MI [27]. There was evidence in the PROMIS data of a dose-response relationship below 10 smokeless products a day. Individuals consuming 5-10 doses a day had an OR of 1.77 (1.43; 2.20) risk of MI, which is a considerable increase in risk. The increased risk of a similar amount of cigarettes was 2.56 (2.07; 3.16).

These findings may help better understand the aetiology of tobacco use in relation to CVD and suggest a role for toxins that are intrinsic to tobacco itself. The pathogenesis of CVD as a result of smokeless tobacco use is not well understood. The proportion of nicotine is very variable and chemical composition differs between products. Unlike smoking, which produces rapid peaks and troughs, smokeless tobacco use causes more prolonged, sustained levels of nicotine, often lasting for 1 hour [9]. The rate of absorption may vary among different forms of smokeless tobacco depending on the pH level of the product, amount of nicotine and size of the tobacco cutting. When testing a range of smokeless tobacco products, venous concentrations have been shown to be higher than for cigarette smoking [28]. In addition to nicotine, smokeless tobacco has been shown to contain nitrosamines, nitrosamine acids, polycyclic aromatic hydrocarbons, aldehydes and metals, whose effects are known to be cancerous and remain unknown relative to CVD [29]. Blood pressure levels are thought to be affected by the high sodium content of smokeless tobacco as well as by nicotine, and acute cardiovascular effects, similar to those caused by cigarette smoking, are seen with the use of smokeless tobacco such as an increase in heart rate and blood pressure levels which is strong enough to activate the sympathetic nervous system [9]. In particular, chewing betel quid has been shown to increase blood pressure and the chance of being hypertensive [29].

9.4 Public health relevance of my findings

Cigarettes account for 96% of global sales of manufactured tobacco by value, and global cigarette production continues to increase dramatically [2, 30]. Around a billion people are addicted to nicotine in the form of cigarettes and many have no immediate plans to quit. The concept of “harm reduction” has been coined to refer to the objective of minimising the net damage to health of continuing tobacco users and the general population by substituting less harmful tobacco products for cigarettes, for example other smoking products such as pipes or cigars or even smokeless

tobacco [15, 30]. I explore below whether, in light of my findings, this concept is applicable to either developed or developing countries.

Smokeless tobacco as a "harm reduction" tool, and *snus* in particular for which evidence regarding CVD risk remains inconclusive, has been advocated in developed countries [31]. In Sweden, consumption of *snus* is thought to play a role in the relatively low prevalence of smoking and low rates of MI compared to other European countries [12]. However, extending the experience to other developed countries remains controversial [32]. Another argument against the use of *snus*, which is also valid for cigars and pipes, includes the many unintended consequences that promoting *snus* as a substitute to cigarettes could have Individuals that never intended to smoke cigarettes could initiate tobacco use.. Some individuals could only partially rather than entirely replace cigarette smoking with use of another tobacco product. In the ERFC, individuals who smoked pipes or cigars on top of cigarettes had a RR for CVD of 1.95 (1.81; 2.10), nearly equivalent to that of cigarette only users who experienced a RR of 1.97 (1.85; 2.11). Individuals who initiate use of cigars, pipes or smokeless tobacco may also be tempted to later switch to cigarettes. *Snus* and cigarettes are sold by the same companies [33] and cigars are increasingly made to look like cigarettes in shape and format [17]. In the US, young men who were not smokers but regularly used smokeless tobacco were more than three times as likely as never users to have become current smokers four years later [34]. Cigar and pipe smoking, which are very similar forms of tobacco use compared to cigarette smoking, could have even higher switch rates. In this context, the steady growth of cigar[35] and smokeless tobacco use[9] over the past decades in the US is a source of concern, and I believe that the best public health policy in developed countries remains to encourage quitting all forms of tobacco use. One example which I was not able to investigate in my thesis because of the lack of data available, and soon to compete in prevalence with pipes and cigars in developed countries, is the *e-cigarette*. *E-cigarettes* are the fastest growing smokeless tobacco industry in developed countries: the industry is already worth £150m in the UK and £520m in the US [36]. The *e-cigarette* is an electronic nicotine delivery device [37] promoted as a safe alternative to cigarette smoking which produces no passive smoking [38]. It is not legislated in EU countries and is sometimes used on pharmacy counters as a smoking cessation product. Evidence regarding the effect of *e-cigarettes* on health is unknown and it is questionable whether it should belong in the category of nicotine replacement products or is a novel form of smokeless tobacco.

In developing countries, and in particular in South Asia, smoking is not yet the dominant culturally accepted form of tobacco use. In my analysis of urban Pakistanis, smoking cigarettes coexisted with bidies smoking and a large panel of smokeless products (*paan*, *gutka*, *supari* and *naswar*). Smoking was preferred by men whilst smokeless tobacco was particularly favoured by women and younger individuals. All these forms of tobacco use were significantly associated with the risk of MI. Current users of both smoking and smokeless products experienced a risk ratio for non-fatal MI higher than smokers alone; and their risk was nearly four times the risk of a never tobacco user. In this context, applying different public health policies to smoking and smokeless tobacco and encouraging smokers to switch to smokeless products as advocated by the "harm reduction" approach would have several harmful effects which may counter any beneficial impact [39]. Firstly, in developing countries, tobacco control is in its infancy [24], and this policy risks damaging educational effort teaching that tobacco is harmful to cardiovascular health. In the PROMIS participants, only 9% of men and 2% of women were past tobacco users, indicating that most smokers have probably never attempted to quit. In this context, attempting to quit, rather than switching to other tobacco products, needs to remain the first step for someone willing to curtail his or her CVD risk. Secondly, it would leave out women who are already using smokeless tobacco. Thirdly, there is evidence that smokeless tobacco is a gateway to smoking. In South Asia, smokeless products are particularly favoured by children who later on will start smoking in teenagehood and adulthood [14]. Fourthly, smokeless products in developing countries have an unknown toxicity. In my study, in contrast to findings regarding *snus* use in developed countries, snuff dipping in the form of *naswar* was significantly associated with MI risk. In the absence of regulatory control that can successfully address the toxicities introduced by the practices of small local producers, any shift from smoking to smokeless tobacco use is likely to be to the indigenous forms of high toxicity rather than to the Swedish-style products manufactured to have low toxicity [39].

For all the reasons listed above, the findings of this thesis encourage a public health policy addressing all forms of tobacco use, rather than cigarettes alone. The Framework Convention on Tobacco Control (FCTC) was adopted in 2003 by 170 countries members of WHO [40]. It advocates a world free of all forms of tobacco use with a target prevalence <5% worldwide. I have shown that CVD is the cause of around 1/3 of smoking deaths in middle-age Western individuals. In the ERFC, the CVD risk of cigarette smoking was rapidly reversible in past smokers, with an 80% reduction in risk over the first 5 years while the decrease was much slower for the

risk of lung cancer. It is plausible that stopping using smokeless tobacco will also rapidly result in a decreased risk. Among PROMIS participants, past users of all tobacco products had a non-significant increased risk of MI. In this context, WHO predicts that full implementation of the FCTC would avert 5.5 million deaths over 10 years in 23 low-income and middle-income countries with a high burden of non-communicable diseases, including Pakistan [3]. Most of the deaths averted would be because of a reduction in the burden of CVD rather than cancer or other non-communicable diseases [3].

9.5 Future work

This thesis contains several shortcomings which could be addressed either using the dataset already available, or in the longer term by creating new appropriate resources.

Firstly, case-control studies are unable to establish causality, and my estimates are subject to recall bias [27]. To prove causality of smokeless tobacco on CVD risk beyond doubt, a randomized control trial would be needed, but would prove ethically challenging, given the evidence established in this thesis. A prospective study with a long follow-up and set in a developing country such as Pakistan where the prevalence of these practices is relatively high (>1 in 10 men and women used smokeless tobacco in PROMIS controls) would provide invaluable information on the relationship between smokeless tobacco and cardiovascular outcomes. Information on cotinine levels, a biomarker of nicotine ingested and therefore of smoking intensity, and tobacco and other contents of the smokeless product would also be useful as they would enable the analyst to address the question of dose-response relationship and help unravel aetiological mechanisms by which smokeless tobacco leads to CVD, in particular whether nicotine plays a role in inducing CVD. It could also help inform strategies regarding other nicotine products such as e-cigarettes.

Secondly, genetic information from a genome wide association study (GWAS) done using Illumina 660 Quad array is available in PROMIS participants. This could allow me (1) to investigate genetic determinants of smoking and smokeless tobacco use, and (2) to test whether some tobacco users have a genetic predisposition to CVD. Before the development of high-throughput genotyping platforms, candidate gene studies investigated whether targeted genes, selected because of their biological relevance to the atherosclerotic process, had an effect on CVD and whether this effect depended on smoking status. Candidate gene studies have generally been

small, underpowered and poorly replicated. Genes found to interact with smoking status on the risk of CVD include the C-allele of CYP1A1, an enzyme present in the lungs that is known to activate smoke carcinogens [41]; eNOS4a, an uncommon polymorphism of the endothelial enzyme eNOS involved in the synthesis of nitric oxide [42]; the p53 protein, a transcription factor that suppresses growth and triggers apoptosis [43]; the glycoprotein IIIa P1(A2) polymorphism [44]; the factor V Arg506 Gln mutation known to reduce the anticoagulation effect of activated protein C; glutathione S-transferases M1 and T1 [45]; PON1192Arg polymorphism in the paroxanase gene; and the L-gene promoter polymorphism of the serotonin transporter gene. More recently, GWASs have identified single nucleotide polymorphisms (SNPs) associated with addiction, initiation and cessation of smoking (**Table 9.1**). The most highly associated locus with number of cigarettes per day has been rs1051730, located on chromosome 15q25 and the overlapping three genes CHRNA5/CHRNA3/CHRNA4 known to encode neuronal nicotinic acetylcholine receptor subunits [46-49]. This locus has also been associated with lung cancer and peripheral vascular diseases but not yet with CVD. Complementary to GWASs of DNA variation, large scale epigenetic studies have identified a locus cg03636183 located in F2RL3 with altered methylation patterns in current versus never smokers [50]. This locus happens to be lying in a gene coding for a potential drug target of cardiovascular importance. An analysis of PROMIS participants, with replication of my findings in other populations to enhance power, could yield novel genetic markers of smoking and smokeless tobacco dependence and help us understand how tobacco use causes CVD in the context of a South Asian population.

Thirdly, a source of passive smoking other than “being in the presence of a smoker” has been recently introduced as “third-hand smoking” and relates to bio-persistent cigarette smoke residue [51]. Nicotine has been shown to persist in ceiling tiles for up to 30 years [52]. Particulates from cigarette smoke deposit on indoor surfaces and can have toxic effect: for example nicotine residue can react with ambient nitrous acid to produce carcinogenic compounds. The effect of third-hand smoking on health and, in particular, on the risk of CVD remains unknown. It is also unknown whether other forms of tobacco use, such as chewing and dipping also produce residue (for example on the table where they are prepared, often by children) which persist for a long time and whether these would have an effect on health. Investigating the effect of third-hand smoking would require first defining more precisely the exposure of interest and second devising reliable methods to measure it on a large scale before being able to assess its effect on population health.

9.7 Conclusion

Tobacco use remains a danger to cardiovascular health. Cigarette smoking is the most popular form of tobacco in developed countries whilst it coexists with a range of smokeless products in a developing country such as Pakistan. Cigarette or bidi smoking carried the greatest risk, whereas pipes, cigars, chewable products and dipping snuff were all associated with a substantially increased risk of MI. Enforcing strict regulations on both smoking and smokeless tobacco use and encouraging individuals to quit any type of tobacco use remain preferable, compared to a policy which advocates replacement of cigarettes with other tobacco products with weak associations with CVD risk.

Table 9.1: Genome-wide association studies reporting on the association with number of cigarettes per day and other smoking phenotypes

Genetic information							Effect size					Reference		
Chr	SNP	Position (Mb)	Coded allele	Non coded allele	Gene	Smoking phenotype	Effect size	se	P-value	Study population	Sample size	Journal	Year	First author
1	rs910696	30,236,689	*MA*		intergenic	CPD	ABS available	NA	3.00E-06	CGEMS (PLCO & NHS)	4.5K	PlosOne	2009	Caporaso
3	rs6437740	108,948,507	*MA*		BBX	CPD	by study	NA	2.40E-07	CGEMS (PLCO & NHS)	4.5K	PlosOne	2009	Caporaso
4	rs5522	149576925			NR3C2	CPD			1.52E-05		15K	Molecular Psychiatry	2008	Berrettini
4	rs5525	149576925			NR3C2	CPD			3.78E-05		15K	Molecular Psychiatry	2008	Berrettini
5	rs2645339	178348669			GRM6	CPD			0.000272		15K	Molecular Psychiatry	2008	Berrettini
7	rs215605	32,303,490	G	T	intergenic	CPD	0.26	0.04	5.4x10-9	Ox-GSK, TAG, ENGAGE	85K	Nature Genetics	2010	Thorgeirsson
7	rs215614	32,313,860	G	A	intergenic	CPD	0.22	0.04	2.1x10-7	Ox-GSK, TAG, ENGAGE	85K	Nature Genetics	2010	Thorgeirsson
7	rs7804771	136783133			CGKI	CPD			9.81E-05		15K	Molecular Psychiatry	2008	Berrettini
8	rs13280604	42,678,743	A	G		CPD	0.31	0.05	1.3x10-8	Ox-GSK, TAG, ENGAGE	85K	Nature Genetics	2010	Thorgeirsson
8	rs6474412	42,669,655	T	C	CHRNB3\CHRNA6	CPD	0.29	0.05	1.4x10-8	Ox-GSK, TAG, ENGAGE	85K	Nature Genetics	2010	Thorgeirsson
9	rs10869409	149576925			RORB	CPD			5.53E-05		15K	Molecular Psychiatry	2008	Berrettini
9	rs13293006	76326716			RORB	CPD			0.000592		15K	Molecular Psychiatry	2008	Berrettini
9	rs7846903	149576925			RORB	CPD			0.000122		15K	Molecular Psychiatry	2008	Berrettini
9	rs7873840	76340109			RORB	CPD			0.001698		15K	Molecular Psychiatry	2008	Berrettini
10	rs1028936	93,349,297	C	A	EGLN2	CPD	−0.4464	0.074	1.29 × 10−9	Ox-GSK, TAG, ENGAGE	74K	Nature Genetics	2015	TAGC
10	rs1329650	93,347,620	T	G	EGLN2	CPD	−0.3673	0.059	5.67 × 10−10	Ox-GSK, TAG, ENGAGE	74K	Nature Genetics	2014	TAGC
15	rs1051730	76,681,394	A	G	CHRNA5/CHRNA3/CHRNB5	CPD	0.8	0.05	2.4x10-69	Ox-GSK, TAG, ENGAGE	85K	Nature Genetics	2010	Thorgeirsson
15	rs1051730	76,681,394	G		CHRNA5/CHRNA3/CHRNB5	CPD			1.71x10-66	Ox-GSK, TAG, ENGAGE	41K	Nature Genetics	2010	Liu
15	rs1051730	76,681,394	T	G	CHRNA5/CHRNA3/ CHRNB5	CPD	0.095		6x10-20	deCODE, Zaragoza, the Netherlands	16K	Nature	2008	Thorgeirsson
15	rs1051730		G	A	CHRNA5/CHRNA3/CHRNB5	CPD	−1.0209	0.056	2.8x10-73	Ox-GSK, TAG,ENGAGE	74K	Nature Genetics	2010	TAGC
15	rs12439738	90336555			SLCO3A1	CPD			0.000531		15K	Molecular Psychiatry	2008	Berrettini
15	rs12439765	90336606			SLCO3A1	CPD			0.000625		15K	Molecular Psychiatry	2008	Berrettini
15	rs16969968	76,669,980	G		CHRNA5/CHRNA3/CHRNB5	CPD			4.29x10*65	Ox-GSK, TAG, ENGAGE	41K	Nature Genetics	2010	Liu
15	rs16969968		G	A	CHRNA5/CHRNA3/CHRNB5	CPD	−1.0029	0.056	5.57 × 10−72	Ox-GSK, TAG, ENGAGE	74K	Nature Genetics	2011	TAGC
15	rs4932597	90338621			SLCO3A1	CPD			0.000245		15K	Molecular Psychiatry	2008	Berrettini
15	rs4932598	90338849			SLCO3A1	CPD			0.000192		15K	Molecular Psychiatry	2008	Berrettini
15	rs55853698	76,652,480	T		CHRNA5/CHRNA3/ CHRNB5	CPD			1.74x10-3	Ox-GSK, TAG, ENGAGE	41K	Nature Genetics	2010	Liu
15	rs6495308	76694711			CHRNA5/CHRNA3/CHRNB5	CPD			6.9x10-5		15K	Molecular Psychiatry	2008	Berrettini
15	rs6495308	76,694,711	T		CHRNA5/CHRNA3/CHRNB5	CPD			5.82x10-44	Ox-GSK, TAG, ENGAGE	41K	Nature Genetics	2010	Liu
16	rs7260329	46,213,478	G	A	CYP2B6	CPD	0.2	0.04	5.5x10-6	Ox-GSK, TAG, ENGAGE	85K	Nature Genetics	2010	Thorgeirsson
17	rs758642	3,733,656	*MA*		CAMKK1	CPD	ABS	NA	7.30E-06	CGEMS (PLCO & NHS)	4.5K	PlosOne	2009	Caporaso
19	rs10411195	19,897,176	*MA*		ZNF505	CPD	ABS	NA	5.80E-06	CGEMS (PLCO & NHS)	4.5K	PlosOne	2009	Caporaso
19	rs1801272	46,046,373	A	T		CPD	0.68	0.18	1.1x10+4	Ox-GSK, TAG, ENGAGE	85K	Nature Genetics	2010	Thorgeirsson
19	rs4105144	46,050,464	C	T	CYP2A6	CPD	0.39	0.06	2.2x10-12	Ox-GSK, TAG, ENGAGE	85K	Nature Genetics	2010	Thorgeirsson
19	rs7937	45,994,546	T	C	RAB4B	CPD	0.24	0.04	2.4x10-9	Ox-GSK, TAG, ENGAGE	85K	Nature Genetics	2010	Thorgeirsson
X	rs7050529	110,961,378	MA		TRPC5	CPD	ABS	NA	6.20E-06	CGEMS (PLCO & NHS)	4.5K	PlosOne	2009	Caporaso

MA: Minor allele; ABS: available by study. Legend: Chr: Chromosome; SNP: Single Nucleotide Polymorphism; CPD: Number of Cigarettes smoked per day; SI: Smoking initiation. Literature search conducted on PubMed in January 2011. (search: ("smoking" OR "tobacco" OR "cigarettes") AND ("genome wide" OR "whole genome" OR "GWA" OR "SNP" OR "single nucleotide polymorphism") as well as retrieving papers through manual inspections of citation lists.

References

(1) Strong K, Mathers C, Leeder S, Beaglehole R. Preventing chronic diseases: how many lives can we save? *Lancet* 2005 October 29;366(9496):1578-82.

(2) World Health Organization. WHO report on the global tobacco epidemic. the MPOWER package. 2008.

(3) Beaglehole R, Bonita R, Horton R et al. Priority actions for the non-communicable disease crisis. *Lancet* 2011 April 23;377(9775):1438-47.

(4) World Health Organization. *The Tobacco Atlas*. 2002.

(5) World Health Organization. Country profiles, Pakistan. 25-8-2011.

(6) Ismail J, Jafar TH, Jafary FH, White F, Faruqui AM, Chaturvedi N. Risk factors for non-fatal myocardial infarction in young South Asian adults. *Heart* 2004 March 1;90(3):259-63.

(7) Saunders C, Geletko K. Adolescent Cigarette Smokers' and Non-cigarette Smokers' Use of Alternative Tobacco Products. *Nicotine Tob Res* 2012 February 7.

(8) Piano MR, Benowitz NL, Fitzgerald GA et al. Impact of smokeless tobacco products on cardiovascular disease: implications for policy, prevention, and treatment: a policy statement from the American Heart Association. *Circulation* 2010 October 12;122(15):1520-44.

(9) Gupta R, Gurm H, Bartholomew JR. Smokeless tobacco and cardiovascular risk. *Arch Intern Med* 2004 September 27;164(17):1845-9.

(10) Loukas A, Batanova MD, Velazquez CE et al. Who uses snus? A study of Texas adolescents. *Nicotine Tob Res* 2012 May;14(5):626-30.

(11) Peeters S, Gilmore AB. How online sales and promotion of snus contravenes current European Union legislation. *Tob Control* 2012 January 21.

(12) Foulds J, Ramstrom L, Burke M, Fagerstrom K. Effect of smokeless tobacco (snus) on smoking and public health in Sweden. *Tob Control* 2003 December;12(4):349-59.

(13) Dobe M, Sinha DN, Rahman K. Smokeless tobacco use and its implications in WHO South East Asia Region. *Indian J Public Health* 2006 April;50(2):70-5.

(14) Gupta PC, Ray CS. Smokeless tobacco and health in India and South Asia. *Respirology* 2003 December;8(4):419-31.

(15) Wald NJ, Watt HC. Prospective study of effect of switching from cigarettes to pipes or cigars on mortality from three smoking related diseases. *BMJ* 1997 June 28;314(7098):1860-3.

(16) Tverdal A, Bjartveit K. Health consequences of pipe versus cigarette smoking. *Tob Control* 2011 March;20(2):123-30.

(17) Baker F, Ainsworth SR, Dye JT et al. Health risks associated with cigar smoking. *Stat Med* 2002 June 15;21(11):1539-58.

(18) Forouhi NG, Sattar N, Tillin T, McKeigue PM, Chaturvedi N. Do known risk factors explain the higher coronary heart disease mortality in South Asian compared with European men? Prospective follow-up of the Southall and Brent studies, UK. *Diabetologia* 2006 November;49(11):2580-8.

(19) Kirkwood B, Sterne J. Measure of association and impact. *Medical statistics*. 2003. p. 447-57.

(20) IRIN. Pakistan: anti-smoking law have little impact. *IRIN Humanitarian news and analysis* 2004 July 14.

(21) Whincup PH, Gilg JA, Emberson JR et al. Passive smoking and risk of coronary heart disease and stroke: prospective study with cotinine measurement. *BMJ* 2004 July 24;329(7459):200-5.

(22) Teo KK, Ounpuu S, Hawken S et al. Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet* 2006 August;19;368(9536):647-58.

(23) Boffetta P, Straif K. Use of smokeless tobacco and risk of myocardial infarction and stroke: systematic review with meta-analysis. *BMJ* 2009 August 18;339:

(24) Hansson J, Galanti MR, Hergens MP et al. Use of snus and acute myocardial infarction: pooled analysis of eight prospective observational studies. *Eur J Epidemiol* 2012 June 22.

(25) Nitzkin JL. Re: "Risk of incident cardiovascular disease among users of smokeless tobacco in the Atherosclerosis Risk in Communities (ARIC) Study". *Am J Epidemiol* 2011 March 1;173(5):593-4.

(26) Gartner CE, Hall WD, Vos T, Bertram MY, Wallace AL, Lim SS. Assessment of Swedish snus for tobacco harm reduction: an epidemiological modelling study. *Lancet* 2007 June 16;369(9578):2010-4.

(27) Last J. *A Dictionary of Epidemiology*. 4th edition ed. Oxford University Press; 2001.

(28) Benowitz NL, Porchet H, Sheiner L, Jacob P, III. Nicotine absorption and cardiovascular effects with smokeless tobacco use: comparison with cigarettes and nicotine gum. *Clin Pharmacol Ther* 1988 July;44(1):23-8.

(29) Heck JE, Marcotte EL, Argos M et al. Betel quid chewing in rural Bangladesh: prevalence, predictors and relationship to blood pressure. *Int J Epidemiol* 2012 April;41(2):462-71.

(30) Warner KE. Tobacco harm reduction: promise and perils. *Nicotine Tob Res* 2002;4 Suppl 2:S61-71.:S61-S71.

(31) Foulds J, Kozlowski L. Snus--what should the public-health response be? *Lancet* 2007 June 16;369(9578):1976-8.

(32) Boffetta P, Hecht S, Gray N, Gupta P, Straif K. Smokeless tobacco and cancer. *Lancet Oncol* 2008 July;9(7):667-75.

(33) Gartner CE, Hall WD, Chapman S, Freeman B. Should the health community promote smokeless tobacco (snus) as a harm reduction measure? *PLoS Med* 2007 July;4(7):e185.

(34) Tomar SL. Is use of smokeless tobacco a risk factor for cigarette smoking? The U.S. experience. *Nicotine Tob Res* 2003 August;5(4):561-9.

(35) State-specific prevalence and trends in adult cigarette smoking--United States, 1998-2007. *MMWR Morb Mortal Wkly Rep* 2009 March 13;58(9):221-6.

(36) Burnettes K. Sales of e-cigarettes soar, but do they really work? *The Guardian* 12 A.D. January 10;www.guardian.co.uk.

(37) Cobb NK, Abrams DB. E-cigarette or drug-delivery device? Regulating novel nicotine products. *N Engl J Med* 2011 July 21;365(3):193-5.

(38) Grana RA, Glantz SA, Ling PM. Electronic nicotine delivery systems in the hands of Hollywood. *Tob Control* 2011 June 9.

(39) yo-Yusuf OA, Burns DM. The complexity of 'harm reduction' with smokeless tobacco as an approach to tobacco control in low-income and middle-income countries. *Tob Control* 2012 March;21(2):245-51.

(40) www.who.int/fctc/en. WHO Framework Convention on Tobacco Control. 24-8-2012.

(41) Wang XL, Greco M, Sim AS, Duarte N, Wang J, Wilcken DE. Effect of CYP1A1 MspI polymorphism on cigarette smoking related coronary artery disease and diabetes. *Atherosclerosis* 2002 June;162(2):391-7.

(42) Wang XL, Sim AS, Badenhop RF, McCredie RM, Wilcken DE. A smoking-dependent risk of coronary artery disease associated with a polymorphism of the endothelial nitric oxide synthase gene. *Nat Med* 1996 January;2(1):41-5.

(43) Wang XL, Wang J, Wilcken DE. Interactive effect of the p53 gene and cigarette smoking on coronary artery disease. *Cardiovasc Res* 1997 August;35(2):250-5.

(44) Barakat K, Kennon S, Hitman GA et al. Interaction between smoking and the glycoprotein IIIa P1(A2) polymorphism in non-ST-elevation acute coronary syndromes. *J Am Coll Cardiol* 2001 November 15;38(6):1639-43.

(45) Li R, Boerwinkle E, Olshan AF et al. Glutathione S-transferase genotype as a susceptibility factor in smoking-related coronary heart disease. *Atherosclerosis* 2000 April;149(2):451-62.

(46) Thorgeirsson TE, Gudbjartsson DF, Surakka I et al. Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet* 2010 May;42(5):448-53.

(47) Thorgeirsson TE, Geller F, Sulem P et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* 2008 April 3;452(7187):638-42.

(48) Liu JZ, Tozzi F, Waterworth DM et al. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nat Genet* 2010 May;42(5):436-40.

(49) Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010 May;42(5):441-7.

(50) Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-Smoking-Related Differential DNA Methylation: 27K Discovery and Replication. *Am J Hum Genet* 2011 April 8;88(4):450-7.

(51) Kuschner WG, Reddy S, Mehrotra N, Paintal HS. Electronic cigarettes and thirdhand tobacco smoke: two emerging health care challenges for the primary care provider. *Int J Gen Med* 2011 February 1;4:115-20.:115-20.

(52) Schick S. Thirdhand smoke: here to stay. *Tob Control* 2011 January;20(1):1-3.

APPENDIX 1: List of publications authored during PhD

Published

1. *Genetic Variants in Novel Pathways Influence Blood Pressure and Cardiovascular Disease Risk.* International Consortium for Blood Pressure Genome-Wide Association Studies. Nature 2011 Sep 11. 478(7367)*;103-9.*

2. *Genetic determinants of major blood lipids in Pakistanis compared with Europeans*. Saleheen D, Soranzo N, Rasheed A, Scharnagl H, Gwilliam R, **Alexander M** & alii. Circ Cardiovasc. Genet. 2010 Aug;3(4):348-57.

3. *Association of the 9p21.3 locus with risk of first-ever myocardial infarction in Pakistanis: case-control study in South Asia and updated meta-analysis of Europeans.* Saleheen D, **Alexander M** & alii. J. Arterioscler Thromb Vasc Biol. 2010 Jul;30(7):1467-73. [Note: Saleheen D. and Alexander M. are joint first authors].

In preparation

4. Behavioural, biological and socio-economic risk factors for myocardial infraction in South Asia: case-control study in urban Pakistan. **Alexander M**, Saleheen D, Di Angelantonio E, Kee Ho W, Johnson L, Danesh J & alii.

5. *Smoking cessation and access to tobacco outlet: English experience.* Han T, Niggebrugge A, **Alexander M**, Battersby J, Hollands G, Marteau T.

6. *Interplay of cigarette smoking with metabolic risk factors on the incidence of major vascular morbidity; an individual participant meta-analysis of 1 million people.* **Alexander M,** The Emerging Risk Factors Collaboration.

Appendix 2: Socio-economic status and the risk of myocardial infarction in Pakistan

Background

In developed countries, lower socioeconomic groups have been shown to carry more CVD risk factors and to be at increased risk of MI [1,2], but results have been inconsistent in other parts of the world [3]. In Pakistan, a developing country with a narrower middle class and a bigger proportion of poor individuals compared to developed countries[4], the importance of socio-economic status in relation to CVD risk has not yet been investigated. PROMIS collected information on several markers of socio-economic status (income, education, household possessions, employment status and category of job), lifestyle, medical and biochemical information on CVD risk factors in more than 12,000 first ever MI cases and age and sex frequency matched controls. This large dataset provides a unique opportunity to investigate more precisely than has been previously possible the relationship between multi-faceted socio-economic status and the risk of MI in Pakistan [5].

Methods

A principal component analysis was used to define socio-economic status using variables of income, education, asset ownership and occupation. As occupation was added later in the questionnaire and available only on a subset of 4826 cases and 5417 controls, sensitivity analyses were done to compare the PCA results with and without occupation on individuals with non missing information. ORs were computed using unconditional logistic regression, and adjusted for at least the "basic covariates": age, sex, recruitment centre, sub-ethnicity, tobacco use, history of diabetes or hypertension, family history of CAD, waist to hip ratio and LDL-cholesterol. Participants were excluded if they did not provide information on all the covariates for basic adjustment, retaining 6057 cases and 6889 controls. To characterize shapes of associations, ORs calculated within quintiles or within pre-defined categories were plotted against mean values within each quintile or category; 95% confidence intervals were represented using "floating absolute variances" which enables graphical comparisons of ORs between every two categories, and not only with the reference group [6]. I investigated the possibility of effect-modification by age, sex and other relevant subgroups fitting interaction terms and representing the shapes of association within subgroups. Analyses were conducted using STATA v10 (StataCorp, Texas).

Results

Amongst controls, 77% were employed and 8% were unemployed, the rest declaring to be retired. The main occupations in Men were professional - 15%, skilled and unskilled labour - respectively 16% and 20% -; while nearly 9 out of 10 women declared being housewives (**Table A2.1**). The median self-reported monthly income was 12,000 Pakistani rupees, which corresponds to approximately 140$ US. The most commonly owned items were in order a television, a mobile phone, a home, a radio, a motorcycle, a bicycle, a computer, air conditioning and land. Around two thirds of men (69%) reported having received a formal education and 50% reported more than 10 years of education, while 51% percent of women reported no formal education.

The principal component analysis of socio-economic indices identified one main gradient labelled "socio-economic status" which explained 14% of the variance of all socio-economic indices (**Figure A2.1 & A2.2**), and was approximately normally distributed (**Figure A2.3**).

In the top third of socio-economic status (labelled "upper class"), the median income was 25,000 Pakistani rupees, 14 years of education and 6 household possessions; in the middle third it was 12,000 Pakistani rupees, 8 years of education and 4 household possessions; and in the bottom third it was 7,000 Pakistani rupees, no formal education and 3 household possessions (**Table A2.2**). In the upper class, professionals were largely over-represented, and 72% reported more than 10 years of education, compared to 21% of the middle class and 2% in the lower class. Individuals belonging to the upper class were more likely to be never smokers and to cook using oil rather than *ghee*. Women, Pathan and Balochi ethnic groups were the most socio-economically deprived groups, with more than 50% of the individuals in these groups belonging to the bottom third of socio-economic status, labelled 'lower class'; while Men, Sindhi and Urdu ethnic groups were the most privileged groups (**Table A2.3**).

The middle and upper classes of the "socio-economic status" were at lower risk of MI risk than the middle class (**Figure A2.4**). The protection conferred by belonging to the upper class was attenuated amongst individuals aged ≤50 years old (**Figure A2.5**).

Discussion

In developed countries a low socio-economic position has been shown to increase MI risk [1,2] whilst results have been inconsistent in middle and low-income regions [3]. In PROMIS, individuals in the middle range of socio-economic status were at increased risk of MI while individuals in top third of the socio-economic gradient were protected. The middle group was more likely to smoke, use *ghee*, to have a self-reported history of diabetes or hypertension, and to have a family history of CAD than individuals belonging to the lower group.

Urdu were over-represented amongst the upper class and were also more likely to report a family history of MI -20% of Urdu versus 16% of Punjabi, 6% of Pathan and 5% of Sindhi - and had higher LDL-cholesterol levels (p-value<0.0001). Urdu speakers have also been

shown to have higher age adjusted prevalence of hypertension than Punjabi and Sindhi individuals [7]. However, adjustment for these covariates as well as ethnicity in the model did not alter the strength of the association between socio-economic status and MI risk. There may be some residual confounding such as amount and duration of tobacco use, but better access to health care and better awareness of the danger of tobacco use in the upper class group is most likely to explain their protection against MI risk. In this respect, lack of awareness of the danger of smoking and smokeless tobacco have been correlated with lower education levels in previous studies in Pakistan [8,9]. In a cross-sectional study among Pakistani students of a private medical university with equal proportion of men and women, the average income of the household of students was 50,000 Pakistani rupees per months, meaning most would be classified in the top third of the socio-economic gradient and labelled "upper class", and only 7% of students were tobacco users, compared to 34% of men and 12% of women in the general Pakistani population [10,11].

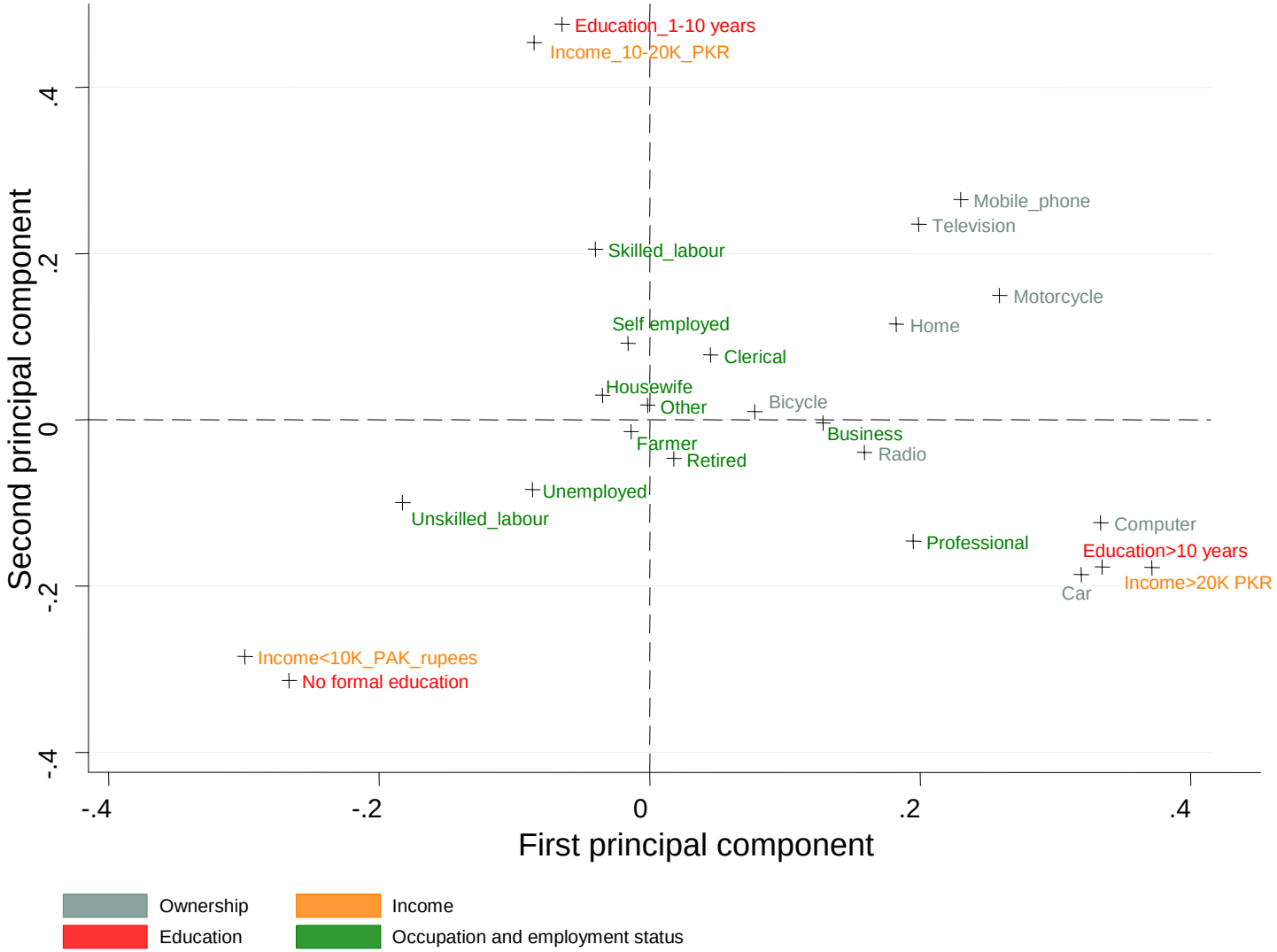
Measurement error in socio-economic indices resulting from mis-reporting of income and asset ownership, errors in recall or mis-categorization into type of jobs, may also have resulted in misclassification of individuals [12]. In addition, socio-economic position determined by principal component analysis of asset-ownership has been shown to correlate poorly in developing countries with consumption expenditure [13], and therefore may not be an appropriate risk factor. Finally, the first component of the principal component analysis labelled as "socio-economic gradient" only explained 14% of the overall variance of socio-economic indices, and therefore only captured a small part of the multidimensional effects of income, education, asset ownership, employment and type of job.

INTERHEART reported a weaker association of socio-economic status in low income countries compared to high income countries, and a weak association worldwide of income with MI risk; whilst education was the marker most consistently associated with risk for acute MI globally [14]. In this context and in light of my findings, the use of a single dimension of socio-economic status rather than a composite variable, namely education, may be more appropriate when assessing MI risk in the developing world and more specifically in Pakistanis. In my study as in INTERHEART, education was continuously and log linearly associated with a decrease in MI risk. OR for ≥20 years of education versus no education: was 0.73 (0.65; 0.82).

Conclusion

Compared to lower socio-economic groups, Individuals classified as middle class on the basis of income, education, household possessions, job and employment status, were at increased risk of MI in a Pakistani urban population, whilst the upper class was protected.

Figure A2.1: Principal component analysis showing second versus first component



Retaining two components without rotation of the matrix of loadings.

Figure A2.2: Scree plot of Eigen values over the number of principal components

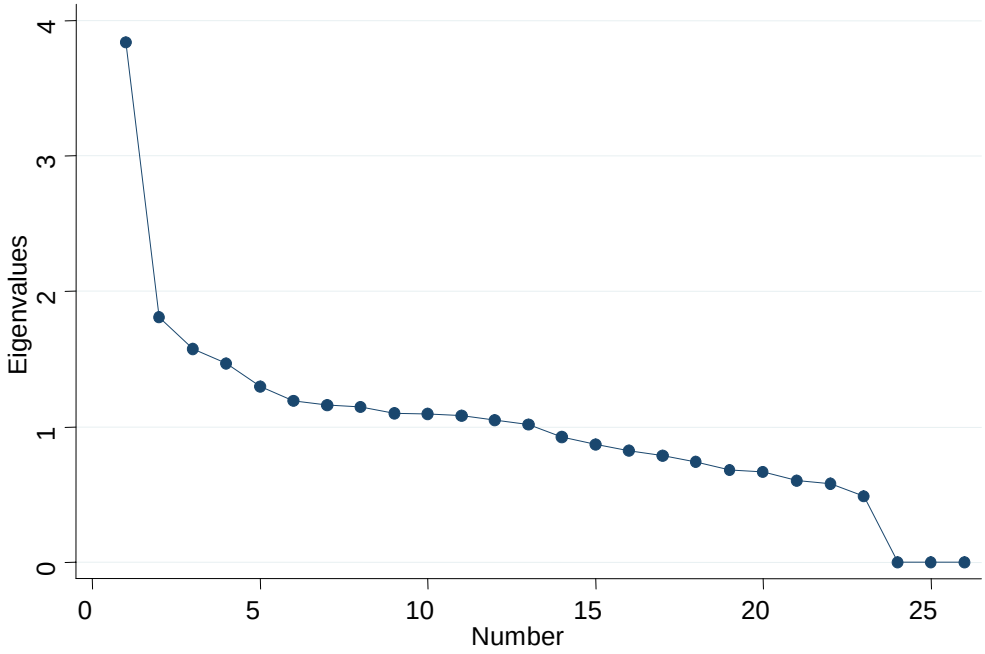
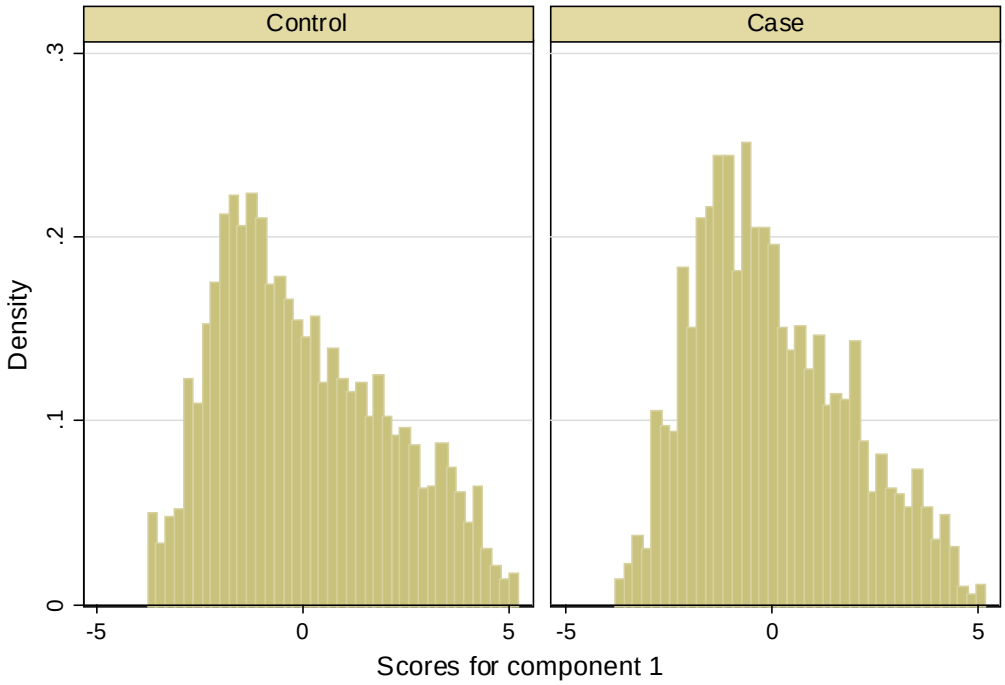
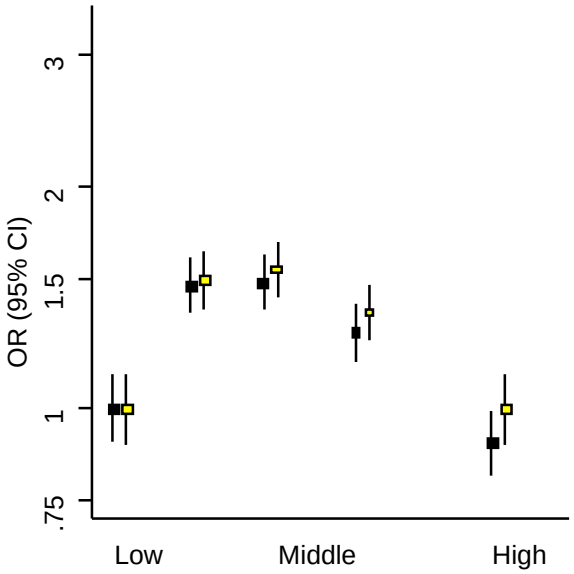


Figure A2.3: Distribution of the socio-economic gradient by case-control status



Graph by case-control status

Figure A2.4: Association of socio-economic status with MI risk



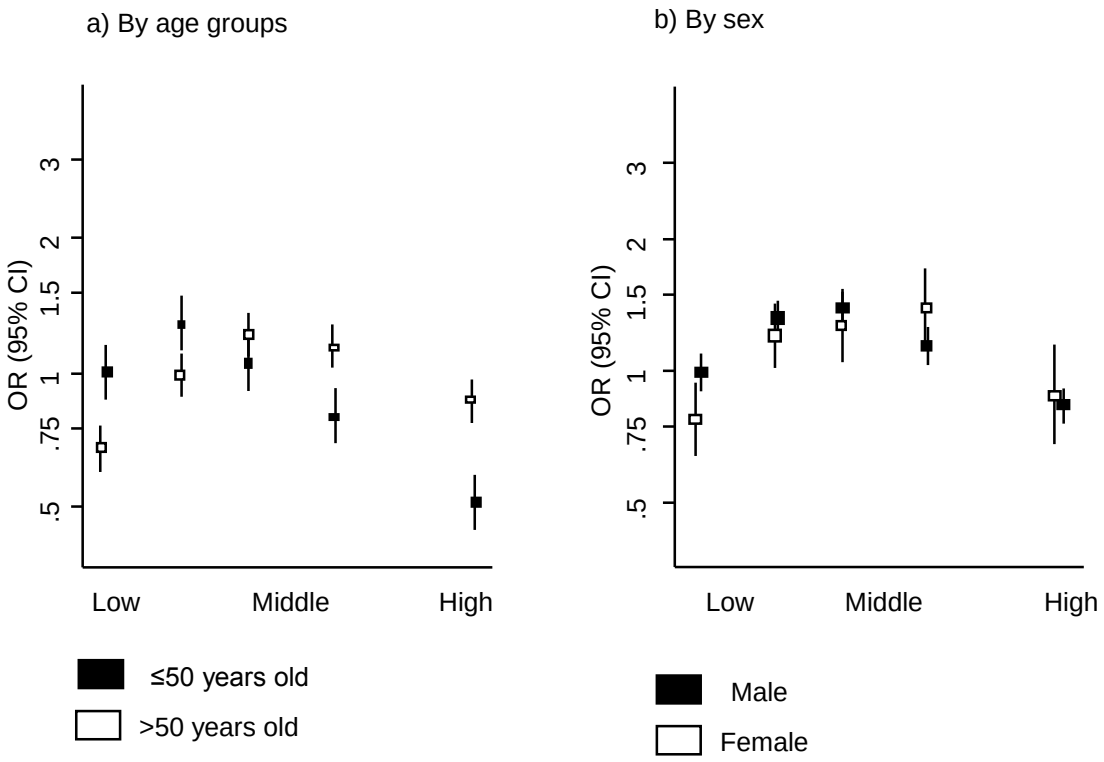
Adjusted for age, sex, ethnicity and centre, diabetes, hypertension, family history of CAD, LDL, tobacco use, WHR

Further adjusted for socio-economic status, dietary patterns and cooking oil

Note: Dataset reduced to 4521 cases and 4603 controls with information on all the covariates

Odds ratios were computed within fifths of income, education and number of household possessions; and were plotted against the arithmetic mean within each fifth; except for income and education where the geometric mean was used to account for the skewness of the distributions. The size of the box is proportional to the inverse of the variance of OR.

Figure A2.5: Association with socio-economic status by age groups and by sex



OR: Odds ratios. Models adjusted for gender, ethnicity and centre, tobacco use, LDL-C, WHR, history of diabetes and hypertension

Table A2.1: Descriptive characteristics of socio-economic status by status and sex

	Cases		Controls		Overall P-value of difference cases versus controls
	% Male / % Female (n=5503/n=1456)	P-value Male v Female	% Male / % Female (n=4643 / n=922)	P-value Male v Female	
Formal education		<0.0001		<0.0001	<0.0001
No	31% / 64%		30% / 60%		
≤10 years	40% / 29%		35% / 26%		
>10 years	30% / 7%		35% / 14%		
Income		0.8		0.003	<0.0001
<10K PKR	26% / 25%		34% / 36%		
10-20K PKR	34% / 36%		33% / 36%		
≥20K PKR	40% / 39%		33% / 28%		
Occupation[1]		<0.0001		<0.0001	<0.0001
Professional	9% / 4%		13% / 7%		
Skilled labour	17% / 3%		14% / 3%		
Unskilled labour	15% / 6%		16% / 5%		
Housewife	0% / 25%		0% / 47%		
Farmer	2% / 0%		2% / 0%		
Business	12% / 0%		10% / 1%		
Clerical	6% / 0%		7% / 0%		
Self-employed	6% / 1%		6% / 0%		
Other	12% / 3%		10% / 3%		
Unemployed	6% / 50%		5% / 30%		
Retired	15% / 7%		16% / 4%		
Ownership					
Home	80% / 80%	0.7	72% / 72%	0.6	<0.0001
Car	17% / 14%	0.03	19% / 14%	<0.0001	0.2
Motorcycle	42% / 38%	0.01	39% / 30%	<0.0001	<0.0001
Bicycle	25% / 22%	0.02	27%/ 22%	<0.0001	0.06
Radio	41% / 38%	0.1	47% / 42%	0.0003	<0.0001
Television	89% / 88%	0.6	83% / 82%	0.5	<0.0001
Land	12% / 7%	<0.0001	16% / 9%	<0.0001	<0.0001
Computer	18% / 17%	0.26	23% / 15%	<0.0001	<0.0001
Mobile phone	88% / 84%	0.001	75% / 69%	<0.0001	<0.0001

[1]: For occupation, in cases information was available in 4391 males and 248 females provided information; and in controls in 4682 males and 662 women; % corresponds to a column percentage. PKR: Pakistani rupees.

Table A2.2: Socio-economic indices across quintiles of the socio-economic gradient

Socio-economic indices	Q1	Q2	Q3	Q4	Q5	P-value
Monthly income, median (PKR)	6,000	8,000	13,000	20,000	30,000	<0.0001
Formal education, median (years)	0	0	8	12	14	<0.0001
Number of household possessions, median	2	3	4	5	6	<0.0001
Occupation/employment (row %)						
Professional	1	2	10	25	62	<0.0001
Skilled labour	15	25	27	22	11	
Unskilled labour	50	29	12	7	2	
Housewife	22	28	29	16	5	
Farmer	18	29	22	23	9	
Business	3	5	17	29	45	
Clerical	4	11	27	36	22	
Self employed	10	23	29	24	13	
Other occupation	15	23	20	22	21	
Unemployed	40	26	19	10	5	
Retired	16	24	21	23	16	

Qi: Quintile number i. P-value of linear trend for income, education and number of household possessions; and p-value from a Chi2 test of independence for occupation/employment.

Table A2.3: Correlates of socio-economic gradient

	'Socio-economic gradient'			
	Low	Middle	High	P-value
Demography				
Male sex	73%	77%	86%	<0.0001
Age, mean (SD)	54.2 (9.83)	53.2 (9.45)	51.3 (8.98)	<0.0001
Major ethnic groups				<0.0001
Punjabi	26%	33%	31%	
Urdu	35%	39%	41%	
Pathan	13%	7%	2%	
Sindhi	14%	9%	14%	
Biochemistry, medical and familial				
Diabetes	13%	15%	15%	0.08
Hypertension	23%	29%	31%	<0.0001
Family history of MI	8%	15%	22%	<0.0001
LDL-cholesterol (mmol/l), mean (SD)	2.76 (1.07)	2.87 (1.05)	2.99 (1.06)	<0.0001
Waist to hip ratio, mean (SD)	.934 (.0659)	.946 (.0657)	.95 (.0638)	<0.0001
Tobacco use				<0.0001
Never/ex	62%	67%	75%	
Dip naswar only	9%	5%	1%	
Chew only	7%	8%	3%	
Smoke (and dip/chew)	21%	20%	21%	
Diet				
Cooking fat				<0.0001
Oil	42%	47%	59%	
Oil & ghee	20%	29%	32%	
Ghee only	38%	25%	9%	
Dietary pattern 1				<0.0001
Low	30%	34%	36%	
Middle	35%	37%	28%	
High	34%	29%	36%	
Dietary pattern 2				<0.0001
Low	53%	33%	16%	
Middle	30%	40%	30%	
High	17%	27%	55%	

Note: Column percentages. Low corresponds to the bottom third, middle to the middle third and high to the top third of the distribution of the socio-economic gradient. *P-values derive from Chi2 tests of independence for categorical variables and from a Fisher test of equality of the means (computed as a test of nullity of all the coefficients in the regression of the continuous row variable - for example LDL-C - over categories of the column variable - for example tobacco use).

References

1. Avendano M, Kunst AE, Huisman M et al. Socioeconomic status and ischaemic heart disease mortality in 10 western European populations during the 1990s. *Heart.* 2006;92:461-467.

2. Mackenbach JP, Stirbu I, Roskam AJ et al. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med.* 2008;358:2468-2481.

3. Manrique-Garcia E, Sidorchuk A, Hallqvist J, Moradi T. Socioeconomic position and incidence of acute myocardial infarction: a meta-analysis. *J Epidemiol Community Health.* 2011;65:301-309.

4. Ravaillon, M. The developing world's bulging (but vulnerable) "middle class". Policy Research Working Paper 4816. The World Bank. 4816, 1-28. 2009. Washington DC, USA.

5. Dennis B, Aziz K, She L et al. High rates of obesity and cardiovascular disease risk factors in lower middle class community in Pakistan: the Metroville Health Study. *J Pak Med Assoc.* 2006;56:267-272.

6. Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med.* 1991;10:1025-1035.

7. Jafar TH, Levey AS, Jafary FH et al. Ethnic subgroup differences in hypertension in Pakistan. *J Hypertens.* 2003;21:905-912.

8. Alam AY, Iqbal A, Mohamud KB, Laporte RE, Ahmed A, Nishtar S. Investigating socio-economic-demographic determinants of tobacco use in Rawalpindi, Pakistan. *BMC Public Health.* 2008;8:50.:50.

9. Khawaja MR, Mazahir S, Majeed A et al. Chewing of betel, areca and tobacco: perceptions and knowledge regarding their role in head and neck cancers in an urban squatter settlement in Pakistan. *Asian Pac J Cancer Prev.* 2006;7:95-100.

10. Nisar N, Qadri MH, Fatima K, Perveen S. A community based study about knowledge and practices regarding tobacco consumption and passive smoking in Gadap Town, Karachi. *J Pak Med Assoc.* 2007;57:186-188.

11. Nanan, D. and White, F. PROCOR: 7/25/99: The National Health Survey of Pakistan: Review and discussion of report findings pertaining to selected risk factors for cardiovascular disease. 25-7-1999.

12. Onwujekwe O, Hanson K, Fox-Rushby J. Some indicators of socio-economic status may not be reliable and use of indices with these data could worsen equity. *Health Econ.* 2006;15:639-644.

13. Howe LD, Hargreaves JR, Huttly SR. Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. *Emerg Themes Epidemiol.* 2008;5:3.:3.

14. Rosengren A, Subramanian SV, Islam S et al. Education and risk for acute myocardial infarction in 52 high, middle and low-income countries: INTERHEART case-control study. *Heart.* 2009;95:2014-2022.

Appendix 3: Diet and the risk of myocardial infarction in Pakistan

Background

A "prudent" pattern of diet high in fruit, vegetables and fish intake, has been shown to reduce prospective CVD risk in Western populations [1-4]. However, there is sparse evidence on the importance of diet as a risk factor for CVD in other populations [5,6]. In addition, whilst Western diets are well characterized, South Asian diet remains little studied in indigenous settings. In studies set amongst immigrants, South Asians reported higher mean total energy intake, including higher intakes of total fat, polyunsaturated fat and protein, than White Europeans [7,8] and Americans [9]. However, migrants varied in their adherence to traditional diet and were shown to have adopted some Western food, such as snacks [10].

In this context, PROMIS represents the largest scale epidemiological resource with detailed dietary information. PROMIS recorded frequency of 43 food items locally relevant to Pakistani diet in more than 12,000 first ever MI cases and matched controls recruited in 5 urban centers across Pakistan, with additional information on demography, lifestyle, as well as medical and family history.

Methods

A locally relevant food frequency questionnaire was developed following a 24 hour recall and a 7-days food diary on a subset of 200 healthy adults as described previously [11]. Questions related to food rather than drink consumption were used in this analysis. Categories of food frequency were converted into weekly consumption (**Table A3.1**).

Principal component analysis on the converted food items with rotation of the matrix of loading was conducted to reveal main dietary patterns. Major components were determined after looking at the scree-plot (**Figure A3.1**), which shows the drop in variance explained by each additional component; and the interpretability of the components obtained (**List A3.1**).

Sensitivity analyses were run which (1) excluded individuals with self reported hypertension or diabetes (2568 case and 2528 controls); (2) defined quintiles of dietary patterns within sex; (3) computed the matrix of loadings in controls only and inferred scores for cases and controls; (4) ran the principal component analysis of food groups rather than food items. Analyses were conducted using STATA v10 (StataCorp, Texas).

Odds ratios (ORs) were computed using unconditional logistic regression, and adjusted for at least the "basic covariates": age, sex, recruitment centre and sub-ethnicity, tobacco use, history of diabetes or hypertension, family history of CAD, waist to hip ratio and LDL-cholesterol. To characterize shapes of associations, ORs calculated within quintiles or within

pre-defined categories were plotted against mean values within each quintile or category. For skewed variables such as income, the geometric mean was reported rather than the arithmetic mean.

When assessing the shape of association with MI, 95% confidence intervals were derived from "floating absolute variances" to enable graphical comparisons of ORs between every two categories, and not only with the reference group [12]. I investigated the possibility of effect-modification by age, sex and other relevant subgroups fitting interaction terms and representing the shapes of association within subgroups. To avoid false-negative findings and because of the large number of tests performed, only p-values <0.001 were emphasized. Analyses were conducted using STATA v10 (StataCorp, Texas).

Results

Principal component analysis of the food frequency questionnaire identified two main dietary patterns: a "high vegetables and carbohydrates diet" (with loadings >0.2 for cooked and raw vegetables, and potatoes) and a "high meat, fish and sweets pattern" (with loadings >0.2 for chicken, beef, mutton, fish and sweets), each pattern accounting for 9% of the total variance of the dietary data (**Figure A3.2**).

There was considerably greater consumption of raw and cooked vegetables (but only slightly higher intake of carbohydrates) across the fifths of values of the first principal component (**Figure A3.3** & **Figure A3.4**). Controls scoring in the top third of the "high vegetables and carbohydrates diet" compared to the bottom third were more likely to be women, Urdu, with a self-reported history of hypertension, oil rather than ghee users and to score high for the "high meat, fish and sweet diet". Controls above the median score for "high meat, fish and sweet diet" were in average younger, more likely be men, Urdu, with a family history of MI, to smoke, to belong to the upper class, to use oil rather than *ghee* for cooking; and less likely to be diabetics.

The "high vegetables and carbohydrates" diet association with MI risk was non linear (**Figure A3.5**), with a positive relationship below the median and a negative relationship above the median. The association did not attenuate upon adjustment for socio-economic status, the other dietary patterns and cooking oil; upon adjustment for individual food items scoring high for this component (cooked and raw vegetables, pulses and nut, potatoes and rice) (**Figure A3.6**). There was no evidence of an effect modification by age and sex (**Figure A3.7**).

The "high meat, fish and sweet" pattern was positively and approximately log-linearly associated with MI risk. Individuals scoring in the top 5th of the distribution had an OR of 1.79 (1.54; 2.08) compared to the bottom 5th. Further adjustment for the other dietary pattern, cooking oil and socio-economic status only modestly attenuated the ORs. Sensitivity analyses identified similar dietary patterns, yielding similar shapes and strength of association with MI risk.

Discussion

Over the past 30 years, South Asian countries have been experiencing a dietary transition, which has been accompanying the rural exodus and rapid urbanization [13,14]. The adoption of a westernized diet especially rich in meat and sweets has been linked to urbanization [15,16] and is also happening amongst Pakistani immigrants settling in Western countries [17].

The nutrition transition has been marked by a shift from diets high in carbohydrates and based on indigenous staple grains or starchy roots, locally grown legumes, vegetables and fruits and limited foods of animal origins; toward more fat-rich diets than include more pre-processed food of animal origin, more sweets and sugar and more fried food cooked in ghee and oil [13,18,19]. In PROMIS, traditional staples correlated positively with fruit and vegetable consumption on the first dietary pattern; while meat consumption correlated positively with sweet snacks on the second dietary pattern.

In Western populations, adherence to a diet high in red and processed meat has been shown to increase MI risk whilst a "prudent diet" high in fruit, vegetables and fish intake was protective [1-4]. In this study, a "high meat, fish and sweets" diet was associated with a positive and approximately log linear relationship with MI risk (adjusted OR for top 5th versus bottom 5th: 1.79 (95% CI: 1.54; 2.08)). Individuals who scored in the top 3rd of this dietary pattern and who also cooked with ghee had more than triple the risk of MI compared to individuals in the bottom 3rd of the high meat diet and cooking with oil only (adjusted OR equal to 2 (2.4; 4.3)). Pakistan is the country with the highest per capita consumption of meat in all of South Asia [20], being twice as high as in India, and therefore public health campaigns highlighting to the health effects of both high meat consumption and cooking with ghee could be considered.

In contrast, the "high vegetables and carbohydrates" pattern was not associated with a reduction in risk, and the shape of association was an inverted "J". Individuals scoring around the median of this dietary pattern experienced an increased risk, even after multiple adjustments. Rather than boiled or steamed, vegetables are usually fried in Pakistan [8] and higher intake of vegetables may mask a greater consumption of edible fat. Quantities of edible fat were not reliably recorded in this dataset and it was not possible to adjust for this variable. In addition, individuals in the middle group of the "high vegetables and carbohydrates" pattern were likely to belong to the middle class, a socio-economic group who has been the most rapid in adopting a Westernized lifestyle of low physical activity and low-quality diet. Tthe lower class has been slower in adopting these practices and the upper class has access to better quality produces, and has been shown to practice leisure time physical activity [19].

This study has several limitations. The dietary questionnaire enquired about the frequency of food consumption and did not capture the quantity of food consumed. As a result, I was unable to adjust for total energy intake. In addition, food frequency questionnaires are prone to measurement error due to imperfect recall and within person variability. This may result in

misclassification of individuals, and is more likely to happen amongst individuals reporting food frequencies not far from the average, such as individuals in the middle tertile of the middle tertile of the distribution of each dietary pattern.

Conclusion

As part of a strategy to reduce the epidemic burden of CVD in Pakistan and more generally in South Asia, dietary programs could be considered which target individuals with a high consumption of meat and sweets.

Figure A3.1: Scree plot of Eigen values over number of principal components

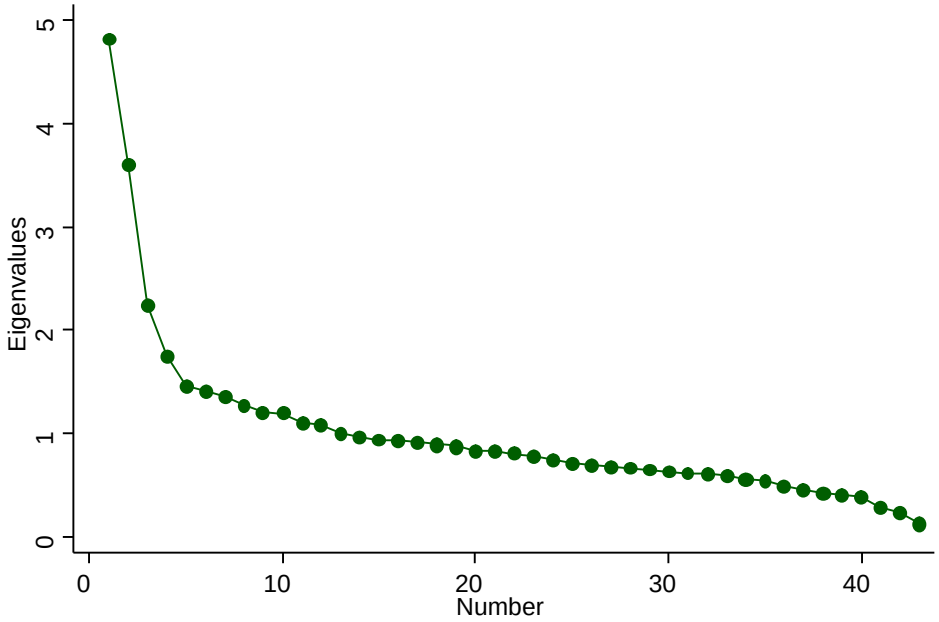
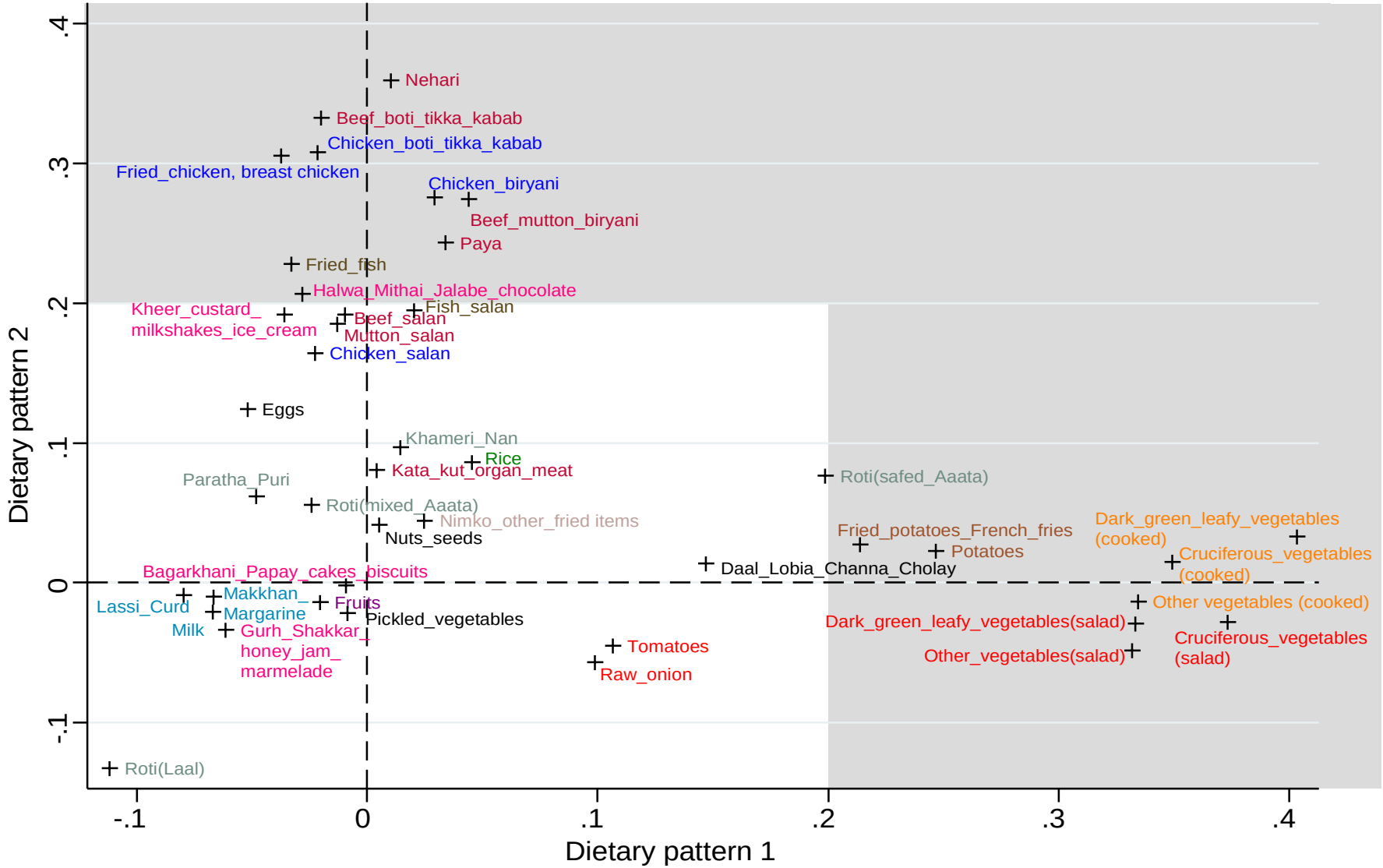


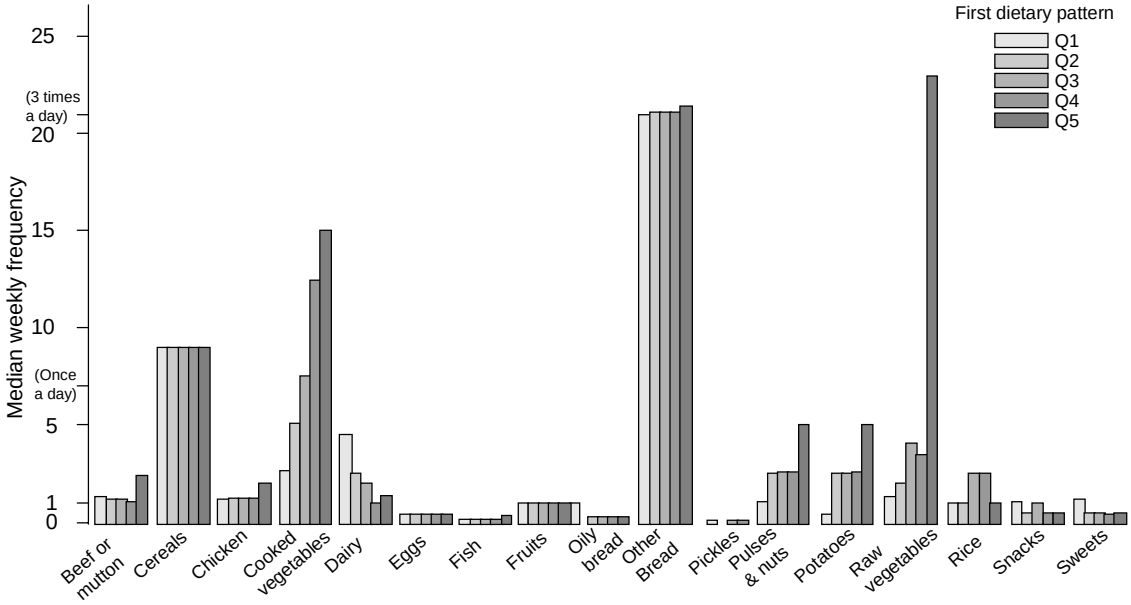
Figure A3.2: Plot of the first and second principal components labelled as “dietary pattern 1” and “dietary pattern 2”



The first component explains 10.2% of the variance, the 2nd component 8.1% and the 3rd component 4.5% of the variance.

Figure A3.3: Median frequency of weekly intake of food groups across fifths of the two dietary patterns

a) By Fifths of the first dietary pattern labelled “high carbohydrates and vegetables” diet



b) By fifths of the second dietary pattern labelled “high meat and sweets” diet

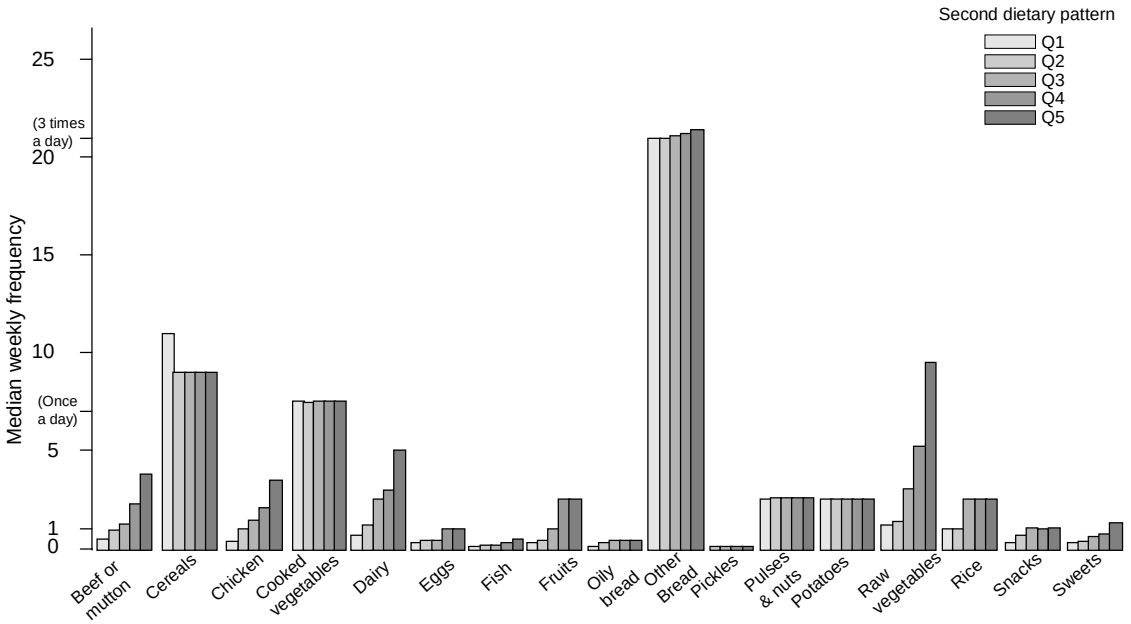
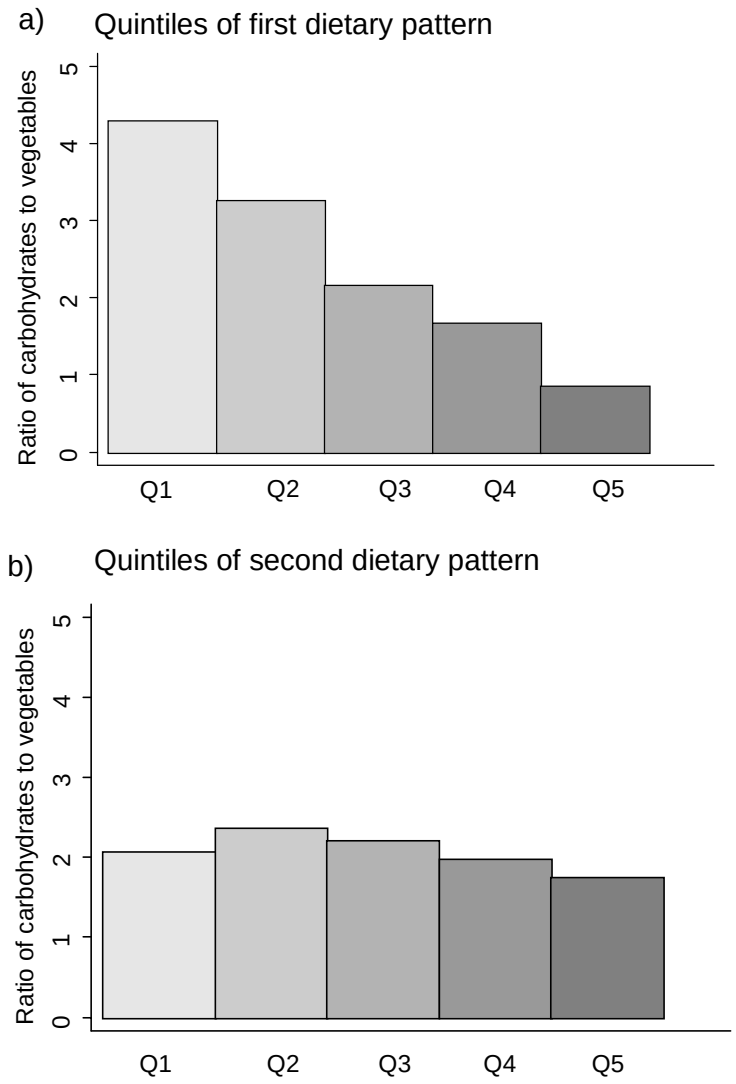
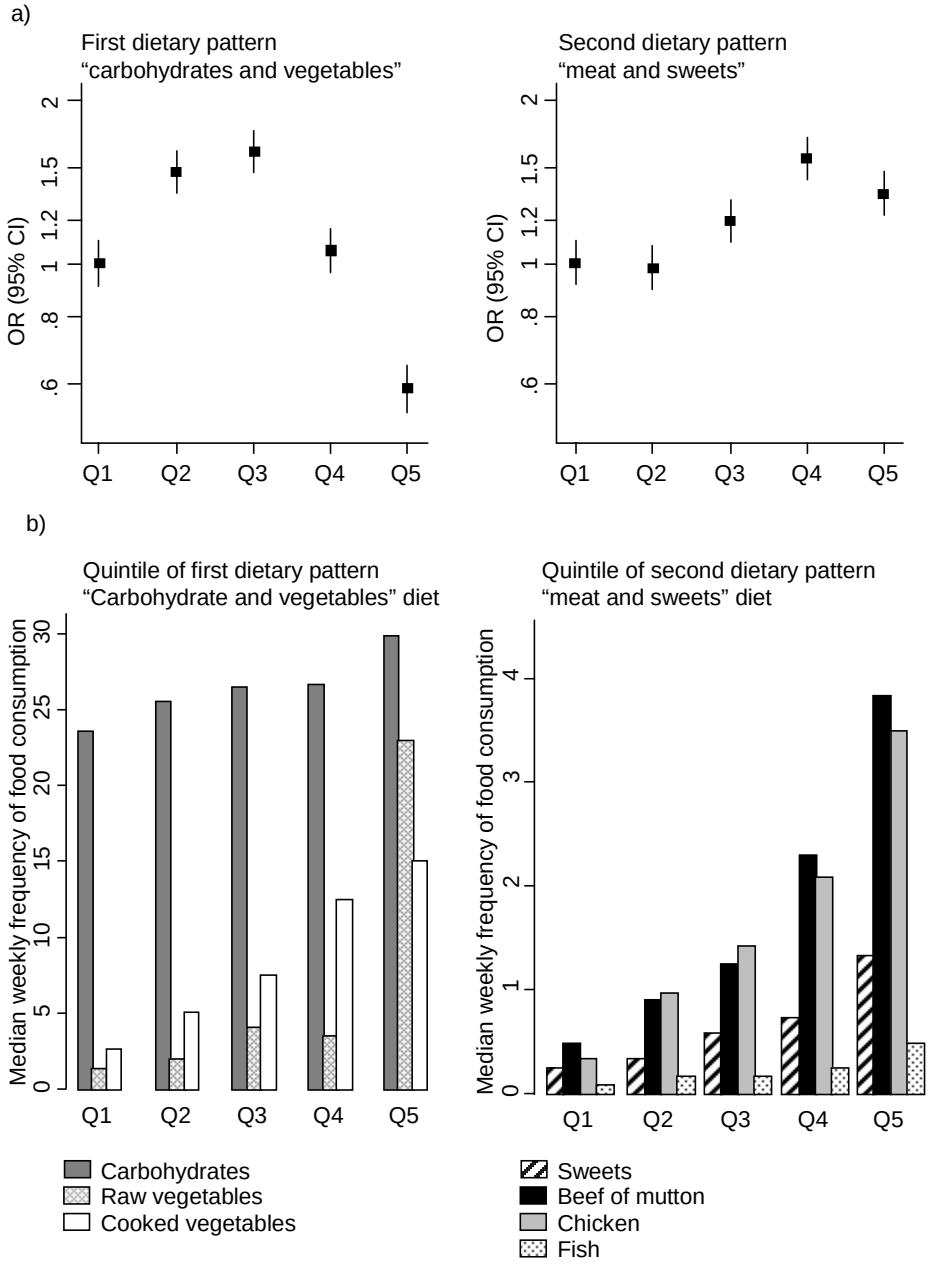


Figure A3.4: Ratio of carbohydrate to vegetable intake across fifths of the two dietary patterns



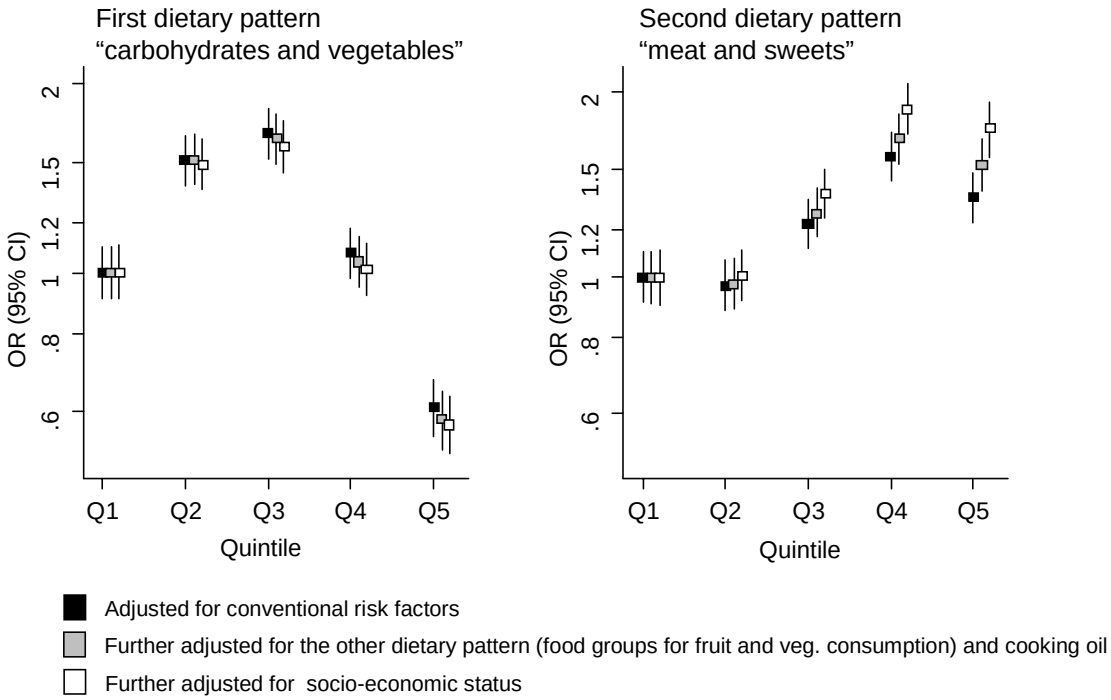
Q: Quintile. The ratio of carbohydrates to vegetables was defined as the ratio of the following food items converted into weekly frequencies: (rice + fried potatoes + potatoes + *roti safed Aaata* + *khameri* + *rotil laal* + *roti mixed Aaata*) divided by (cooked dark green vegetables + cooked cruciferous vegetables + other cooked vegetables + dark green vegetables in salad + cruciferous vegetables in salad + other vegetables in salad + tomatoes + raw onions).

Figure A3.5: Odds ratios for myocardial infarction with dietary patterns



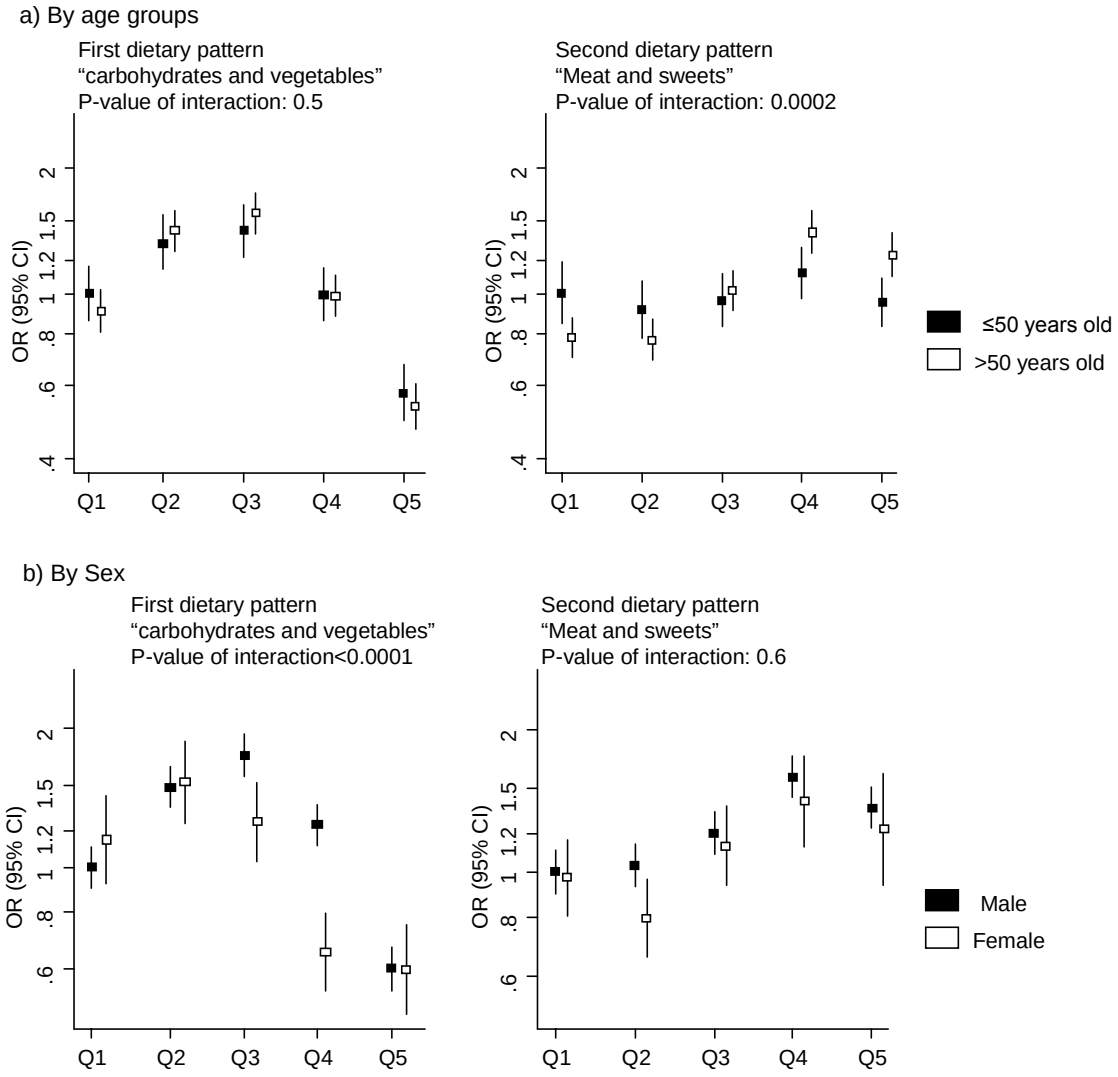
Panel a) OR: Odds Ratio, CI: confidence intervals, Q: Quintiles of the exposure variable. All analyses were adjusted for age, sex, self-reported ethnicity, recruitment centre, tobacco use, diabetes status, hypertension status; LDL-cholesterol levels, waist-to-hip ratio, and family history of MI. Confidence intervals were calculated using "floating absolute variances". Participants in lowest fifths are referents. Dietary patterns were calculated using principal component analyses (PCA). **Panel b)** across fifths of each dietary pattern, in the control participants, median weekly consumption of major food groups that have a loading value of ≥0.20 on PCA analyses.

Figure A3.6: Progressive adjustment of the association between dietary patterns and MI risk



OR: Odds Ratio, CI: confidence intervals, Q: Quintiles of the exposure variable.

Figure A3.7: Association between dietary patterns and MI risk by age groups and by sex



Models adjusted for gender, age, ethnicity and centre, LDL-C, tobacco use, WHR, family history of CAD, history of hypertension or diabetes.

Table A3.1: Conversion of food frequency questions into weekly consumption

Nominal Level	Coded as	Converted to weekly consumption	Converted to daily consumption
More than 4 times per day	1	4X7=28	4
Between 2 and 4 times per day	2	3X7=21	3
Once per day	3	1X7=7	1
More than 3 times per week	4	5	5/7 = .71428571
2 to 3 times per week	5	2.5	2.5/7 = .35714286
Once per week	6	1	1/7 = .14285714
Less than once per week	7	1/2.5=0.4	1/14 = .07142857
Once per month	8	1/4=0.25	1/30 = .03333333
Less than once per month	9	1/12=0.083333	1/80 = .01250000
During Ramadan only	10	.	.
Never	11	0	.
Everything else	*	.	.

Table A3.2: Correlates of dietary patterns in controls

	Dietary pattern 1				Dietary pattern 2			
	Low (n=1817)	Middle (n=1684)	High (n=2326)	*P-value*	Low (n=2167)	Middle (n=1940)	High (n=1720)	*P-value*
Demography								
Male sex	82%	73%	79%	<0.0001	74%	75%	85%	<0.0001
Age, mean (SD)	53.2 (9.49)	52.47 (9.33)	53.17 (9.49)	0.02	54.91 (9.28)	53.22 (9.37)	50.71 (9.19)	<0.0001
Major ethnic groups				<0.0001				<0.0001
Punjabi	30%	39%	25%		33%	37%	24%	
Urdu	38%	35%	41%		33%	35%	46%	
Pathan	6%	7%	8%		8%	7%	6%	
Sindhi	10%	11%	15%		14%	9%	12%	
Biochemistry, medical and familial history								
Diabetes	16%	13%	14%	0.1	17%	16%	11%	<0.0001
Hypertension	25%	28%	32%	<0.0001	26%	30%	29%	0.002
Family history of MI	16%	18%	14%	0.002	11%	17%	19%	<0.0001
LDL-cholesterol (mmol/l), mean (SD)	2.86 (1.03)	2.92 (1.08)	2.83 (1.06)	0.03	2.85 (1.1)	2.85 (1.01)	2.91 (1.05)	0.1
Waist to hip ratio, mean (SD)	0.95 (0.07)	0.95 (0.06)	0.94 (0.06)	0.3	0.94 (0.07)	0.95 (0.07)	0.95 (0.06)	0.02
Tobacco use								<0.0001
Never/ex	69%	70%	66%		68%	71%	66%	
Dip naswar only	4%	4%	6%	0.001	6%	5%	3%	
Chew only	5%	6%	7%		6%	5%	7%	
Smoke (and dip/chew)	22%	20%	22%		21%	20%	24%	
Socio-economic status				<0.0001				<0.0001
Low	29%	34%	33%		50%	29%	16%	
Middle	34%	37%	29%		33%	40%	27%	
High	37%	29%	38%		17%	31%	57%	
Diet								
Cooking fat				<0.0001				<0.0001
Oil	43%	54%	53%		44%	51%	55%	
Oil & ghee	30%	20%	31%		23%	25%	33%	
Ghee only	27%	26%	16%		34%	24%	12%	
Dietary pattern 1				<0.0001				<0.0001
Low	100%	0%	0%		39%	35%	26%	
Middle	0%	100%	0%		33%	36%	31%	
High	0%	0%	100%		28%	29%	43%	
Dietary pattern 2				<0.0001				<0.0001
Low	39%	33%	28%		100%	0%	0%	
Middle	35%	36%	29%		0%	100%	0%	
High	26%	31%	43%		0%	0%	100%	

Note: Column percentages. *P-value derives from Chi2 tests of independence for categorical variables and from a Fisher test of equality of the means (computed as a test of nullity of all the coefficients in the regression of the continuous row variable - for example LDL-C - over categories of the column variable - for example tobacco use).

List A3.1: Grouping of food items into food groups

Colour	Food Groups	Food items
	Beef and mutton	Beef or mutton biryani; Kata kut or organ meat; beef salan; mutton salan; beef boti, tikka, kabab, beef shawarma and others; Nehari; Paya
	Breads	Khameri or Nan (refined bread); Paratha/Puri (oily bread); Roti safed Aaata (unrefined bread); Roti Laal Aaata/chakki (refined bread); Roti mixed Aaata (refined bread)
	Cereals	Daliya
	Chicken	Chicken biryani; Chicken salan; Chicken boti, tikka, kabab, chicken roll; chicken shawarma and others; Chicken fried, chicken breast
	Cooked vegetables	Dark green leafy vegetables and yellow vegetables; cruciferous vegetables (Gobi, phool gobi, bang gobi, sursoon, others); other vegetables excluding potatoes
	Dairy products	Milk; Makkhan, Margarine; Lassi; Curd
	Eggs	
	Fish	Fish Salan, Fried Fish
	Fruits	
	Nimkod	
	Pickled vegetables	
	Potatoes	French fries, potatoes
	Pulses and nuts	Daal, lobia, channa, cholay; nuts and seeds
	Raw vegetables (salad)	Dark green leafy vegetables and yellow vegetables; cruciferous vegetables (Gobi, phool gobi, bang gobi, sursoon, others); other vegetables excluding potatoes; tomatoes as salad; raw onion
	Rice	
	Sweet snacks & desserts	Bagarkhani; Papay; Cakes; Biscuits; Gurrh; Sakkar; honey; jam; marmelade; Kheer; custard; Milkshakes; Ice cream; Halwa, Mithai, Jalabe, chocholate

List A3.2: Definition of local food items

Bagarkhani: Slightly sweet and salty phyllo puff-pastry

Bang gobi: Cabbage

Biryani: A means of cooking rice, where rice is mixed with a curry including meat or fish, eggs and

Boti: Pieces of meat

Channa: Chick peas

Cholay: Chick peas

Curd: Yoghurt

Daal: Lentil

Ghee: Clarified butter

Gurrh: Unrefined sugar

Halwa: Confectionary generally made from grain flour (typically semolina), oil and sugar

Hool gobi: Cauliflower

Jalabe chocolate: Confectionary made by deep-frying a kind of pretzel and then soak it in syrup

Katakut: Combination of spices, brains, liver, kidney and other organ meats

Khameri: type of Nan

Khees: Generally corn in milk gravy

Lassi: Yoghurt drink made by blending yogurt with water or milk and adding spices

Lobia: Black eyed beans

Makkhan: Butter

Mitahi: Confectionary made from unrefined sugar and milk/butter fried

Nan: Flat bread, stone baked in a clay oven (tandoor)

Nehari: Curry made usually from beef shank and more rarely lamb

Nimko: Mixture of spicy dried ingredients, which may include fried lentils, peanuts, chickpea flour noodles, corn, vegetable oil, chickpeas, flaked rice, fried onion and curry leaves. These are flavoured with salt and a blend of spices that may include coriander and mustard seed

Paratha: Unleavened flat bread made of whole-wheat and fried

Paya: Cow hoof

Puri: Unleavened bread prepared from aata (stone-ground whole-wheat flour) deep fried in ghee or vegetable oil.

Roti chakki: Bread whole-wheat flour

Roti Laal Aaata: Bread whole-wheat flour

Roti mixed aaata – bread, mixed flour

Roti safed Aaata: Bread, white flour

Sakkar: Fried flour sweet

Salan: a means of cooking curry

References

1. Oliveira A, Rodriguez-Artalejo F, Gaio R, Santos AC, Ramos E, Lopes C. Major habitual dietary patterns are associated with acute myocardial infarction and cardiovascular risk markers in a southern European population. *J Am Diet Assoc.* 2011;111:241-250.

2. Heidemann C, Schulze MB, Franco OH, van Dam RM, Mantzoros CS, Hu FB. Dietary patterns and risk of mortality from cardiovascular disease, cancer, and all causes in a prospective cohort of women. *Circulation.* 2008;118:230-237.

3. Brunner EJ, Mosdol A, Witte DR et al. Dietary patterns and 15-y risks of major coronary events, diabetes, and mortality. *Am J Clin Nutr.* 2008;87:1414-1421.

4. Hu FB, Rimm EB, Stampfer MJ, Ascherio A, Spiegelman D, Willett WC. Prospective study of major dietary patterns and risk of coronary heart disease in men. *Am J Clin Nutr.* 2000;72:912-921.

5. Iqbal R, Anand S, Ounpuu S et al. Dietary patterns and the risk of acute myocardial infarction in 52 countries: results of the INTERHEART study. *Circulation.* 2008;118:1929-1937.

6. Radhika G, Sudha V, Mohan SR, Ganesan A, Mohan V. Association of fruit and vegetable intake with cardiovascular risk factors in urban south Indians. *Br J Nutr.* 2008;99:398-405.

7. Donin AS, Nightingale CM, Owen CG et al. Nutritional composition of the diets of South Asian, black African-Caribbean and white European children in the United Kingdom: the Child Heart and Health Study in England (CHASE). *Br J Nutr.* 2010;104:276-285.

8. Lip GY, Malik I, Luscombe C, McCarry M, Beevers G. Dietary fat purchasing habits in whites, blacks and Asian peoples in England--implications for heart disease prevention. *Int J Cardiol.* 1995;48:287-293.

9. Yagalla MV, Hoerr SL, Song WO, Enas E, Garg A. Relationship of diet, abdominal obesity, and physical activity to plasma lipoprotein levels in Asian Indian physicians residing in the United States. *J Am Diet Assoc.* 1996;96:257-261.

10. Simmons D, Williams R. Dietary practices among Europeans and different South Asian groups in Coventry. *Br J Nutr.* 1997;78:5-14.

11. Saleheen D, Zaidi M, Rasheed A et al. The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia. *Eur J Epidemiol.* 2009;24:329-338.

12. Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med.* 1991;10:1025-1035.

13. Drewnowski A, Popkin BM. The nutrition transition: new trends in the global diet. *Nutr Rev.* 1997;55:31-43.

14. Pingali, P. and United Nations. Westernization of Asian Diets and the transformation of food systems: Implications for research and policy. EAS Working Paper No.04-17.Agricultural and Development Economics Division.The Food and Agriculture Organization of the United Nations . 2004.

15. Singh RB, Niaz MA, Ghosh S et al. Association of trans fatty acids (vegetable ghee) and clarified butter (Indian ghee) intake with higher risk of coronary artery disease in rural and urban populations with low fat consumption. *Int J Cardiol.* 1996;56:289-298.

16. Hakeem R, Thomas J, Badruddin SH. Food habits and nutrient density of diets of Pakistani children living in different urban and rural settings. *J Health Popul Nutr.* 2002;20:255-263.

17. Wandel M, Raberg M, Kumar B, Holmboe-Ottesen G. Changes in food habits after migration among South Asians settled in Oslo: the effect of demographic, socio-economic and integration factors. *Appetite.* 2008;50:376-385.

18. Misra A, Singhal N, Sivakumar B, Bhagat N, Jaiswal A, Khurana L. Nutrition Transition in India: Secular Trends in Dietary Intake and their Relationship to Diet-related Non-communicable Diseases. *J Diabetes.* 2011;10-0407.

19. Popkin BM. The nutrition transition: an overview of world patterns of change. *Nutr Rev.* 2004;62:S140-S143.

20. Speedy A. Animal source foods to improve micronutrient nutrition in developing countries. *Journal of Nutrition.* 2003;4048S-4054S.

Appendix 4: Association of the 9p21.3 locus with risk of first-ever myocardial infarction in Pakistan

[+]Danish Saleheen[1,2], [+]Myriam Alexander[2], Asif Rasheed[1], David Wormser[2], Nicole Soranzo[3,4], Naomi Hammond[3], Adam Butterworth[2], Moazzam Zaidi[1], Philip Haycock[2], Suzannah Bumpstead[3], Simon Potter[3], Hannah Blackburn[3], Emma Gray[3], Emanuele Di Angelantonio[2], Stephen Kaptoge[2], Nabi Shah[1], Maria Samuel[1], Ahmedyar Janjua[1], Nasir Sheikh[2], Shajjia Razi Haider[1], Muhammed Murtaza[1], Usman Ahmad[1], Abdul Hakeem[1], Muhammad Ali Memon[1], Nadeem Hayat Mallick[5], Muhammad Azhar[5], Abdus Samad[6], Syed Zahed Rasheed[6], Ali Raza Gardezi[7], Nazir Ahmed Memon[8], Abdul Ghaffar[8], Fazal-ur-Rehman[9], Khan Shah Zaman[10], Assadullah Kundi[10], Zia Yaqoob[10], Liaquat Ali Cheema[10], Nadeem Qamar[10], Azhar Faruqui[10], Rashid Jooma[11], Jawaid Hassan Niazi[11], Mustafa Hussain[1], Kishore Kumar[1], Asim Saleem[1], Kishwar Kumar[1], Muhammad Salman Daood[1], Fatima Memon[1], Aftab Alam Gul[1], Shahid Abbas[1], Junaid Zafar[1], Faisal Shahid[1], Zehra Memon[1], Shahzad Majeed Bhatti[1], Waleed Kayani[1], Syed Saadat Ali[1], Muhammad Fahim[1], Muhammad Ishaq[6], Philippe Frossard[1], *Panos Deloukas[3], *John Danesh[2] ,

+ These authors have contributed equally to this work; * These authors have contributed equally to this work.

[1]Center for Non-Communicable Diseases (CNCD) Karachi Pakistan, [2]Department of Public Health and Primary Care, University of Cambridge, UK; [3]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK; [4]Department of Twin Research & Genetic Epidemiology, King's College London, St Thomas' Hospital Campus, Lambeth Palace Rd, London, UK; [5]Punjab Institute of Cardiology Lahore Pakistan; [6]Karachi Institute of Heart Diseases, Karachi, Pakistan; [7] Ch. Perwaiz Elahi Institute of Cardiology, Multan, Pakistan; [8]Deewan Mushtaq CCU, Liaquat University of Medical and Health Sciences, Civil Hospital; [9]Red Crescent Institute of Cardiology, Hyderabad, Pakistan; [10]National Institute of Cardiovascular Diseases, Karachi, Pakistan. [11]Jinnah Post-graduate Medical Centre, Karachi, Pakistan

Arterioscle Thromb Vasc Biol. 2010;30:1467-1473.

Correspondence:

Dr. Danish Saleheen
Center for Non-Communicable Diseases (CNCD),
Karachi, Pakistan &
Department of Public Health and Primary Care
University of Cambridge
Strangeways Research Laboratory
Cambridge CB1 8RN, UK
danish.saleheen@cncdpk.com,
danish.saleheen@medschl.cam.ac.uk

Text: 2188 words

Abstract: 233 words

Total number of figures and tables: 3 tables and 2 figures

Summary

We studied variants at the 9p21 locus in a case-control study of acute myocardial infarction (MI) in Pakistan, and did an updated meta-analysis of published studies in people of European ancestry to help contextualize these data. A total of 1851 patients with first-ever confirmed MI and 1903 controls were genotyped for 89 tagging SNPs at 9p21, including the lead variant (rs1333049) identified by the Wellcome Trust Case Control Consortium. Minor allele frequencies and extent of linkage disequilibrium observed in Pakistanis were broadly similar to those seen in Europeans. In the Pakistani study, six variants were associated with MI (P<10-2) in the initial sample set, as well as in an additional 741 cases and 674 controls in whom further genotyping was done for these variants. For Pakistanis, the odds ratio for MI was 1.13 (95% CI: 1.05-1.22; P=2 X 10-3) for each copy of the C allele at rs1333049. By comparison, meta-analysis of studies in Europeans yielded a higher odds ratio of 1.31 (1.26-1.37) for the same variant (P=1x10-3 for heterogeneity). Meta-analyses of 23 variants, in up to 38,250 cases and 84,820 controls, quantified odds ratios for CAD generally yielding higher odds ratios in Europeans than in Pakistanis. This study has provided the first demonstration that variants at the 9p21 locus are significantly associated with MI risk in Pakistanis. Association signals at this locus were, however, considerably weaker in Pakistanis than those previously reported in Europeans.

Background

Variants at the 9p21.3 locus have been established as among the strongest common genetic factors associated with the risk of coronary artery disease (CAD) in people of European continental ancestry.[1-5] These variants are in high linkage disequilibrium (LD) and span a 58 kb region that has multiple neighbouring genes (*CDKN2A*, *CDNK2B* and *MTAP*), without annotating to any single protein coding sequence.[6] An RNA coding gene, *ANRIL*, has been identified that overlaps with the risk locus associated with CAD, suggesting a regulatory role in gene expression.[7] Although associations of variants at 9p21.3 with CAD have been established in several non-European populations (e.g., East Asians), they have not been well-studied in South Asians, populations at high risk of vascular disease.[8]

We report the first large-scale study of variants at the 9p21 locus in relation to risk of acute myocardial infarction (MI) in Pakistan. This study involved 1851 patients with confirmed diagnoses of first-ever myocardial infarction (MI) and 1903 controls from the Pakistan Risk of Myocardial Infarction Study[9] (PROMIS). Genotyping involved 89 tagging single nucleotide polymorphisms (SNPs) at the 9p21.3 locus, including the lead variant (rs1333049) identified by the Wellcome Trust Case Control Consortium in association with MI.[10,11] To place our findings in context, we also report a literature-based meta-analysis of relevant studies, encompassing information on 23 variants at the 9p21 locus in up to 38,250 CAD cases and 84,820 controls. The current meta-analysis substantially updates a previous relevant review[12], involving data from an additional 82,117 participants and 20 additional variants.

Methods

Study design This paper follows the reporting recommendations of STREGA.[13] PROMIS is a case-control study of acute first-ever MI in urban Pakistan[9]. MI cases had: (1) symptoms within 24 hours of hospital presentation; (2) typical ECG characteristics (e.g., 1 mm or more ST elevation in any two or more contiguous limb leads, new onset left bundle branch block); and (3) a positive troponin test (>1ng/ml). Controls were individuals without a history of cardiovascular disease frequency-matched to cases by sex and age in 5 year age bands, and concurrently identified in the same hospitals as the index cases by virtue of being: (1) visitors of patients attending the outpatient department; (2) patients attending the outpatient department for routine non-cardiac complaints, or (3) non-blood related visitors of index MI cases. Controls with recent illnesses or infections were excluded. A locally-piloted and -validated epidemiological questionnaire was administered to participants by medically qualified research officers that sought >200 items of information in relation to: ethnicity, demographic characteristics, lifestyle factors (e.g., tobacco and alcohol consumption, dietary intake and physical activity); personal and family history of cardiovascular disease; and medication usage. Non-fasting blood samples were drawn from each participant and centrifuged within 45 minutes of venepuncture. Samples were stored at -80 ºC. The study has received approval from

relevant institutional review boards in each recruitment centre and the Center for Non-Communicable Diseases (CNCD) Karachi, Pakistan. Informed consent was obtained from all the participants (including consent to use samples in genetic, biochemical and other analyses).

Genotyping DNA was extracted from leucocytes using a reference phenol-chloroform protocol.[9] Genotyping was performed at the Wellcome Trust Sanger Institute, Hinxton, UK. To minimise any systematic biases arising from plate- or batch-specific genotyping error, genotyping plates contained a mixture of cases and controls, including negative and positive controls. 1851 cases and 1903 controls were genotyped using version 1 of the IBC array of about 2000 candidate genes. 169 SNPs tagged the 9p21.3 locus at r2 > 0.8 and were available in the current analyses. This array employed a cosmopolitan tagging approach, using information from the Hap Map Caucasian (CEU), East Asian (Han Chinese and Japanese), and African (Yoruba) populations[14]. SNPs were excluded if: (1) the call rate was <95% (2 SNPs); (2) there was evidence of departure from Hardy-Weinberg Equilibrium in controls at a P-value of <0.05 (17 SNPs); or (3) the minor allele frequency (MAF) was <1% (63 SNPs), with most of such omissions due to SNPs relevant for Africans being monomorphic in Pakistanis. Seven individuals were excluded either because self-reported gender did not match chromosomal sex status, evidence of cryptic relatedness or more than 2% missing genotypes. After such quality control, 89 SNPs remained. The six SNPs most significantly associated with MI risk were genotyped in a further 741 cases and 674 controls (iPLEX: Sequenom).

Statistical analysis Analyses involved PLINK 1.06[15] and STATA version 10.0 (StataCorp). Assuming an additive model, associations with each SNP were tested fitting a logistic regression model adjusted for age, sex and the first two principal components (calculated using all 45,000 SNPs genotyped in the array, as described previously).[16],[17] Effect-modification was investigated by tests of interaction in fully adjusted models. LD was assessed using Haploview, with blocks graphically identified from the LD intensity expressed in D'.[18] Haplotype association analyses were performed with the THESIAS software implementing the Stochastic-EM algorithm,[19] enabling simultaneous estimation of haplotypes frequencies and their effects on MI, again assuming additive effects. Using a parsimonious approach, the most informative tagging SNPs were chosen for haplotype analyses that accounted for at least 85% of the haplotypic block variability.[20] To compare LD patterns in Pakistanis with those in other ethnic groups, data were downloaded from the HapMap website for Caucasians, East Asians (Han Chinese and Japanese), and Africans (Yoruba), and drawn using Haploview.[21]

Systematic review We sought studies reporting on associations between variants at the 9p21 locus and risk of CAD before January 2010 (**Figure A4.1**). Electronic searches, not limited to the English language, used the MEDLINE database and involved search terms related to the locus (e.g., chromosome 9, *CDKN2A*, *CDNK2B*, *MTAP*, *ANRIL*, rs-numbers of variants previously reported) and coronary artery disease (eg, coronary heart disease, myocardial

infarction, atherosclerosis, coronary stenosis). These searches were supplemented by scanning reference lists, hand searching relevant journals, and discussion with authors. Two investigators independently extracted the following: genotype frequencies; unadjusted additive odds ratios; ethnicity; geographical location; CAD definition; study type (e.g., genomewide association study, candidate replication study); genotyping platform; study design; and source of controls. Additive odds ratios were computed for each study using genotype counts and were compared with reported odds ratios, when available. For prospective cohorts, hazard ratios were assumed to approximate odds ratios. Summary odds ratios and 95% confidence interval (CI) for each SNP were calculated fitting a random-effects model that included between-study heterogeneity. Heterogeneity was assessed by the I2 statistic[22] and the Q statistic, and investigated by pre-specified study-level characteristics, notably: study size, case definition, study design, and genotyping platform used.

Results

Analyses in Pakistanis As would be expected, baseline levels of conventional risk factors were significantly higher in MI cases than controls (**Table A4.1**). Genetic similarity did not correlate strongly with self-reported ethnicity among the 8 ethnic and linguistic groups studied in PROMIS (**Figure A4.2**) and LD patterns at the 9p21 locus were similar among the four major Pakistani sub-ethnicities (**Figure A4.3**). Of the 89 relevant SNPs assessed, 6 were associated with MI at p<10-2 (**Figure A4.4** & **Table A4.2**), including the lead variant (rs1333049) identified by Wellcome Trust Case Control Consortium.[23] Odds ratios for MI with each of these 6 SNPs were about 1.13 (p<10-3 for each; **Table A4.3**). In the case of rs1333049, the odds ratio for MI was 1.12 (95% CI: 1.04-1.20; P=2x10-3) per C allele. Odds ratios for MI were not significantly different under a range of circumstances (**Figures A4.5**), though there was limited power to evaluate potential effect-modification (e.g., it was not possible to confirm or refute possible differences noted in men and women and between ethnic groups for some SNPs).

There were 5 distinct LD blocks at the 9p21 locus in Pakistanis, each having strong intra-block LD (**Figure A4.6**). All 6 variants and haplotypes associated with MI were located within block 3. Using a parsimonious model, haplotypic associations observed in block 3 could be explained by only two tagging SNPs (i.e., rs1412832 and rs1333049; **Table A4.4**). Haplotype analyses involving these 2 tagging SNPs generated 4 haplotypes, each with a frequency greater than 2%, and the global test of association with MI was significant (p=0.028). The odds ratio for MI with the AG haplotype was 0.85 (0.76-0.95; P=0.003) compared with the most frequent AC haplotype.

Three SNPs (rs7865618, rs1292136, rs7044859) located in the genomic region of block 2 had previously been found to modulate the effect of SNP rs1333049 on the risk of MI in Europeans, with the two most frequent haplotypes being in "yin yang".[24] To evaluate these findings in Pakistanis, we analysed proxy SNPs (rs518394, rs10965212 and rs7049105) based on the

HapMap CEU data (0.69≤r2≤0.90). We did not find any significant evidence for a heterogeneous effect of the rs1333049-C allele according to the haplotypic backgrounds generated by these three proxy SNPs, or SNPs tagging block 2 (data available upon request).

Meta-analysis We identified 26 relevant studies of 23 variants at the 9p21 locus, comprising 38,250 cases and 84,820 controls (**Table A4.5**). 19 studies were based in Europe or the USA[25-29,30-43], 6 in East Asia[44-49] and 1 in North Africa[50]. 5 were prospective in design[51,52,53,54,55], 2 were nested case-control studies[56,57] and 19 were retrospective case-control studies. 7 studies included only MI patients, while 20 studies included patients with MI or coronary stenosis (defined by intervention procedures). Of the 23 variants studied, 21 were associated with CAD risk in Europeans and 6 with CAD risk in East Asians (**Figure A4.7** & **Figures A4.8a-w**). For most variants, there was null to moderate heterogeneity in ethnic-specific combined odds ratios (e.g., I2 values typically ranged from 0% to 20%, with 6 SNPs with I2>50%). For Europeans and East Asians, odds ratios for CAD were 1.31 (1.26-1.37; P<10-35) and 1.25 (1.13–1.39; P<10-5), respectively, per C allele of rs1333049. Excluding the original report, the corresponding odds ratio was 1.29 (1.23-1.36) in Europeans. Of the 23 variants studied, data were available on 12 in PROMIS (**Figure A4.7**). Odds ratios for MI with variants at this locus were generally lower in Pakistanis than Europeans (e.g., heterogeneity P=1x10-3 for rs1333049), although the allelic frequencies in Pakistanis and Europeans were similar (**Figure A4.7**). HapMap suggested a high degree of LD between rs1333049 and other variants studied at the 9p21 locus, with a similar pattern of LD in Europeans and Pakistanis (**Figure A4.9**).

Discussion

In the first large-scale genetic study of MI in Pakistan, we investigated 89 SNPs spanning 350 kb at the 9p21.3 locus and identified 6 SNPs significantly associated with risk of MI. In an updated literature-based meta-analysis of 38,250 CAD cases and 84,820 controls, we confirmed associations of 21 variants at this locus with CAD in Europeans (as well as confirming associations of 6 variants at this locus with CAD in East Asians). We observed that odds ratios for MI with variants at the 9p21.3 locus were generally lower (typically by about half as much) in Pakistanis than those reported in studies of people of European continental ancestry. As discussed below, however, further studies are needed to determine whether such differences are mainly related to ethnicity or study design features or both.

The LD structure in PROMIS was similar to the one previously observed in Europeans (HapMap CEU) and somewhat stronger than in East Asians and Africans. The 6 SNPs associated with MI in PROMIS were located within one block that had high intra-block LD spanning 53kb. Two haplotype tagging SNPs (rs1333049 and rs1412832) were sufficient to explain the observed association with MI in Pakistanis. This block partially overlapped with the region of association previously identified in Europeans.[58] Although we did not observe in Pakistanis the previously reported significant modulation in Europeans of the effect of

rs1333049 on MI by either proxy or tag SNPs of haplotype block 2[59], this null finding may have been due to limited statistical power and/or inadequate proxy markers in the Pakistan study.

For Pakistanis, Europeans, and East Asians, analyses of the present data and HapMap indicated that CAD-related variants at the 9p21.3 were in high LD with one another and localized within a region that is devoid of a single protein coding sequence.[60] This locus coincides with the *ANRIL* gene, a recently discovered antisense non-coding RNA postulated to enhance gene expression.[61,62,63] There are multiple neighbouring protein coding genes in this region, including *CDKN2A*, *CDKN2B* and *MTAP*, proposed to play important roles in cell-cycle progression, cellular proliferation, apoptosis and cellular senescence.[64,65,66]

Although allelic frequencies for most variants and LD pattern at the 9p21.3 locus were similar in Pakistanis and Europeans, odds ratios for CAD with most variants at this locus were lower in Pakistanis. There are previous reports of cardiometabolic diseases having differential genetic effects by ethnicity, such as stronger associations of a *LTA4H* haplotype with MI in Africans than Europeans,[67] and *TCF7L2* being relevant to type 2 diabetes in West Africans but not East Asians.[68]

Nevertheless, further study is needed to determine whether the apparently different odds ratios for CAD with variants at the 9p21.3 locus noted in this study of Pakistanis versus those of Europeans are due mainly to non-ethnic factors, such as differences in case definitions, age of disease onset, and epidemiological design. In the absence of ethnically-mediated differences in disease susceptibility, however, the present Pakistani study might have been expected to yield higher odds ratios than most previous studies in Europeans owing to this Pakistani study's presumed enrichment for genetic signals (e.g., through use of strict phenotyping and inclusion of early-onset first-ever MI cases with high degrees of familial clustering). Such considerations, therefore, reinforce the need for further studies of MI in Pakistanis to help validate and discover its population-relevant genetic determinants.

Conclusion

This study has provided the first demonstration that variants at the 9p21 locus are significantly associated with MI risk in Pakistanis. Association signals at this locus were, however, considerably weaker in Pakistanis than those previously reported in Europeans.

Acknowledgements

We would like to acknowledge the contributions of the following individuals:

Epidemiological fieldwork in Pakistan: Zeeshan Ozair, Fahad Shuja, Mustafa Qadir Hameed, Imad Hussain, Hamza Khalid, Ali Memon, Kamran Shahid, Ali Kazmi, Sana Nasim, Muhammad Ahsan Javed, Zahir Hussain, Kanwal Aamir, Mazhar Khan, Muhammad Zafar, Faisal Majeed, Madiha Ishaq, Turkey Hussain Marmoos, Faud Khurshid, Farhat Abdul Muntaqim, Sarosh Fatima, Rehan Ahmed, Muhammad Nabeel, , Syed Shazad Hussain, Muhammad Zuhair Yusuf, Nadeem Baig, Madad Ali Ujjan, Parveen Sultan, Asghar Ali, Ayaz Ali, Mir Alam, Hassan Zaib, Abdul Ghafoor, Saeed Ahmed, Muhammad Riazuddin, Muhammad Waqar Khan, Muhammad Irshad Javed, Jabir Furqan, Abdul Ghaffar, Muhammad Shahid, Tanveer Baig Mirza, Muhammad Naeem, Afzal Hussain, Abdul Hakeem, Zahid Hussain, Tanveer Abbas, Muhammad Khurram Shahzad, Iqbal Ali, Muhammad Imran Nisar, Altaf Hussain, Muhammad Shazad, Mehmood Jafree and Ayeesha Kamal.

Laboratory assays: Sadaf Raza, Naeem Khan, Asad Ali Shah, Sobia Naz, Farina Hanif, Shaheen Khanum, Aisha Nazir, Aisha Sultana, Mehwish Jabar, Zahid Hussain, Madiha Yameen, Nadir Khan, Inosh Hasan.

Data management: Sarfaraz Sher Ali, Touqeer Ahmed, Syed Waqas Ahmed, Azfar Hussain, Matthew Walker and Sarah Watson

We are grateful to Dr. David Tregouet for providing help in the haplotype analyses.

Administration: Kashif Saleheen, Hannah Sneath and Sarah Drummond.

Source of funding

Epidemiological fieldwork in PROMIS has been supported by unrestricted grants to investigators at the University of Cambridge. Genotyping for this study was funded by the Wellcome Trust. The British Heart Foundation has supported some biochemical assays. The Yousef Jameel Foundation supports Dr Saleheen. The Marie Curie Bloodomics 2 scholarship supports Myriam Alexander. The cardiovascular disease epidemiology group of Professor Danesh is underpinned by programme grants from the British Heart Foundation, the UK Medical Research Council and the Wellcome Trust.

Disclosure

None.

Figure A4.1: Literature search strategy used in the current meta-analyses

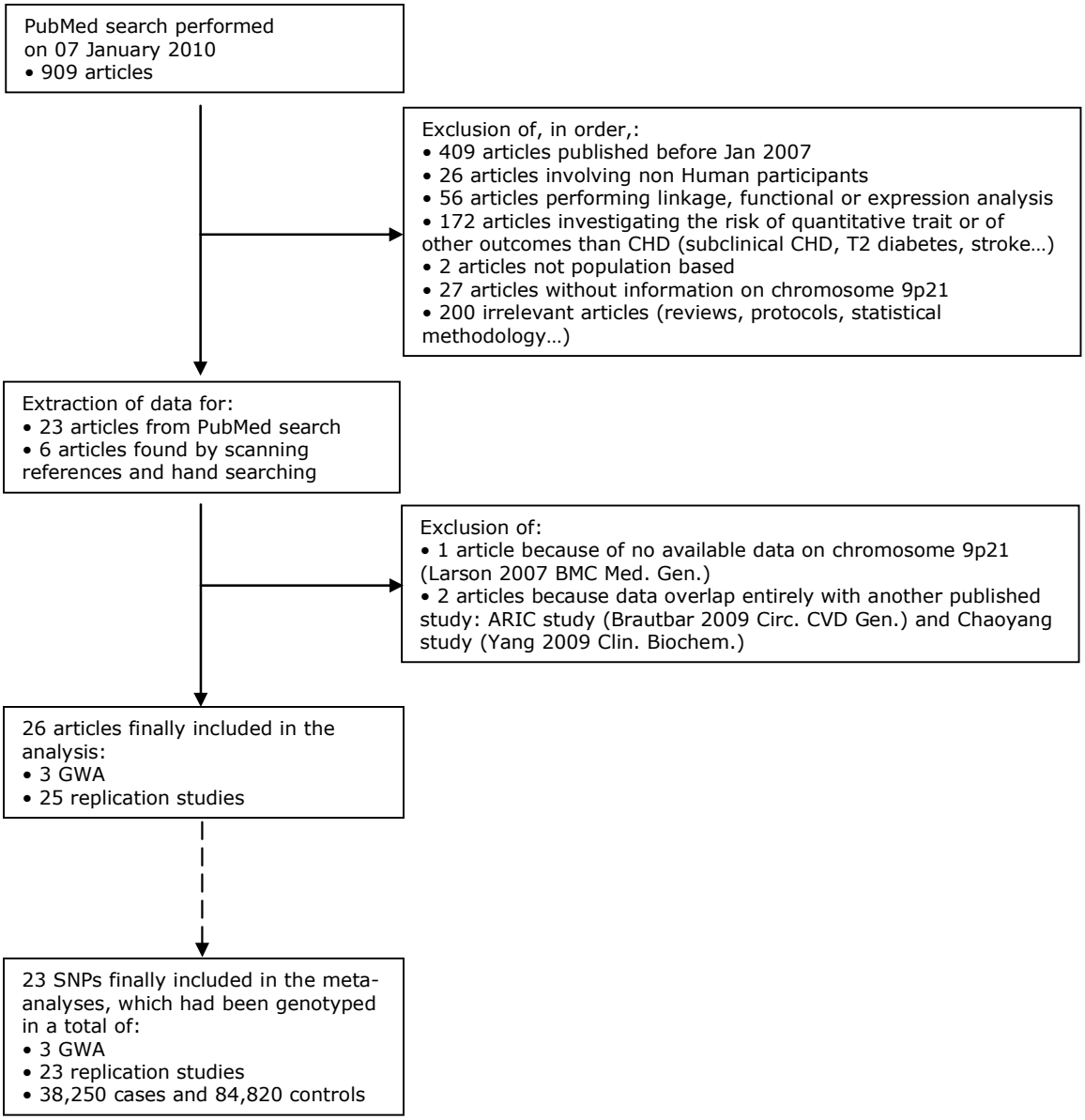
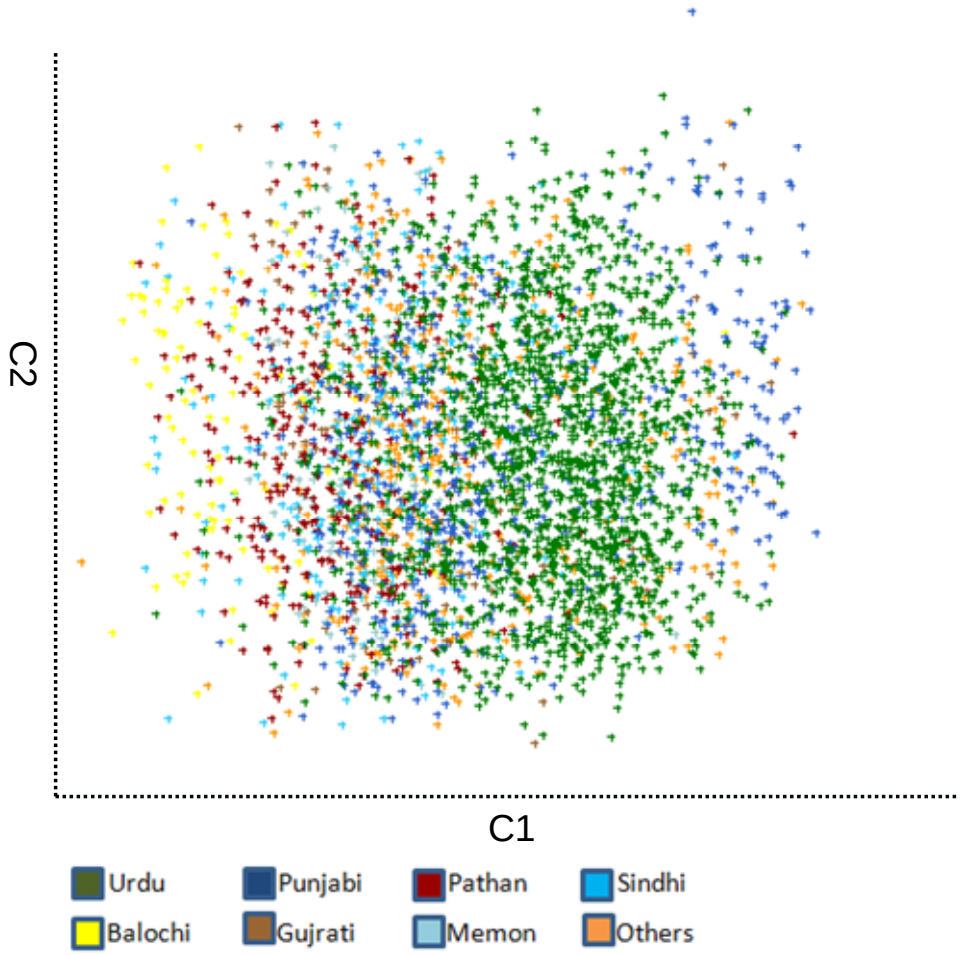
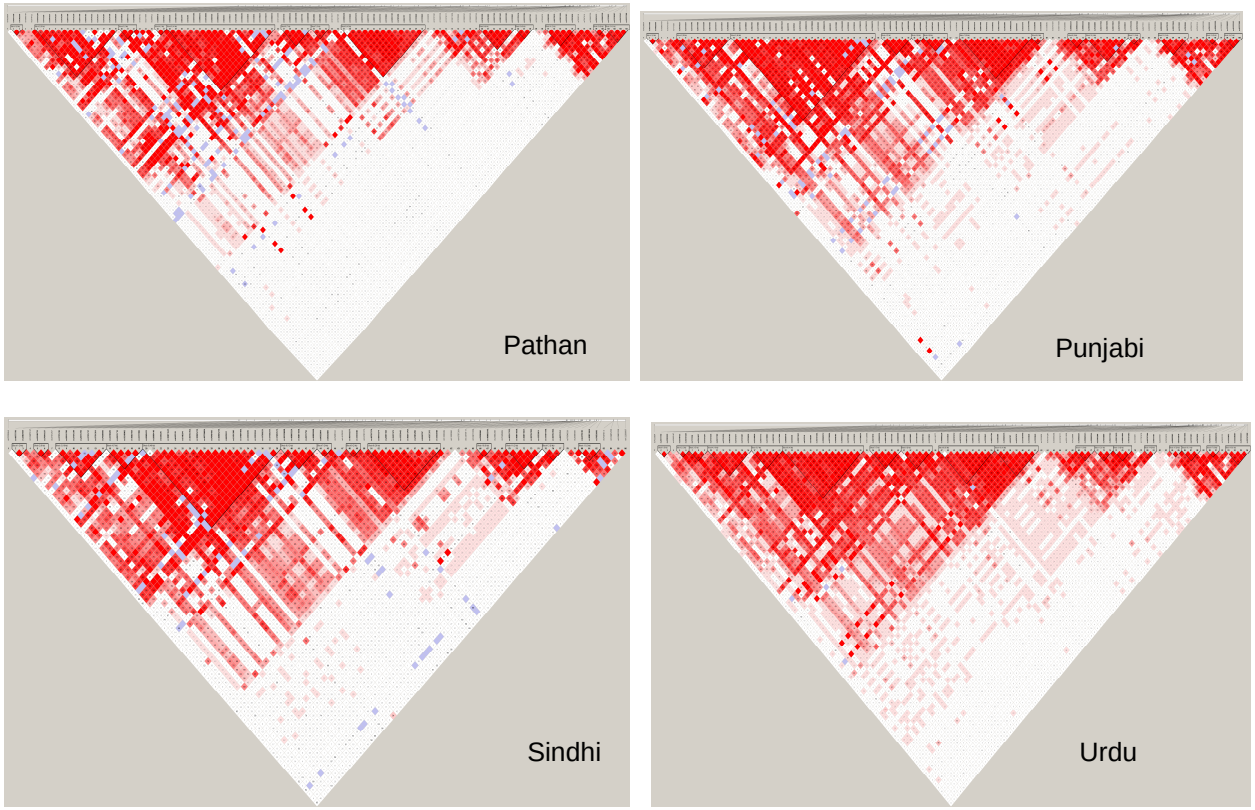


Figure A4.2: Scatter plot of the first two principal components and self reported ethnicities in PROMIS control participants



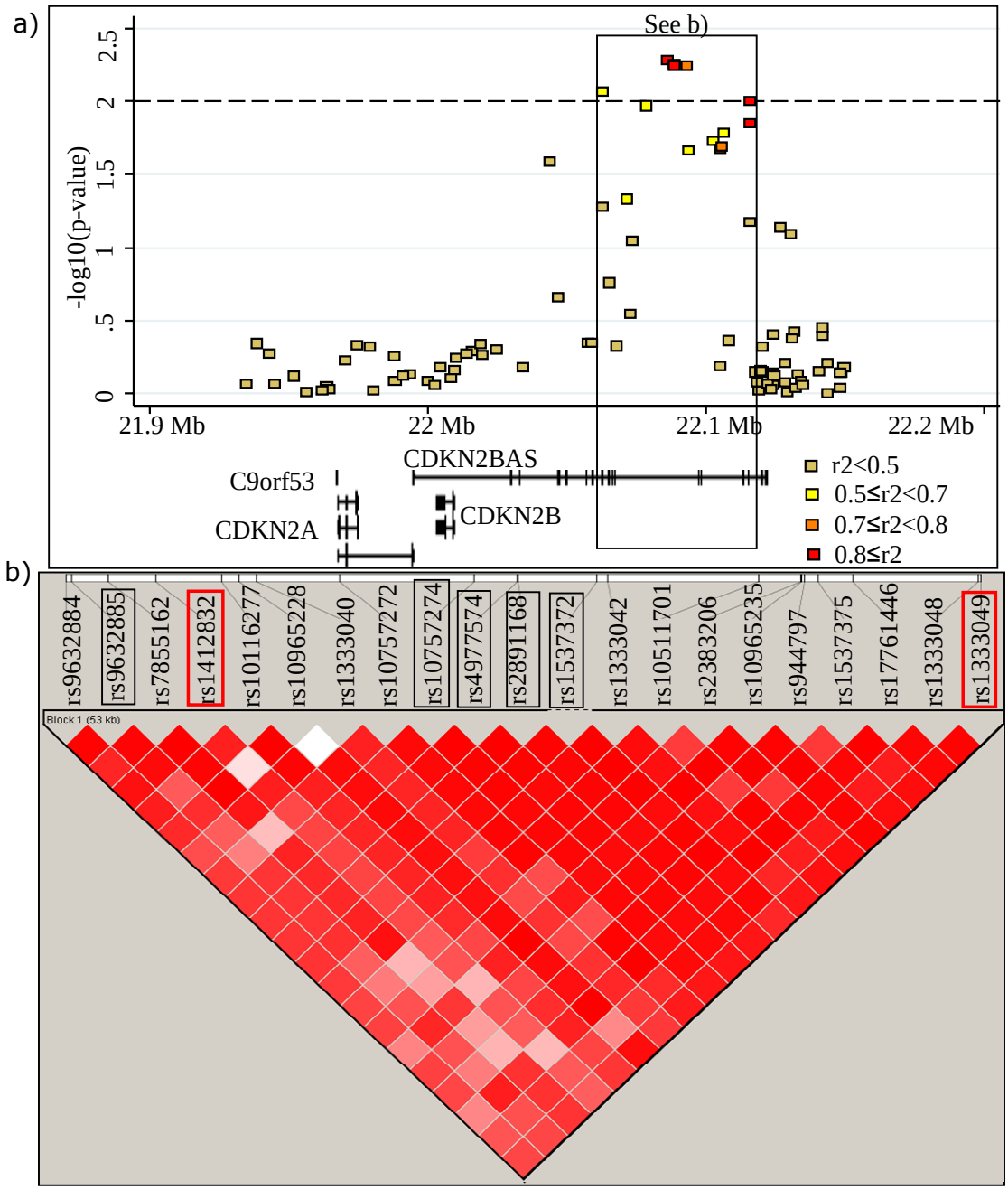
Scatter plot of the first two principal components identified by principal component analysis of the identity-by-state matrix in PROMIS IBC data on 1851 cases and 1903 controls genotyped on 45,000 SNPs. The colours of points refer to self reported-ethnicities in PROMIS participants. C1 and C2 axis represent to the first and second principal components.

Figure A4.3: Linkage disequilibrium patterns in four major Pakistani ethnic groups



There were 302 Pathan, 588 Punjabi, 280 Sindhi and 17510 Urdu. Regions of strong LD are in red, of moderate LD in light red, of very moderate LD in blue and of no LD in white.

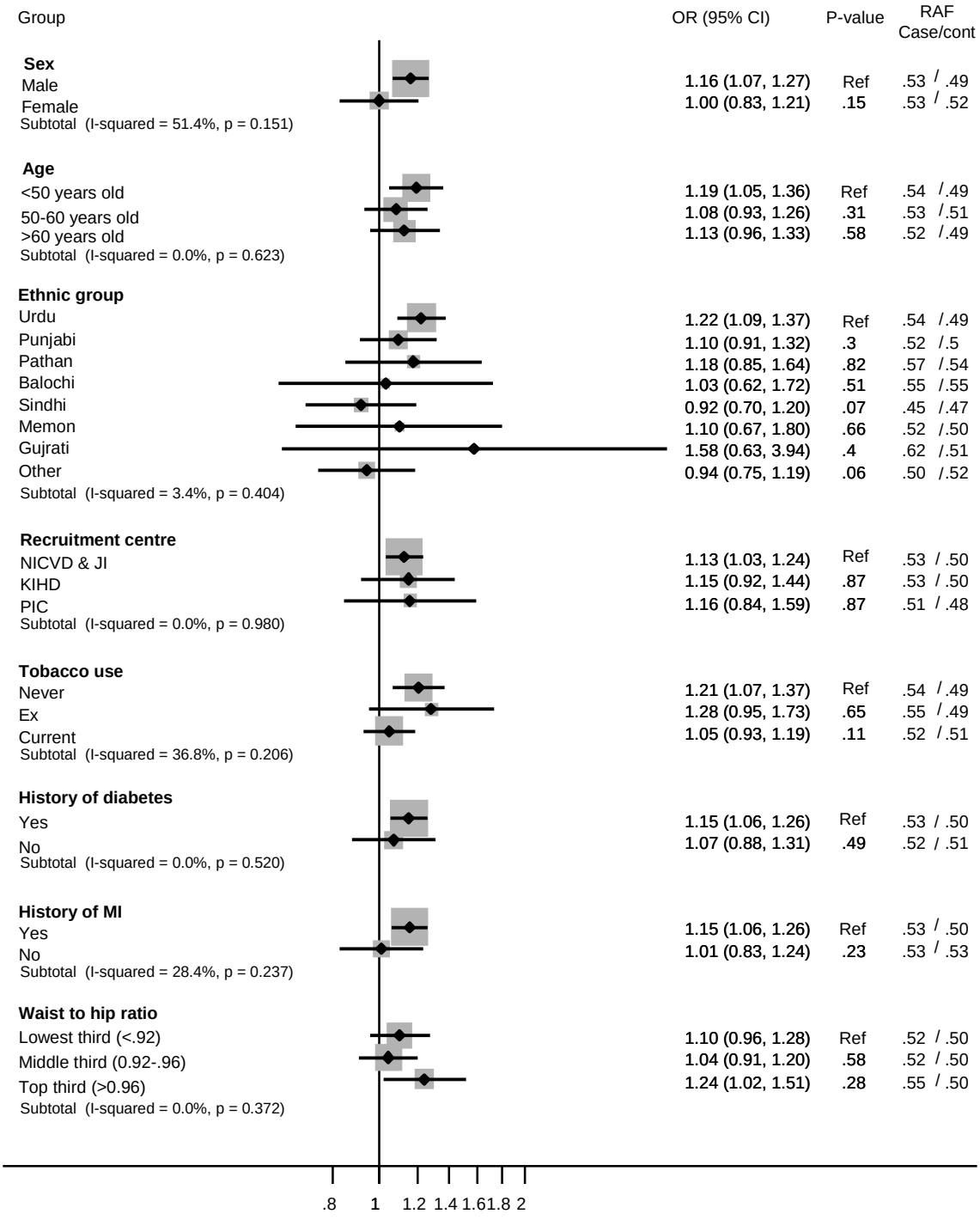
Figure A4.4: Association of variants at 9p21.3 locus with myocardial infarction in PROMIS participants



a) Regional plot of association for all the 89 variants at 9p21.3 locus genotyped in PROMIS participants (1851 cases and 1903 controls) based on r^2 values between SNPs and rs1333049. The dotted line represents the nominal threshold of significance (P-value<10^{-2}). The association between SNPs and MI was tested fitting an additive model adjusting for age, sex and the first two principal components. The lowest part of the panel represents the exons-introns structure of annotated genes. Gene information was downloaded from www.ucsc.edu. *CDKN2BAS* is also called *ANRIL*.

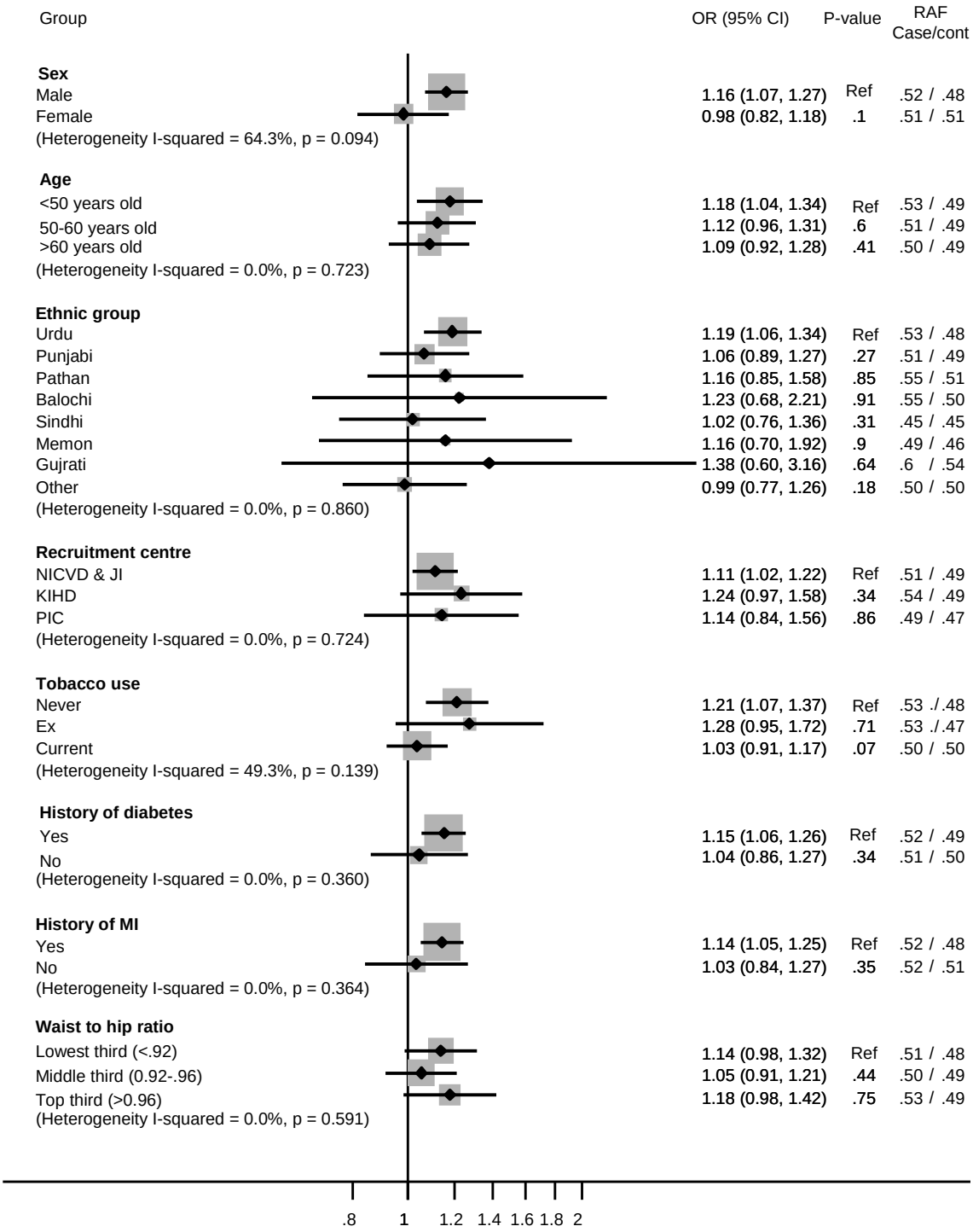
b) LD plot in PROMIS participants for the region highlighted in panel (a). SNPs associated with MI are highlighted in black boxes and the tagging SNPs selected for haplotype analyses are in red boxes.

Figure A4.5a: Effect modification of variant rs10757274 by different factors in PROMIS participants



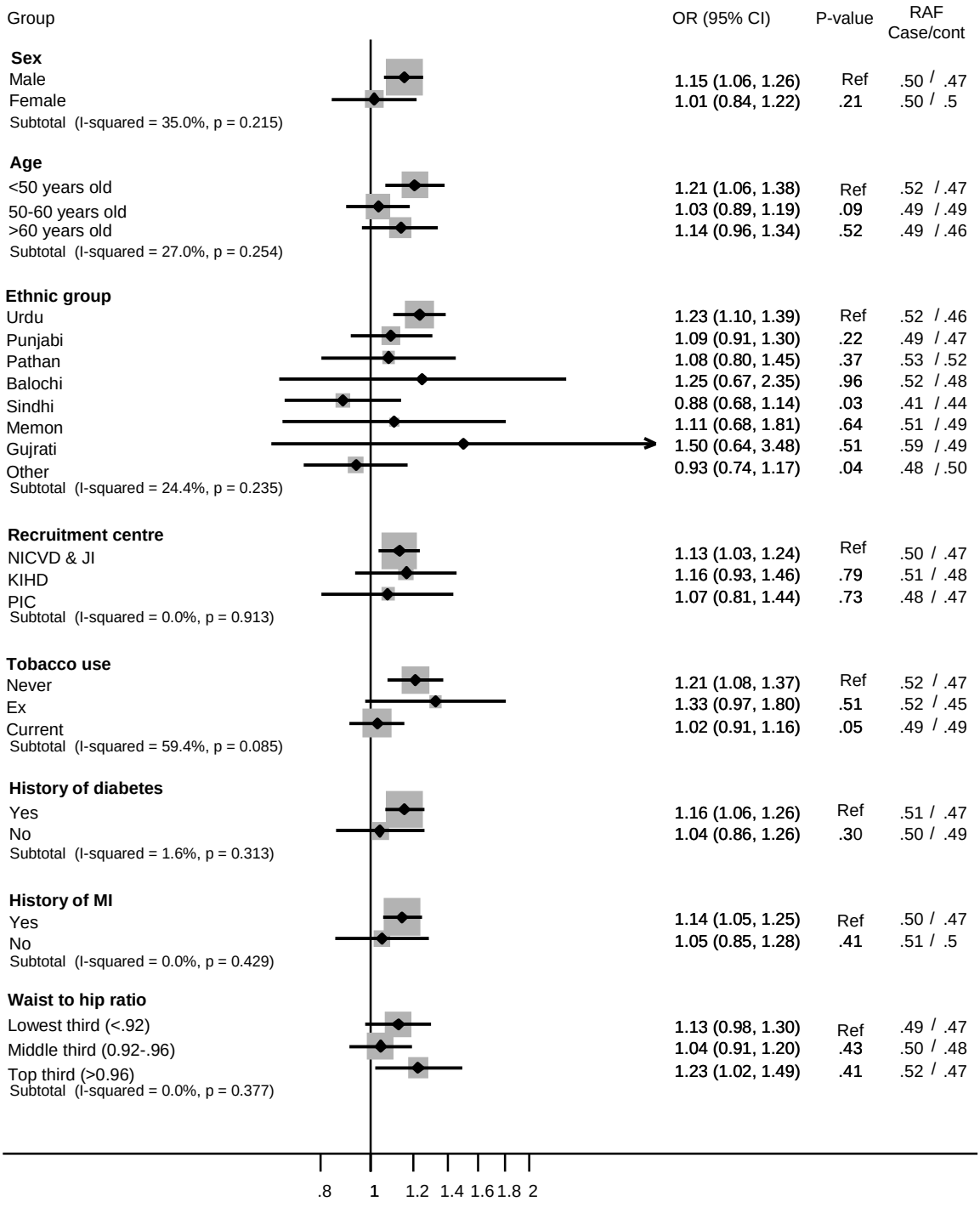
Ref: reference group; P-value: P-value of difference from the reference group; Interaction between each variable and the genotype was assessed by introducing an interaction term in the additive model adjusting for age, gender and self reported ethnicity. For variables with more than two categories, an overall test of heterogeneity is reported below the estimates for all subgroups of this variable.

Figure A4.5b: Effect modification of variant rs1333049 by different factors in PROMIS participants



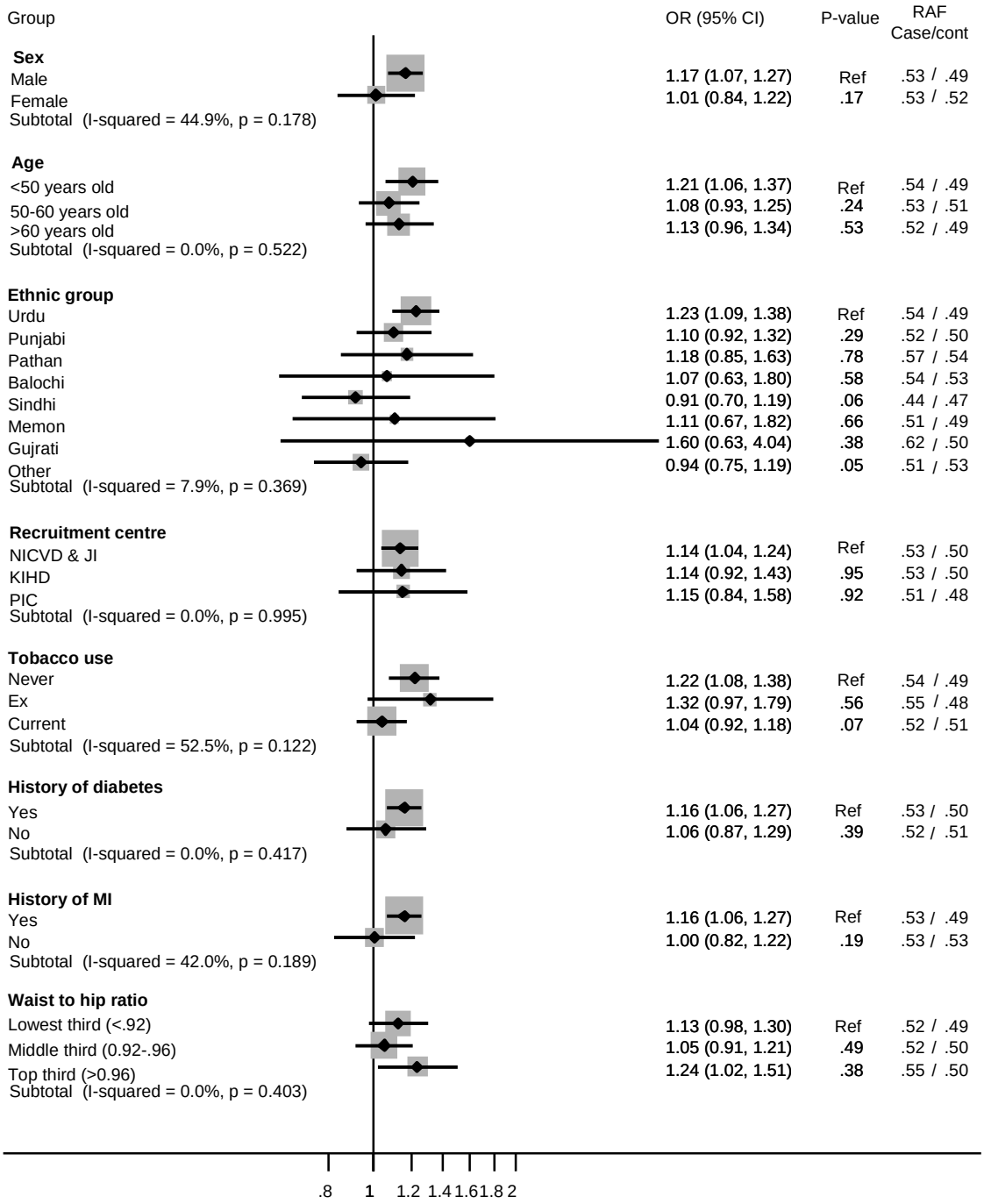
Ref: reference group; P-value: P-value of difference from the reference group; Interaction between each variable and the genotype was assessed by introducing an interaction term in the additive model adjusting for age, gender and self reported ethnicity. For variables with more than two categories, an overall test of heterogeneity is reported below the estimates for all subgroups of this variable.

Figure A4.5c: Effect modification of variant rs1537372 by different factors in PROMIS participants



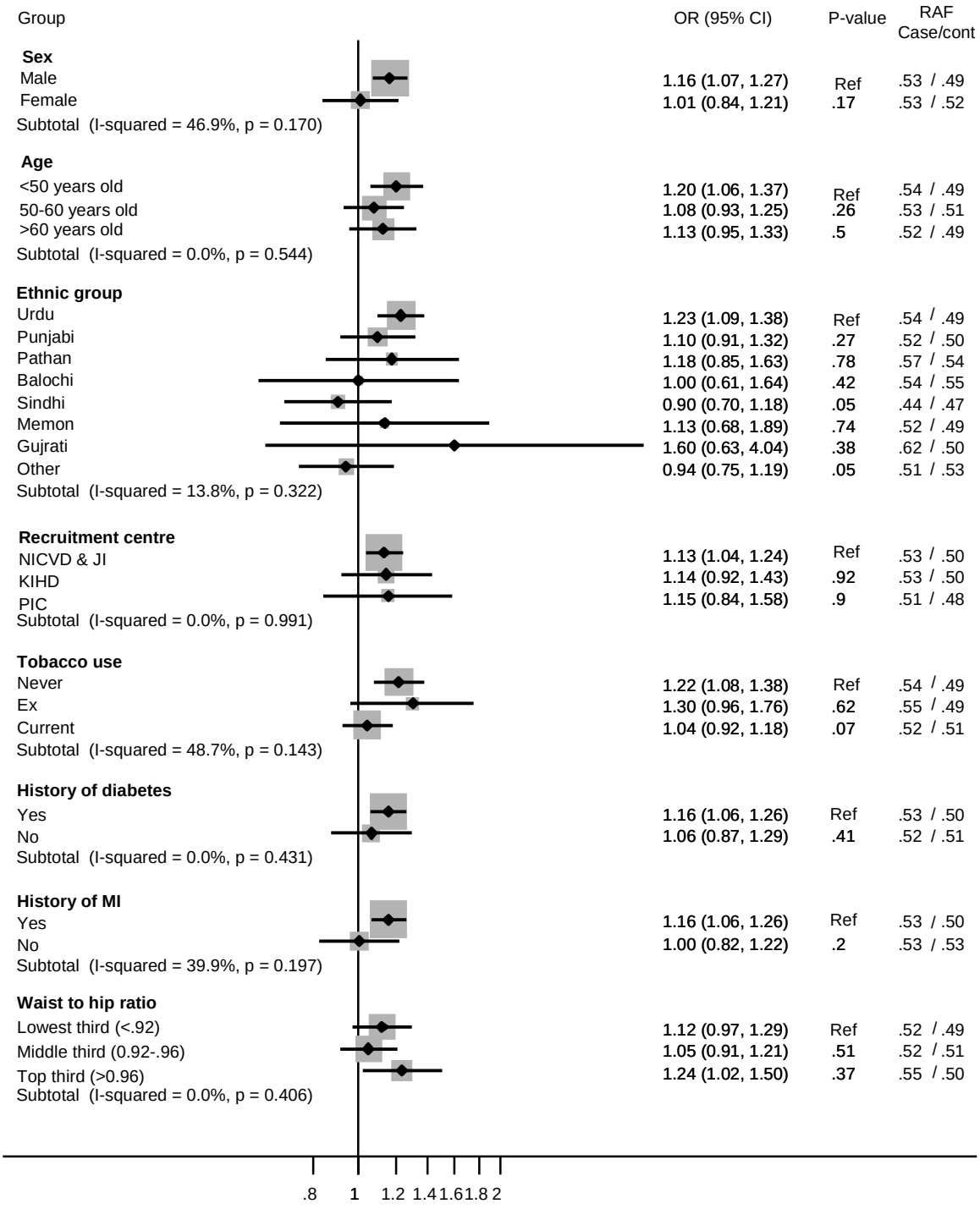
Ref: reference group; P-value: P-value of difference from the reference group; Interaction between each variable and the genotype was assessed by introducing an interaction term in the additive model adjusting for age, gender and self reported ethnicity. For variables with more than two categories, an overall test of heterogeneity is reported below the estimates for all subgroups of this variable.

Figure A4.5d: Effect modification of variant rs2891168 by different factors in PROMIS participants



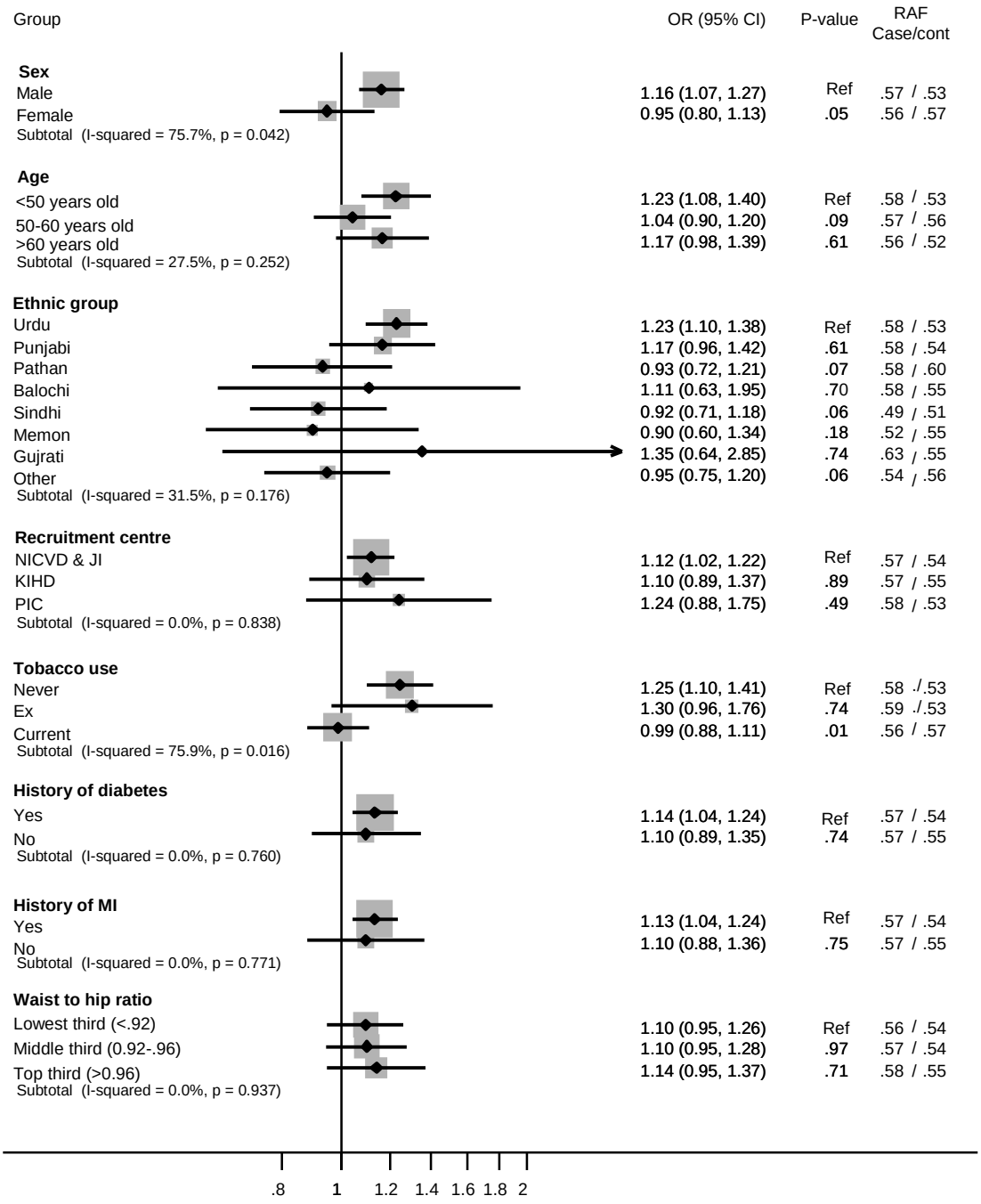
Ref: reference group; P-value: P-value of difference from the reference group; Interaction between each variable and the genotype was assessed by introducing an interaction term in the additive model adjusting for age, gender and self reported ethnicity. For variables with more than two categories, an overall test of heterogeneity is reported below the estimates for all subgroups of this variable.

Figure A4.5e: Effect modification of variant rs4977574 by different factors in PROMIS participants



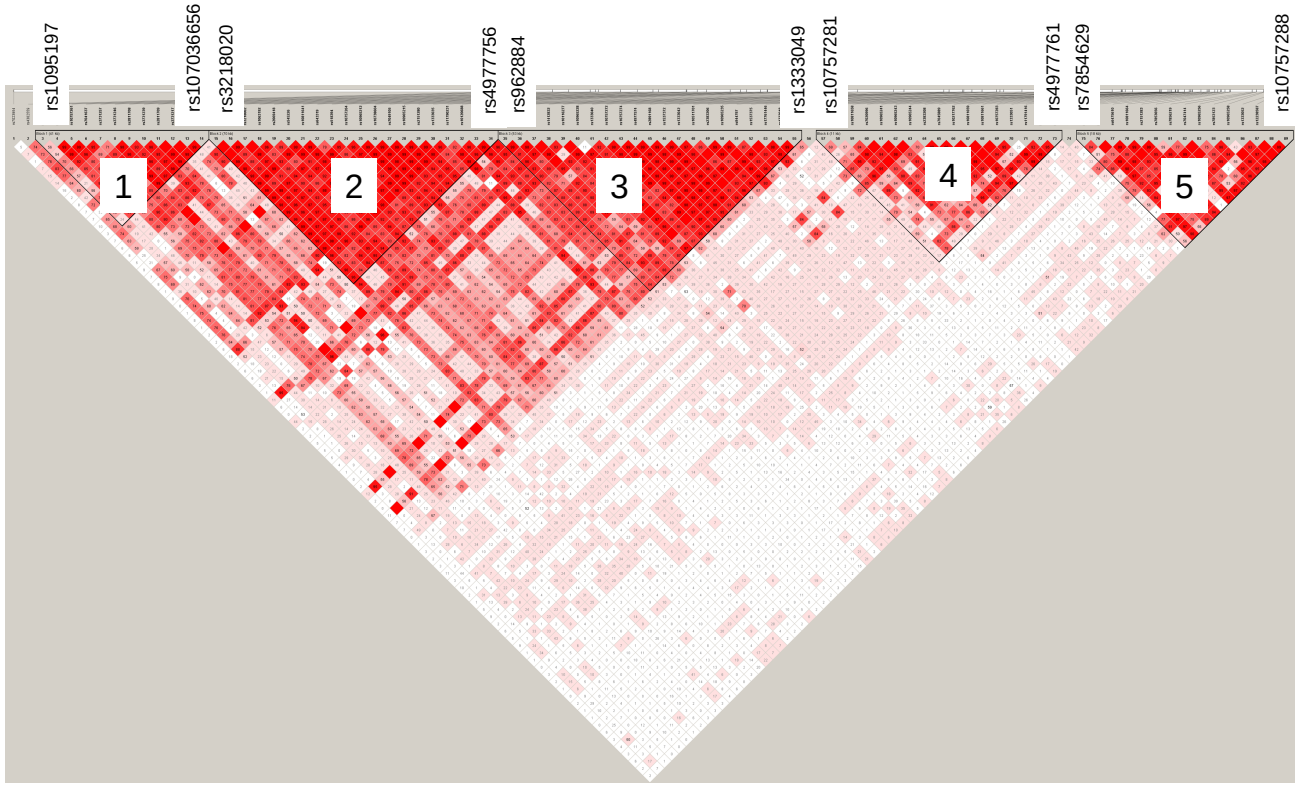
Ref: reference group; P-value: P-value of difference from the reference group; Interaction between each variable and the genotype was assessed by introducing an interaction term in the additive model adjusting for age, gender and self reported ethnicity. For variables with more than two categories, an overall test of heterogeneity is reported below the estimates for all subgroups of this variable.

Figure A4.5f: Effect modification of variant rs9632885 by different factors in PROMIS participants



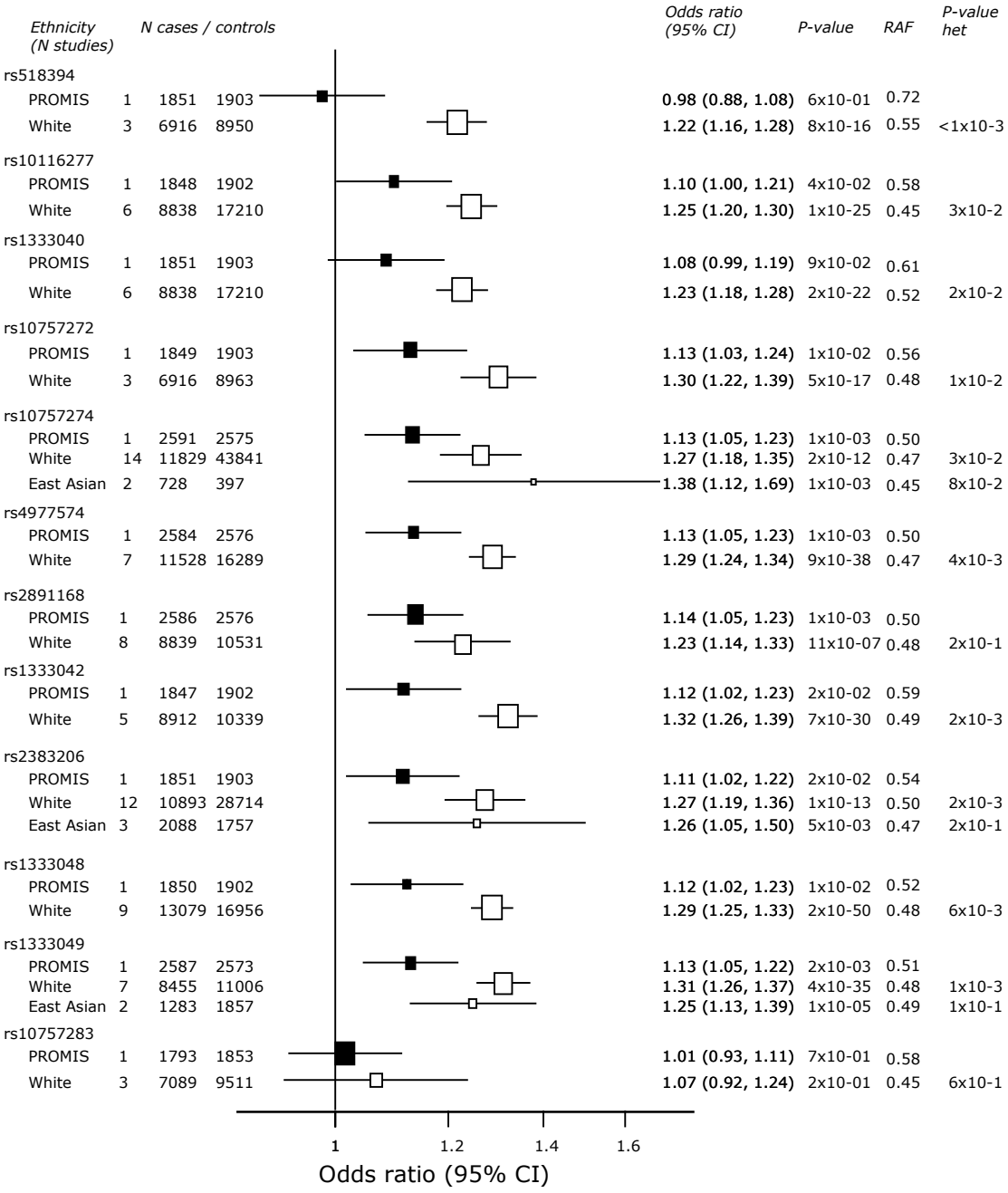
Ref: reference group; P-value: P-value of difference from the reference group; Interaction between each variable and the genotype was assessed by introducing an interaction term in the additive model adjusting for age, gender and self reported ethnicity. For variables with more than two categories, an overall test of heterogeneity is reported below the estimates for all subgroups of this variable.

Figure A4.6: Linkage disequilibrium in Pakistanis



Block 1 extended from rs1095197 to rs107036656, block 2 from rs3218020 to rs4977756, block 3 from rs9632884 to rs1333049, block 4 from rs10757281 to rs4977761 and block 5 from rs7854629 to rs10757288. The number of selected tag SNPs were 6 for block 1 (rs10965197 , rs1077261, rs7041637, rs3731246, rs2811708, rs3731239), 5 for block 2 (rs3218002, rs2069418, rs10738604, rs11790231, rs10965224), 8 for block 3 (rs9632885, rs1412832, rs10965228, rs1333040, rs10757274, rs1537372, rs1333042 and rs1333049), 9 for block 4 (rs10757281, rs10965241, rs10965243, rs10965244, rs2383208, rs7045889, rs10217762, rs10811661, rs11791416, rs4977761) and 6 for block 5 (rs7854629, rs2065505, rs215283, rs7856219, rs10965256, rs7853123). Colours of the blocks are blue or bright red if D'=1 and white or shades of red otherwise.

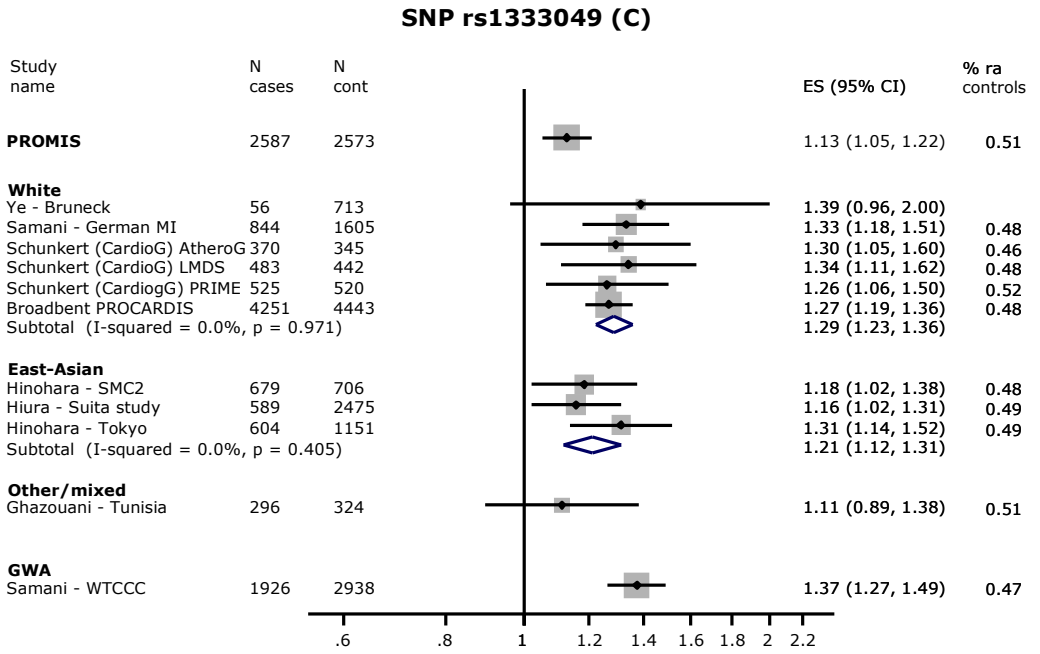
Figure A4.7: Association with MI in Pakistanis, Europeans and East Asians for 12 variants genotyped in PROMIS



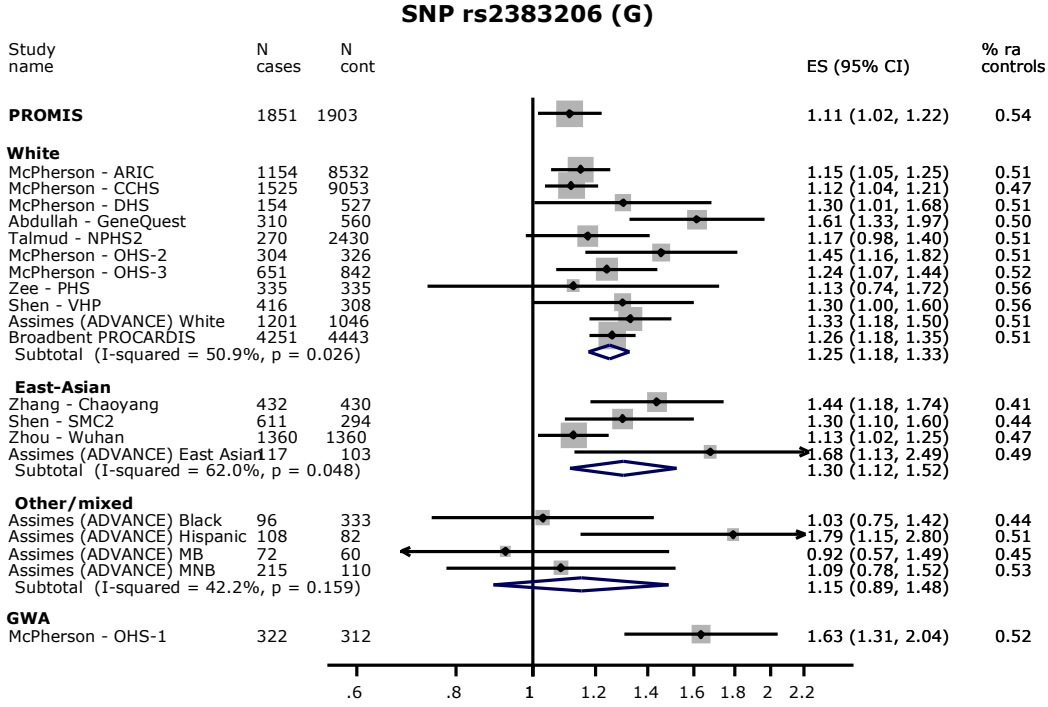
P-het: P-value for heterogeneity; RAF: Risk allele frequency in controls (in %). Odds ratios in PROMIS were computed using an additive model adjusted for age, sex and the first two principal components. Analyses for rs1333409, rs10757274, rs4977574 and rs2891168 are adjusted for age, gender and self-reported ethnicity. Effect estimates in Europeans and East Asians are derived from literature based meta-analyses The P-value heterogeneity corresponds to a heterogeneity test comparing effect estimates in the different ethnic groups. Individual plots for each meta-analysis are presented in efigures 6a-w.

Figures A4.8: Meta-analyses of the 23 SNPs previously published in the literature

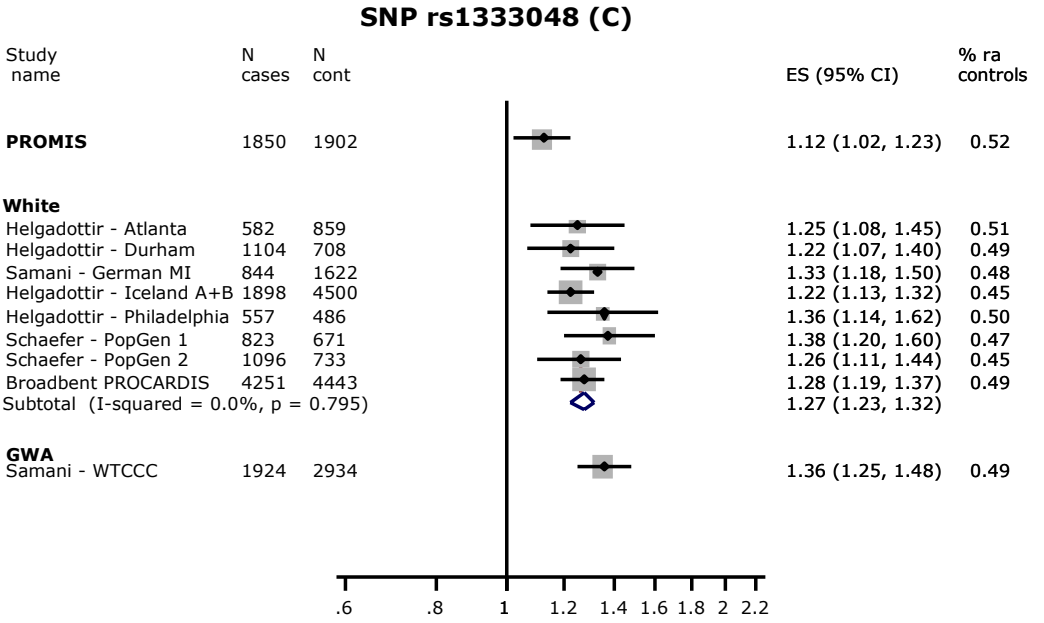
a)



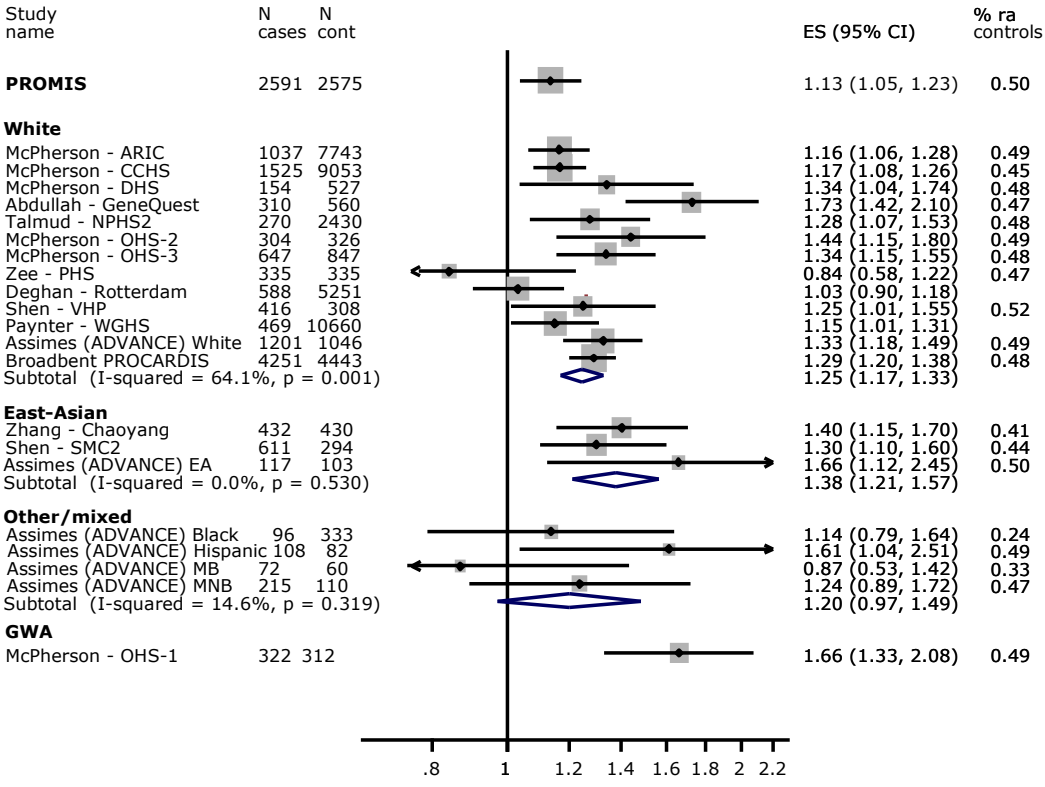
b)



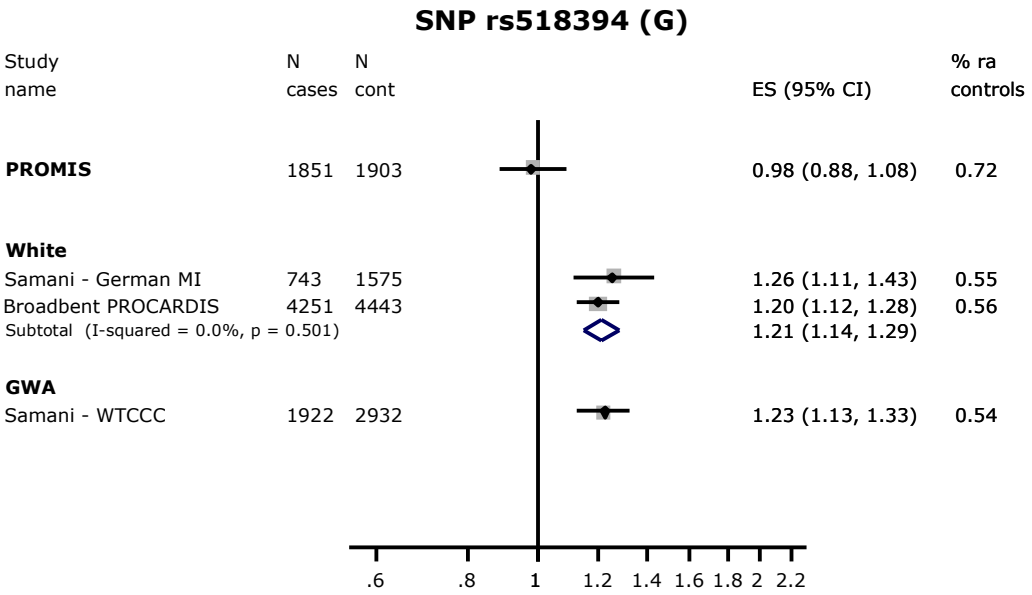
c)


SNP rs1333048 (C)
Study name
N cases
N cont
ES (95% CI)
% ra controls
PROMIS 1850 1902 1.12 (1.02, 1.23) 0.52
White
Helgadottir - Atlanta 582 859 1.25 (1.08, 1.45) 0.51
Helgadottir - Durham 1104 708 1.22 (1.07, 1.40) 0.49
Samani - German MI 844 1622 1.33 (1.18, 1.50) 0.48
Helgadottir - Iceland A+B 1898 4500 1.22 (1.13, 1.32) 0.45
Helgadottir - Philadelphia 557 486 1.36 (1.14, 1.62) 0.50
Schaefer - PopGen 1 823 671 1.38 (1.20, 1.60) 0.47
Schaefer - PopGen 2 1096 733 1.26 (1.11, 1.44) 0.45
Broadbent PROCARDIS 4251 4443 1.28 (1.19, 1.37) 0.49
Subtotal (I-squared = 0.0%, p = 0.795) 1.27 (1.23, 1.32)
GWA
Samani - WTCCC 1924 2934 1.36 (1.25, 1.48) 0.49
.6 .8 1 1.2 1.4 1.6 1.8 2 2.2

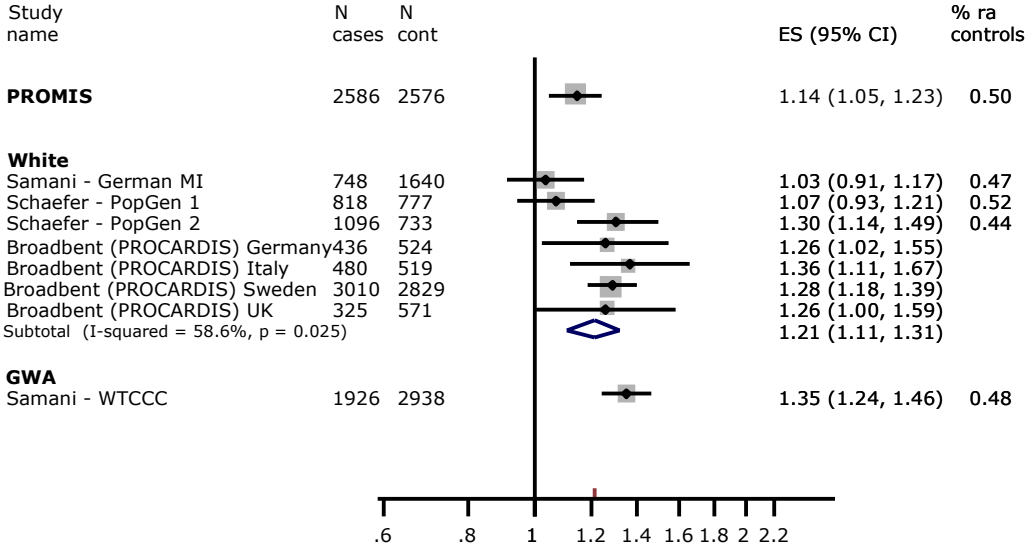
d)


SNP rs10757274 (G)
Study name
N cases
N cont
ES (95% CI)
% ra controls
PROMIS 2591 2575 1.13 (1.05, 1.23) 0.50
White
McPherson - ARIC 1037 7743 1.16 (1.06, 1.28) 0.49
McPherson - CCHS 1525 9053 1.17 (1.08, 1.26) 0.45
McPherson - DHS 154 527 1.34 (1.04, 1.74) 0.48
Abdullah - GeneQuest 310 560 1.73 (1.42, 2.10) 0.47
Talmud - NPHS2 270 2430 1.28 (1.07, 1.53) 0.48
McPherson - OHS-2 304 326 1.44 (1.15, 1.80) 0.49
McPherson - OHS-3 647 847 1.34 (1.15, 1.55) 0.48
Zee - PHS 335 335 0.84 (0.58, 1.22) 0.47
Deghan - Rotterdam 588 5251 1.03 (0.90, 1.18)
Shen - VHP 416 308 1.25 (1.01, 1.55) 0.52
Paynter - WGHS 469 10660 1.15 (1.01, 1.31)
Assimes (ADVANCE) White 1201 1046 1.33 (1.18, 1.49) 0.49
Broadbent PROCARDIS 4251 4443 1.29 (1.20, 1.38) 0.48
Subtotal (I-squared = 64.1%, p = 0.001) 1.25 (1.17, 1.33)
East-Asian
Zhang - Chaoyang 432 430 1.40 (1.15, 1.70) 0.41
Shen - SMC2 611 294 1.30 (1.10, 1.60) 0.44
Assimes (ADVANCE) EA 117 103 1.66 (1.12, 2.45) 0.50
Subtotal (I-squared = 0.0%, p = 0.530) 1.38 (1.21, 1.57)
Other/mixed
Assimes (ADVANCE) Black 96 333 1.14 (0.79, 1.64) 0.24
Assimes (ADVANCE) Hispanic 108 82 1.61 (1.04, 2.51) 0.49
Assimes (ADVANCE) MB 72 60 0.87 (0.53, 1.42) 0.33
Assimes (ADVANCE) MNB 215 110 1.24 (0.89, 1.72) 0.47
Subtotal (I-squared = 14.6%, p = 0.319) 1.20 (0.97, 1.49)
GWA
McPherson - OHS-1 322 312 1.66 (1.33, 2.08) 0.49
.8 1 1.2 1.4 1.6 1.8 2 2.2

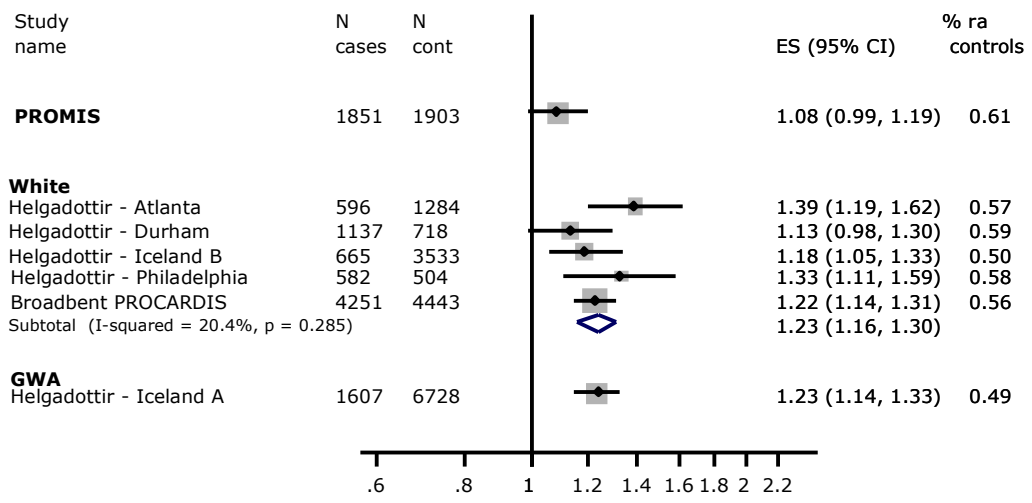
e)


SNP rs518394 (G)
Study name
N cases
N cont
ES (95% CI)
% ra controls
PROMIS 1851 1903 0.98 (0.88, 1.08) 0.72
White
Samani - German MI 743 1575 1.26 (1.11, 1.43) 0.55
Broadbent PROCARDIS 4251 4443 1.20 (1.12, 1.28) 0.56
Subtotal (I-squared = 0.0%, p = 0.501) 1.21 (1.14, 1.29)
GWA
Samani - WTCCC 1922 2932 1.23 (1.13, 1.33) 0.54
.6 .8 1 1.2 1.4 1.6 1.8 2 2.2

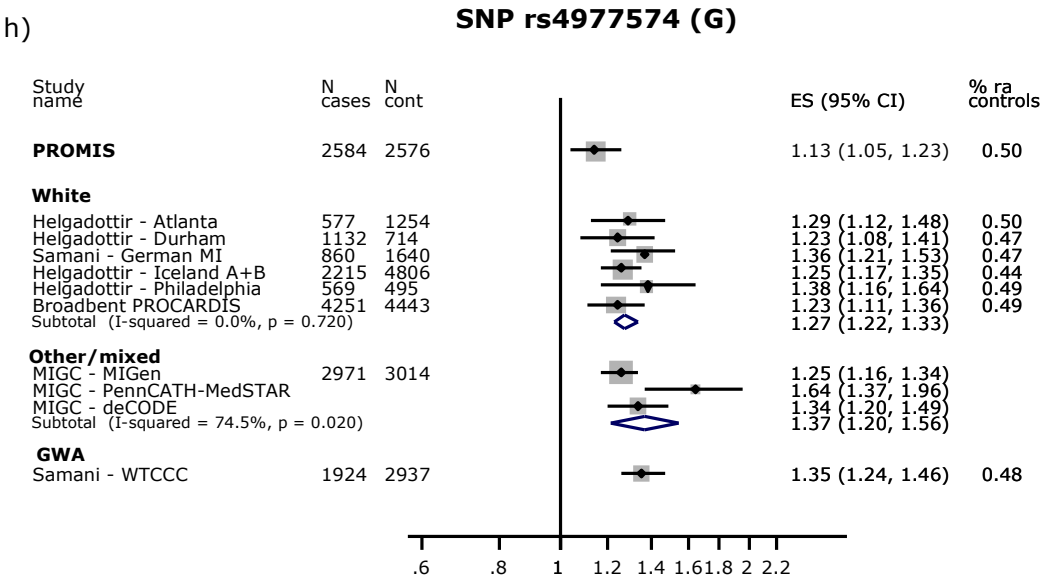
f)


SNP rs2891168 (G)
Study name
N cases
N cont
ES (95% CI)
% ra controls
PROMIS 2586 2576 1.14 (1.05, 1.23) 0.50
White
Samani - German MI 748 1640 1.03 (0.91, 1.17) 0.47
Schaefer - PopGen 1 818 777 1.07 (0.93, 1.21) 0.52
Schaefer - PopGen 2 1096 733 1.30 (1.14, 1.49) 0.44
Broadbent (PROCARDIS) Germany 436 524 1.26 (1.02, 1.55)
Broadbent (PROCARDIS) Italy 480 519 1.36 (1.11, 1.67)
Broadbent (PROCARDIS) Sweden 3010 2829 1.28 (1.18, 1.39)
Broadbent (PROCARDIS) UK 325 571 1.26 (1.00, 1.59)
Subtotal (I-squared = 58.6%, p = 0.025) 1.21 (1.11, 1.31)
GWA
Samani - WTCCC 1926 2938 1.35 (1.24, 1.46) 0.48
.6 .8 1 1.2 1.4 1.6 1.8 2 2.2

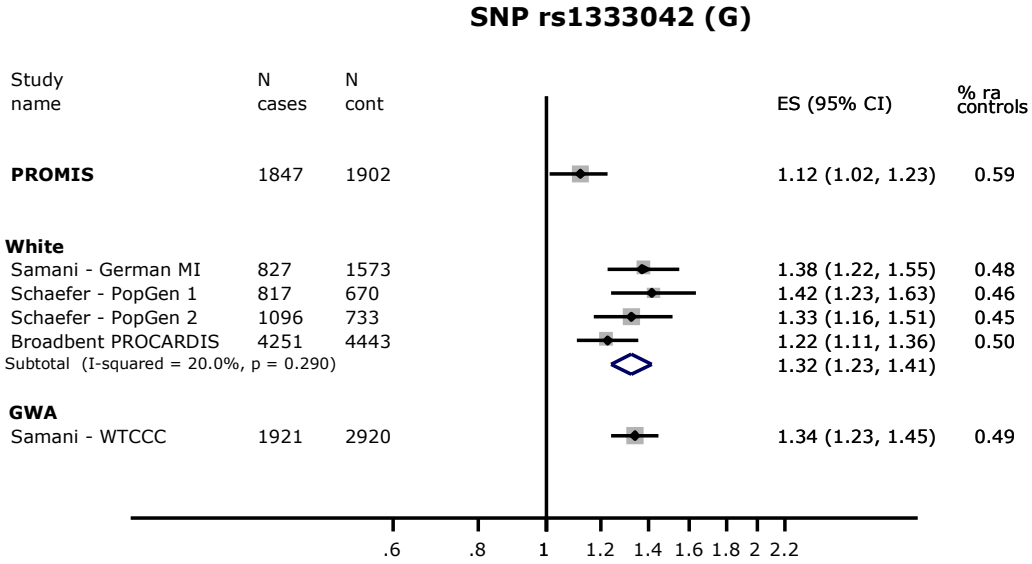
g)


SNP rs1333040 (T)
Study name
N cases
N cont
ES (95% CI)
% ra controls
PROMIS 1851 1903 1.08 (0.99, 1.19) 0.61
White
Helgadottir - Atlanta 596 1284 1.39 (1.19, 1.62) 0.57
Helgadottir - Durham 1137 718 1.13 (0.98, 1.30) 0.59
Helgadottir - Iceland B 665 3533 1.18 (1.05, 1.33) 0.50
Helgadottir - Philadelphia 582 504 1.33 (1.11, 1.59) 0.58
Broadbent PROCARDIS 4251 4443 1.22 (1.14, 1.31) 0.56
Subtotal (I-squared = 20.4%, p = 0.285) 1.23 (1.16, 1.30)
GWA
Helgadottir - Iceland A 1607 6728 1.23 (1.14, 1.33) 0.49
.6 .8 1 1.2 1.4 1.6 1.8 2 2.2

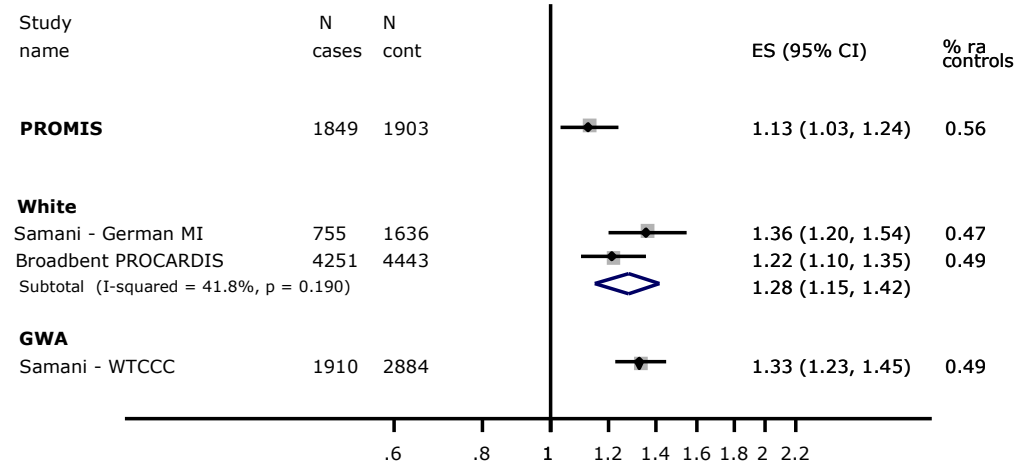
h)


SNP rs4977574 (G)
Study name
N cases
N cont
ES (95% CI)
% ra controls
PROMIS 2584 2576 1.13 (1.05, 1.23) 0.50
White
Helgadottir - Atlanta 577 1254 1.29 (1.12, 1.48) 0.50
Helgadottir - Durham 1132 714 1.23 (1.08, 1.41) 0.47
Samani - German MI 860 1640 1.36 (1.21, 1.53) 0.47
Helgadottir - Iceland A+B 2215 4806 1.25 (1.17, 1.35) 0.44
Helgadottir - Philadelphia 569 495 1.38 (1.16, 1.64) 0.49
Broadbent PROCARDIS 4251 4443 1.23 (1.11, 1.36) 0.49
Subtotal (I-squared = 0.0%, p = 0.720) 1.27 (1.22, 1.33)
Other/mixed
MIGC - MIGen 2971 3014 1.25 (1.16, 1.34)
MIGC - PennCATH-MedSTAR 1.64 (1.37, 1.96)
MIGC - deCODE 1.34 (1.20, 1.49)
Subtotal (I-squared = 74.5%, p = 0.020) 1.37 (1.20, 1.56)
GWA
Samani - WTCCC 1924 2937 1.35 (1.24, 1.46) 0.48
.6 .8 1 1.2 1.4 1.61.8 2 2.2

i)

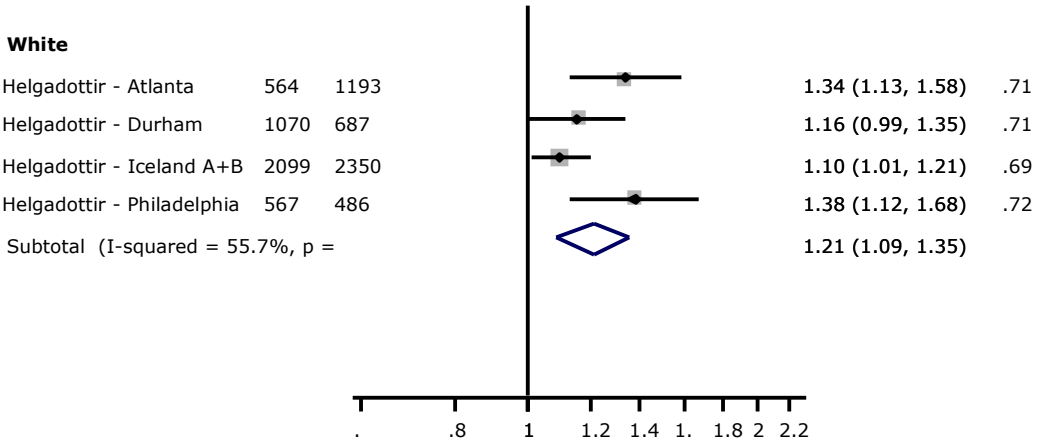

SNP rs1333042 (G)
Study name
N cases
N cont
ES (95% CI)
% ra controls
PROMIS 1847 1902 1.12 (1.02, 1.23) 0.59
White
Samani - German MI 827 1573 1.38 (1.22, 1.55) 0.48
Schaefer - PopGen 1 817 670 1.42 (1.23, 1.63) 0.46
Schaefer - PopGen 2 1096 733 1.33 (1.16, 1.51) 0.45
Broadbent PROCARDIS 4251 4443 1.22 (1.11, 1.36) 0.50
Subtotal (I-squared = 20.0%, p = 0.290) 1.32 (1.23, 1.41)
GWA
Samani - WTCCC 1921 2920 1.34 (1.23, 1.45) 0.49
.6 .8 1 1.2 1.4 1.6 1.8 2 2.2

j)


SNP rs10757272 (T)
Study name
N cases
N cont
ES (95% CI)
% ra controls
PROMIS 1849 1903 1.13 (1.03, 1.24) 0.56
White
Samani - German MI 755 1636 1.36 (1.20, 1.54) 0.47
Broadbent PROCARDIS 4251 4443 1.22 (1.10, 1.35) 0.49
Subtotal (I-squared = 41.8%, p = 0.190) 1.28 (1.15, 1.42)
GWA
Samani - WTCCC 1910 2884 1.33 (1.23, 1.45) 0.49
.6 .8 1 1.2 1.4 1.6 1.8 2 2.2

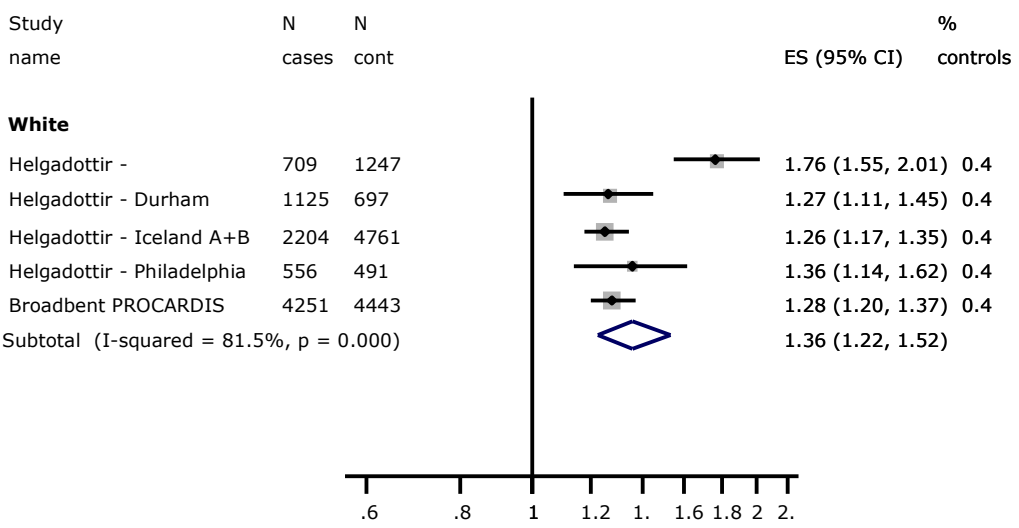
p)

SNP rs6475608 (C)



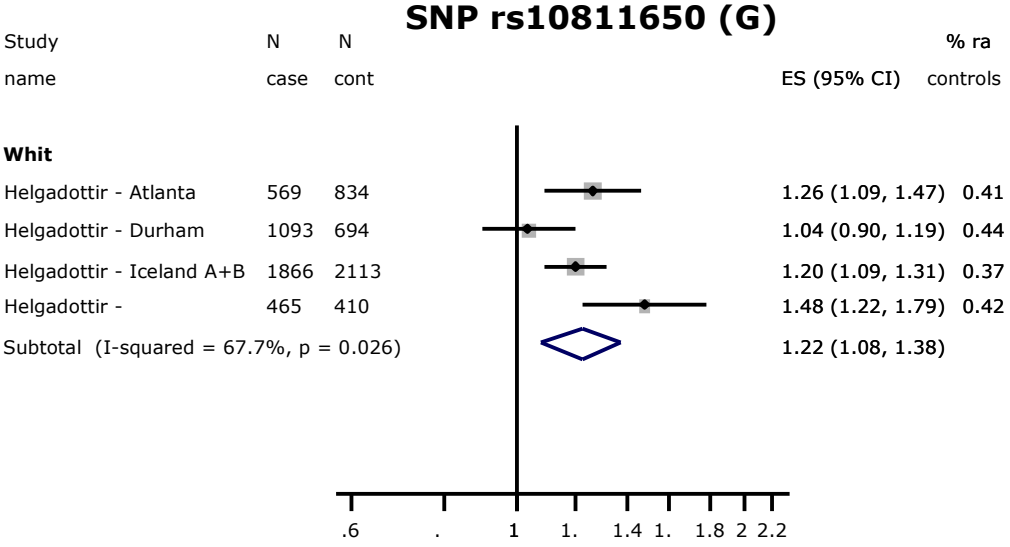
SNP rs10738607

q)

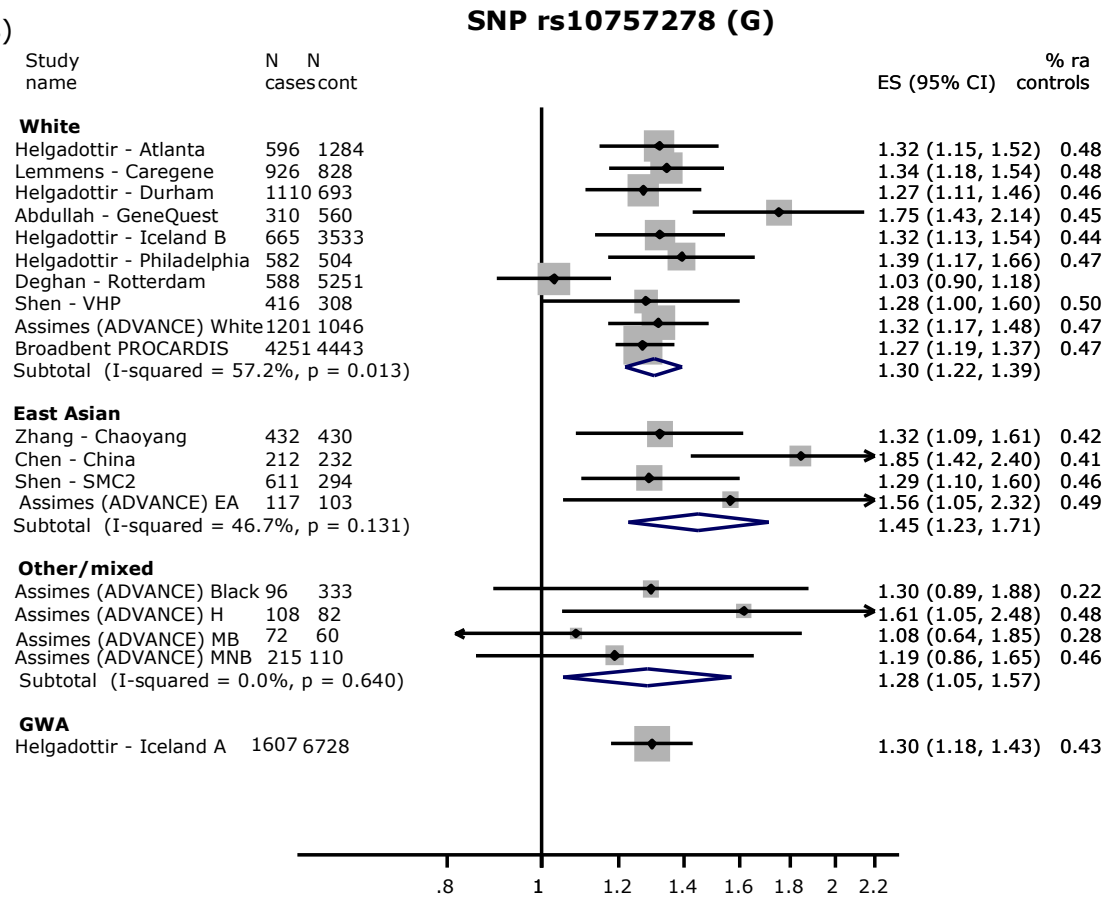


SNP rs10811650 (G)

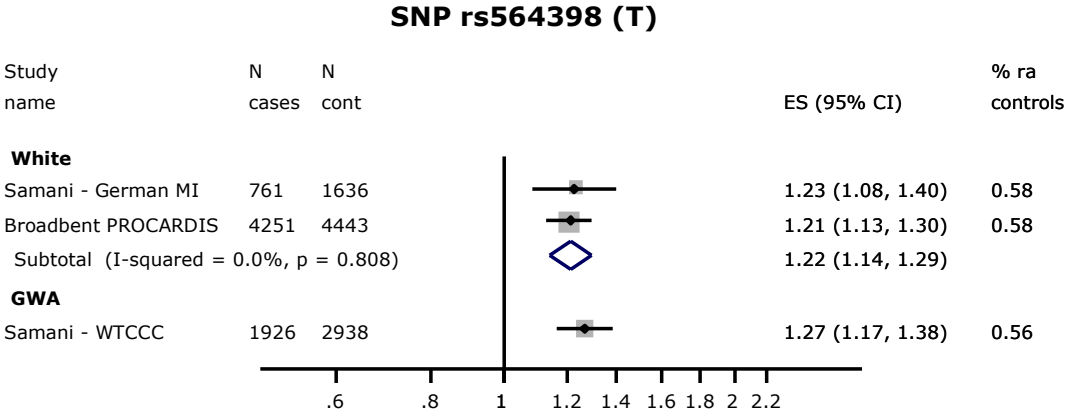
r)



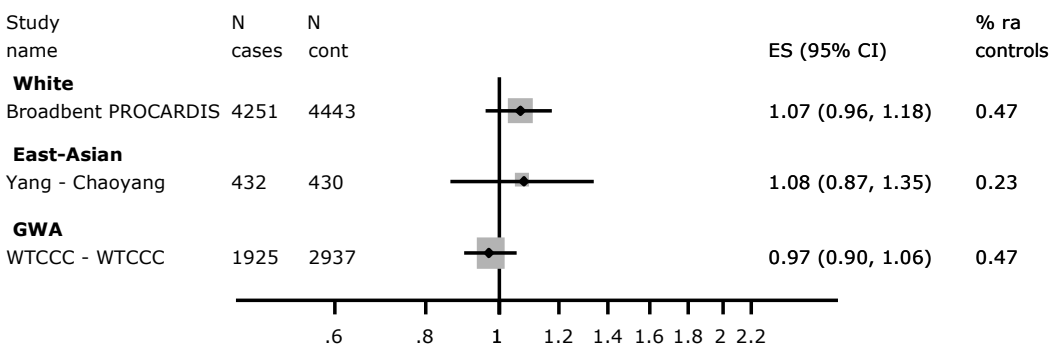
s)


SNP rs10757278 (G)
Study name
N cases
N cont
ES (95% CI)
% ra controls
White
Helgadottir - Atlanta 596 1284 1.32 (1.15, 1.52) 0.48
Lemmens - Caregene 926 828 1.34 (1.18, 1.54) 0.48
Helgadottir - Durham 1110 693 1.27 (1.11, 1.46) 0.46
Abdullah - GeneQuest 310 560 1.75 (1.43, 2.14) 0.45
Helgadottir - Iceland B 665 3533 1.32 (1.13, 1.54) 0.44
Helgadottir - Philadelphia 582 504 1.39 (1.17, 1.66) 0.47
Deghan - Rotterdam 588 5251 1.03 (0.90, 1.18)
Shen - VHP 416 308 1.28 (1.00, 1.60) 0.50
Assimes (ADVANCE) White 1201 1046 1.32 (1.17, 1.48) 0.47
Broadbent PROCARDIS 4251 4443 1.27 (1.19, 1.37) 0.47
Subtotal (I-squared = 57.2%, p = 0.013) 1.30 (1.22, 1.39)
East Asian
Zhang - Chaoyang 432 430 1.32 (1.09, 1.61) 0.42
Chen - China 212 232 1.85 (1.42, 2.40) 0.41
Shen - SMC2 611 294 1.29 (1.10, 1.60) 0.46
Assimes (ADVANCE) EA 117 103 1.56 (1.05, 2.32) 0.49
Subtotal (I-squared = 46.7%, p = 0.131) 1.45 (1.23, 1.71)
Other/mixed
Assimes (ADVANCE) Black 96 333 1.30 (0.89, 1.88) 0.22
Assimes (ADVANCE) H 108 82 1.61 (1.05, 2.48) 0.48
Assimes (ADVANCE) MB 72 60 1.08 (0.64, 1.85) 0.28
Assimes (ADVANCE) MNB 215 110 1.19 (0.86, 1.65) 0.46
Subtotal (I-squared = 0.0%, p = 0.640) 1.28 (1.05, 1.57)
GWA
Helgadottir - Iceland A 1607 6728 1.30 (1.18, 1.43) 0.43
.8
1
1.2
1.4
1.6
1.8
2
2.2

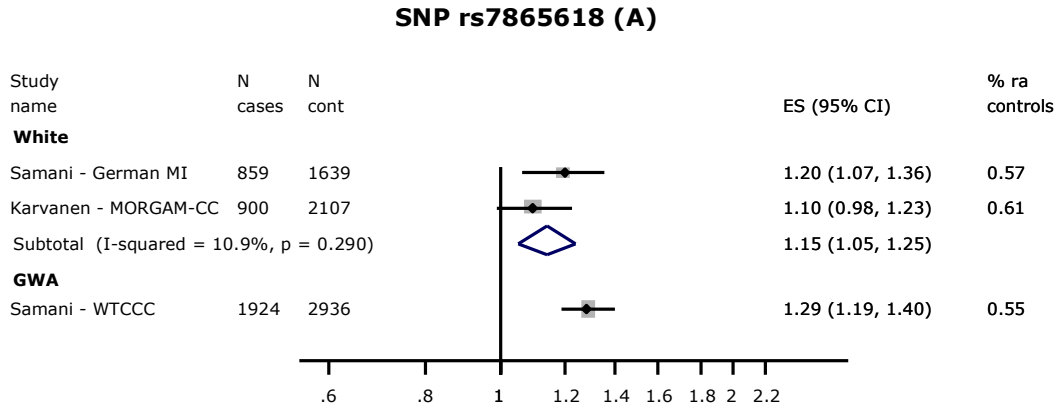
t)


SNP rs564398 (T)
Study name
N cases
N cont
ES (95% CI)
% ra controls
White
Samani - German MI 761 1636 1.23 (1.08, 1.40) 0.58
Broadbent PROCARDIS 4251 4443 1.21 (1.13, 1.30) 0.58
Subtotal (I-squared = 0.0%, p = 0.808) 1.22 (1.14, 1.29)
GWA
Samani - WTCCC 1926 2938 1.27 (1.17, 1.38) 0.56
.6
.8
1
1.2
1.4
1.6
1.8
2
2.2

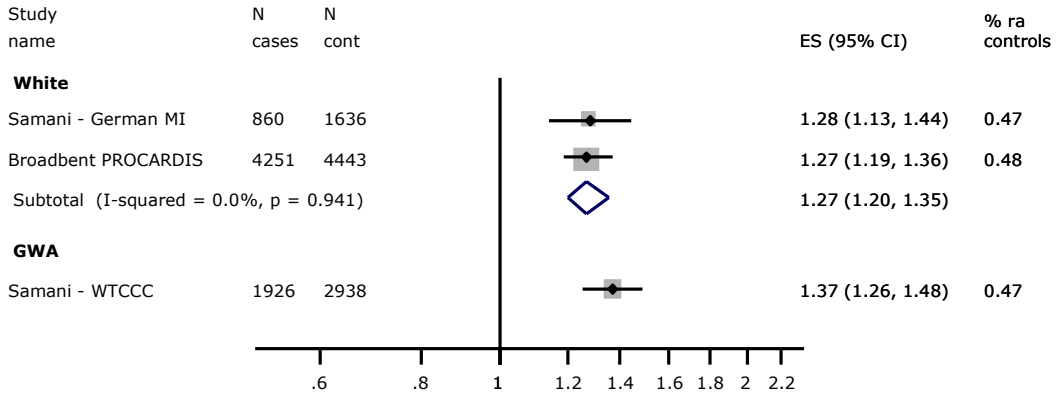
u)


SNP rs7027989 (A)
Study name
N cases
N cont
ES (95% CI)
% ra controls
White
Broadbent PROCARDIS 4251 4443 1.07 (0.96, 1.18) 0.47
East-Asian
Yang - Chaoyang 432 430 1.08 (0.87, 1.35) 0.23
GWA
WTCCC - WTCCC 1925 2937 0.97 (0.90, 1.06) 0.47
.6
.8
1
1.2
1.4
1.6
1.8
2
2.2

v)

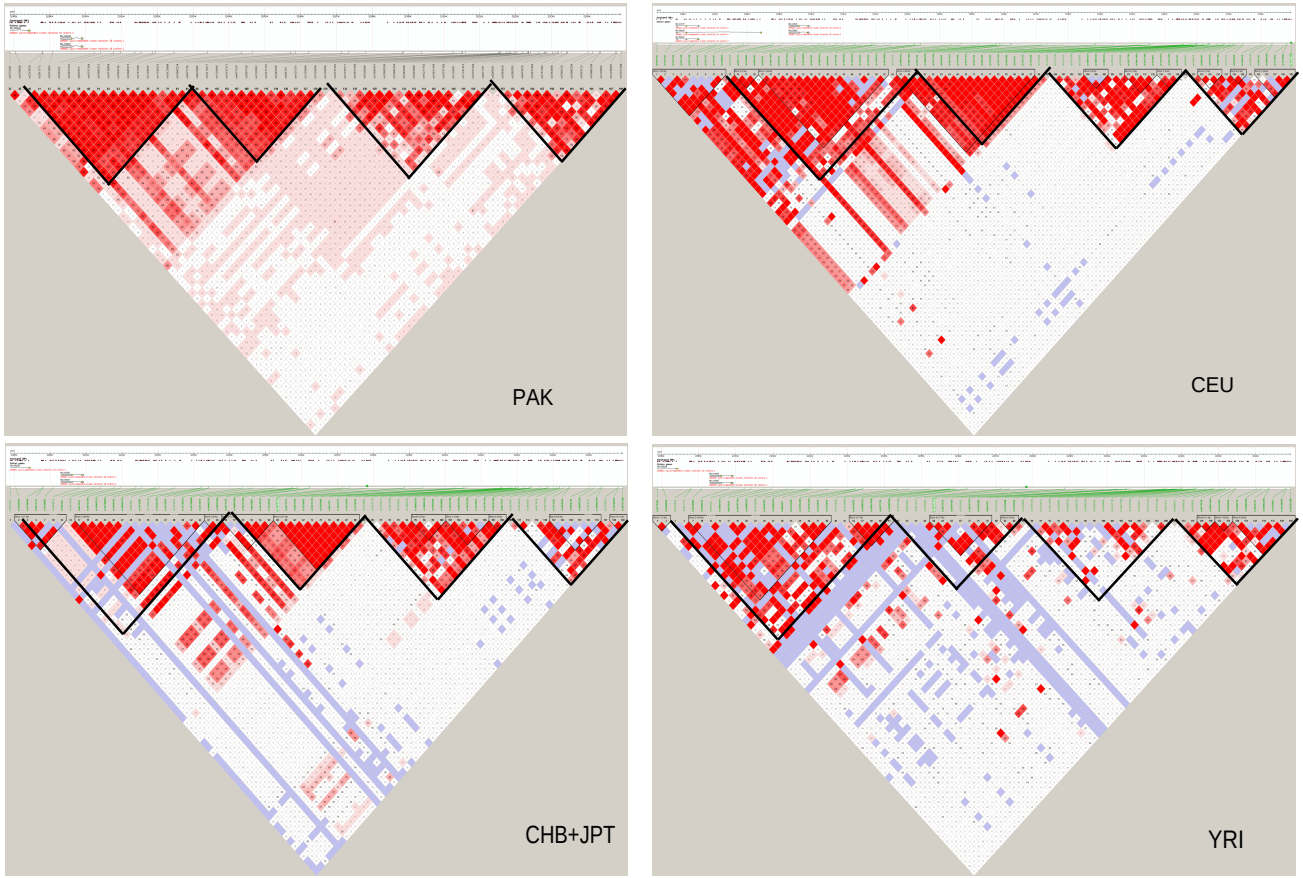


w)



% ra: frequency of the risk allele in controls, N cont: number of controls, N cases: number of cases; SNP names are followed by the name of the risk allele. Weights are from a random effect meta-analysis. Assimes H: Assimes Hispanic; Assimes MB: Assimes Mixed Blacks; Assimes MNB: Assimes Mixed non Blacks; Assimes EA: Assimes East Asians. Study acronyms: ADVANCE: Atherosclerosis Disease Vascular FuNction & GenetiC Epidemiology; PROCARDIS; AMC-PAS: Academic Medical Centre Amsterdam Premature Atherosclerosis Study; ECTIM: Etude Cas-Temoin sur l'Infarctus du Myocarde; EPIC: European Prospective Investigation into Cancer and Nutrition Study; GerMIFS: German MI Family Study; LURIC: Ludwigshafen Risk and Cardiovascular Heatlh; PopGen: PopGen biobank; UKMI: UK MI; SMC: Samsung Medical Centre; OHS: Ottawa Heart Study,; ARIC: Atherosclerosis Risk in Community; CCHS: Copenhagen City Heart Study; DHS: Dallas Heart Study; AMI Gene: Acute MI Gene; VHS: Verona Heart Study; MAHI: Mid America Heart Institute IFS: Irish Family Study; PennCATH: University of Pennsylvania Catheterization Study; MedSTAR: Washington based Study; VHP: Verona Heart Project; NPHS: Northwick Park Health Study; WTCCC: Wellcome Trust Case Control Consortium

Figure A4.9: Linkage disequilibrium patterns in Pakistanis, Europeans, Chinese and West African ethnicities



PAK: Pakistani from the PROMIS controls; CEU: Utah residents with Northern and Western European ancestry from the CEPH collection, YRI: Yoruba in Ibadan Nigeria, CHB+JPT: Han Chinese in Beijing, China and Japanese in Tokyo, Japan. Regions of strong LD are in red, of moderate LD in light red, of very moderate LD in blue and of no LD in white. PROMIS blocks are superposed to LD patterns in CEU, CHB+JPT and YRI. Only common variants genotyped in PROMIS as well as in HapMap CEU, CHB+JPT and YRI populations are used in these plots.

Table A4.1: Baseline characteristics of the participants from PROMIS in 2,592 cases and 2,577 controls

	Cases	Controls	P-value
Age, years	54.2 (10.6)	52.9 (10.1)	<0.001
Female (%)	16%	20%	<0.001
History of diabetes	22%	15%	<0.001
Family history of MI	21%	12%	<0.001
Tobacco users			<0.001
Never	35%	51%	
Ex	12%	11%	
Current	54%	38%	
Ethnic group			<0.001
Urdu (%)	46%	44%	
Punjabi (%)	23%	21%	
Pathan	7%	9%	
Balochi	2%	2%	
Sindhi	6%	9%	
Memon	3%	2%	
Gujrati	1%	4%	
Other	11%	8%	
Waist / Hip ratio	0.941 (0.001)	0.933 (0.001)	<0.001

Data are mean (standard deviation) or percentages; P-values come from a t-test of differences between the means for continuous variables and a $\chi 2$ test of independence for categorical variables. Information on family history of MI was only available in 5,068 participants for history of MI and history of diabetes, 5,041 for tobacco use and 4,509 for waist to hip ratio. All participants had information on age, sex and ethnicity.

Table A4.2: Association of all the 89 variants at the 9p21.3 locus genotyped in 1851 cases and 1903 controls using the IBC-array

SNP	Position	Minor allele	Major allele	Cases (MM/Mm/mm)	Controls (MM/Mm/mm)	HWE p-value in controls	Odds ratio (95%CI)	P-value	R2 value with rs1333049
rs7023954	21806758	T	C	113/676/10	122/695/10	0.46	1.02 (0.91; 1.13)	0.75	0.00
rs4382559	21934818	A	G	8/219/1623	11/235/165	0.36	1.02 (0.85; 1.23)	0.85	0.00
rs10965197	21938666	T	C	200/781/86	193/779/93	0.12	0.96 (0.87; 1.06)	0.45	0.03
rs2518722	21942926	T	C	36/370/144	30/388/148	0.44	0.96 (0.83; 1.1)	0.53	0.00
rs10757261	21944953	T	C	198/774/87	195/812/89	0.60	1.01 (0.91; 1.11)	0.85	0.03
rs7041637	21951866	T	G	334/836/67	339/889/67	0.12	0.99 (0.9; 1.08)	0.76	0.22
rs3731257	21956221	A	G	303/832/71	315/879/70	0.14	1 (0.91; 1.1)	0.98	0.18
rs3731246	21961989	G	C	10/198/164	12/205/168	0.06	1 (0.82; 1.2)	0.96	0.01
rs2811708	21963422	T	G	69/499/128	61/538/130	0.52	1.01 (0.89; 1.14)	0.89	0.00
rs3731239	21964218	C	T	116/672/10	116/690/10	0.62	1 (0.89; 1.11)	0.93	0.10
rs2811709	21970151	A	G	10/192/164	10/195/169	0.13	0.95 (0.78; 1.15)	0.59	0.00
rs3731217	21974661	G	T	21/346/148	30/350/152	0.07	1.06 (0.91; 1.23)	0.47	0.01
rs3731201	21978896	G	A	14/243/159	14/243/164	0.16	0.94 (0.79; 1.12)	0.48	0.01
rs7036656	21980457	C	T	68/517/126	62/550/129	0.69	1 (0.89; 1.13)	0.96	0.00
rs3218020	21987872	T	C	320/851/67	328/887/68	0.17	0.97 (0.89; 1.07)	0.55	0.29
rs2811712	21988035	G	A	28/377/144	34/379/148	0.10	1.02 (0.88; 1.17)	0.81	0.02
rs3218002	21990841	T	C	27/367/145	34/369/150	0.05	1.02 (0.89; 1.18)	0.75	0.01
rs1063192	21993367	G	A	131/719/99	134/746/10	0.95	1.02 (0.92; 1.13)	0.73	0.19
rs2069418	21999698	G	C	148/757/94	149/770/98	0.95	0.99 (0.89; 1.1)	0.83	0.15
rs545226	22002422	C	T	331/858/64	360/882/64	0.05	0.99 (0.9; 1.09)	0.87	0.26
rs10811641	22004137	G	C	332/848/66	343/884/67	0.08	0.98 (0.89; 1.07)	0.66	0.28
rs643319	22007836	A	C	242/843/76	258/840/80	0.10	1.01 (0.92; 1.11)	0.78	0.20
rs518394	22009673	C	G	152/764/92	149/781/97	0.69	0.98 (0.88; 1.08)	0.68	0.15
rs10757264	22009732	A	G	299/885/65	303/878/71	0.25	0.97 (0.89; 1.07)	0.57	0.18
rs10965212	22013795	A	T	259/873/71	283/872/74	0.28	1.03 (0.94; 1.13)	0.53	0.24
rs10738604	22015493	A	G	325/852/67	330/888/68	0.15	0.97 (0.88; 1.06)	0.51	0.28
rs7049105	22018801	A	G	257/874/71	285/873/74	0.26	1.04 (0.94; 1.14)	0.46	0.24
rs10965215	22019445	G	A	262/878/71	286/876/74	0.31	1.03 (0.94; 1.13)	0.54	0.24
rs2151280	22024719	G	A	266/876/70	289/880/73	0.36	1.03 (0.94; 1.14)	0.49	0.24
rs1333035	22034059	C	T	24/370/145	31/375/149	0.18	1.03 (0.89; 1.19)	0.66	0.02
rs11790231	22043591	T	C	36/392/142	16/390/149	0.10	0.85 (0.74; 0.98)	0.03	0.09
rs10120688	22046499	G	A	257/847/74	282/870/75	0.26	1.06 (0.97; 1.17)	0.22	0.26
rs10965224	22057276	T	A	120/694/10	123/731/10	0.81	1.04 (0.94; 1.16)	0.45	0.23
rs4977756	22058652	C	T	121/694/10	124/731/10	0.81	1.04 (0.94; 1.16)	0.45	0.23
rs9632884	22062301	C	G	267/818/76	297/872/73	0.16	1.1 (1; 1.2)	0.05	0.43
rs9632885	22062638	C	T	370/853/62	407/923/57	0.33	1.13 (1.03; 1.24)	0.009	0.57
rs7855162	22064793	C	T	8/219/1622	13/238/165	0.16	1.14 (0.94; 1.37)	0.18	0.06
rs1412832	22067543	G	A	98/643/111	102/673/11	0.90	1.04 (0.93; 1.16)	0.47	0.19
rs10116277	22071397	G	T	302/882/66	339/924/64	0.89	1.1 (1; 1.21)	0.05	0.55
rs10965228	22072380	C	T	6/183/1659	5/206/1687	0.83	1.12 (0.91; 1.37)	0.29	0.06
rs1333040	22073404	G	A	263/867/72	302/888/71	0.36	1.08 (0.99; 1.19)	0.09	0.48
rs10757272	22078260	C	T	333/876/64	373/945/58	0.82	1.13 (1.03; 1.24)	0.01	0.68
rs10757274	22086055	A	G	424/892/53	464/968/47	0.46	1.14 (1.04; 1.25)	0.005	0.86
rs4977574	22088574	A	G	425/888/53	464/966/47	0.52	1.14 (1.04; 1.25)	0.006	0.86
rs2891168	22088619	T	C	426/888/53	465/965/47	0.55	1.14 (1.04; 1.25)	0.006	0.87
rs1537372	22093183	A	C	491/880/48	425/956/52	0.75	0.88 (0.8; 0.96)	0.006	0.78
rs1333042	22093813	A	G	281/869/70	321/915/66	0.81	1.12 (1.02; 1.23)	0.02	0.61
rs10511701	22102599	T	C	316/881/65	355/940/60	0.82	1.12 (1.02; 1.23)	0.02	0.69
rs2383206	22105026	T	C	367/890/59	401/939/55	0.89	1.11 (1.02; 1.22)	0.02	0.78
rs10965235	22105105	A	C	6/203/1630	10/202/168	0.14	1.05 (0.86; 1.28)	0.64	0.06
rs944797	22105286	A	G	364/891/59	401/939/55	0.89	1.12 (1.02; 1.22)	0.02	0.78
rs1537375	22106071	T	C	316/873/65	352/941/60	0.74	1.12 (1.02; 1.23)	0.02	0.69
rs17761446	22108102	G	T	6/197/1647	10/202/169	0.14	1.08 (0.89; 1.32)	0.43	0.06
rs1333048	22115347	A	C	401/886/56	437/952/51	0.93	1.12 (1.02; 1.23)	0.01	0.89
rs1333049	22115503	C	G	440/904/50	487/964/45	0.58	1.13 (1.03; 1.24)	0.01	1.00
rs1333050	22115913	C	T	407/886/55	444/911/54	0.09	1.09 (0.99; 1.19)	0.07	0.38
rs10757281	22117613	T	C	59/513/127	60/512/133	0.21	0.98 (0.87; 1.1)	0.71	0.00
rs12379111	22118180	C	G	15/258/157	12/263/162	0.62	0.98 (0.83; 1.17)	0.84	0.01
rs10811658	22118600	T	C	126/707/10	142/709/10	0.15	1 (0.9; 1.11)	0.95	0.03
rs7020996	22119579	T	C	29/361/146	19/368/151	0.63	0.97 (0.84; 1.13)	0.71	0.00
rs10965241	22119594	G	C	53/492/130	66/514/132	0.07	1.02 (0.91; 1.16)	0.70	0.04
rs10965243	22120065	C	T	23/342/148	16/336/155	0.79	0.95 (0.81; 1.1)	0.48	0.00
rs10965244	22120389	T	A	53/501/129	67/523/131	0.11	1.02 (0.9; 1.16)	0.71	0.04
rs2383208	22122076	G	A	36/416/139	29/443/143	0.48	1.01 (0.88; 1.16)	0.85	0.01
rs7045889	22123251	C	T	333/848/66	346/882/67	0.07	1 (0.92; 1.1)	0.94	0.07
rs10217762	22123645	C	T	337/875/63	335/912/65	0.57	1.02 (0.93; 1.12)	0.72	0.02
rs10811659	22123716	G	A	71/553/122	69/584/125	0.94	0.99 (0.88; 1.11)	0.87	0.02
rs10811661	22124094	G	A	35/401/141	32/438/143	0.92	1.06 (0.93; 1.22)	0.39	0.00
rs10757283	22124172	T	C	339/875/63	335/913/65	0.60	1.01 (0.93; 1.11)	0.75	0.02
rs1333051	22126489	T	A	11/243/159	13/286/160	0.88	1.17 (0.99; 1.39)	0.07	0.01
rs11791416	22128105	C	T	101/665/10	117/687/10	0.49	1.03 (0.92; 1.15)	0.61	0.03
rs10757284	22128458	G	C	281/819/74	272/900/73	0.88	1.01 (0.92; 1.11)	0.83	0.05
rs4977761	22128762	T	C	317/838/69	291/940/67	0.21	1 (0.91; 1.1)	0.98	0.05
rs2065501	22130224	T	G	106/637/11	128/688/10	0.20	1.1 (0.99; 1.22)	0.08	0.00
rs7854629	22131034	C	T	149/749/95	173/775/95	0.38	1.04 (0.94; 1.15)	0.42	0.01
rs2065505	22131790	G	A	92/619/113	99/671/113	1.00	1.05 (0.94; 1.17)	0.37	0.01
rs6475610	22131894	G	A	240/805/80	231/869/80	0.88	0.99 (0.9; 1.09)	0.90	0.00
rs10811664	22132907	A	G	13/275/156	14/278/161	0.54	0.97 (0.82; 1.15)	0.73	0.00
rs2151283	22134305	T	G	271/858/72	281/886/73	0.59	1.01 (0.92; 1.11)	0.82	0.00
rs7853656	22134530	C	A	135/729/98	149/722/10	0.16	1.01 (0.91; 1.12)	0.88	0.00
rs7856219	22140261	G	A	382/895/57	374/928/60	0.64	0.98 (0.9; 1.08)	0.70	0.00
rs7047414	22141412	T	G	24/315/151	19/317/156	0.50	0.93 (0.8; 1.08)	0.35	0.00
rs10965256	22141465	A	G	40/437/137	49/459/139	0.14	1.06 (0.93; 1.2)	0.40	0.00
rs7853123	22143360	T	C	205/758/88	211/812/88	0.25	1.02 (0.93; 1.13)	0.63	0.00
rs10965258	22143663	G	A	71/533/124	72/533/129	0.07	1 (0.89; 1.13)	1.00	0.00
rs1333052	22147908	G	T	378/909/56	382/940/58	0.96	1.01 (0.92; 1.1)	0.91	0.00
rs12238587	22148168	T	A	20/265/156	16/302/158	0.67	1.03 (0.87; 1.21)	0.72	0.00
rs10122243	22148924	C	T	300/859/69	313/910/68	0.78	1.02 (0.93; 1.12)	0.72	0.00
rs10757288	22149416	G	A	189/752/90	188/815/90	0.87	1.02 (0.93; 1.13)	0.66	0.00

M: major allele; m: minor allele according to the frequency in the overall population, r^2: Linkage disequilibrium between rs1333049 and other SNPs; MAF: Minor allele frequency. The odds ratio represent the per major allele increase in risk of MI adjusted for age, sex and the first two PCs. HWE: Hardy Weinberg Equilibrium test was performed in control participants only.

Table A4.3: Association of SNPs at the 9p21.3 locus significantly associated with MI in PROMIS

SNP	Position	Allele m/M	RAF	N cases (MM/Mm/mm)	N controls (MM/Mm/mm)	OR (95%CI)	P-value	HWE p-value	R^2
rs10757274	22086055	A/G	0.50	742/1253/596	633/1302/640	1.13 (1.05;1.23)	2.E-03	0.57	0.86
rs1333049	22115503	G/C	0.51	697/1273/617	609/1290/674	1.13 (1.05;1.22)	2.E-03	0.86	1.00
rs1537372	22093183	A/C	0.47	670/1237/683	718/1273/581	1.13 (1.04;1.22)	2.E-03	0.71	0.78
rs2891168	22088619	T/C	0.50	746/1243/597	634/1298/644	1.14 (1.05;1.23)	1.E-03	0.69	0.86
rs4977574	22088574	A/G	0.50	746/1242/596	636/1298/642	1.13 (1.05;1.23)	1.E-03	0.69	0.86
rs9632885	22062638	C/T	0.54	857/1223/502	772/1240/560	1.12 (1.04;1.21)	3.E-03	0.14	0.56

M: major allele; m: minor allele according to the frequency in the overall population; RAF: Risk allele frequency in controls; N cases: Number of cases; OR: Odds ratio representing the increased risk of MI per risk allele adjusted for age, sex and self reported ethnicity.R2: Linkage disequilibrium between rs1333049 and other SNPs; HWE: Hardy Weinberg Equilibrium test was performed in control participants only. The risk allele corresponds to the minor allele for rs1537372 and to the major allele for the other SNPs

Table A4.4: Haplotype analyses of SNPs rs1333049 and rs1412832

Polymorphisms		Haplotype Frequencies		Haplotypic Odds Ratio [95% CI]	
rs1412832	rs1333049	Controls	Cases	Model 1	Model 2
G	G	0.201	0.205	0.949 [0.841 – 1.072] P-value = 0.403	0.937 [0.828 – 1.060] P-value = 0.304
G	C	0.028	0.022	0.748 [0.523 – 1.068] P-value = 0.110	0.778 [0.541 – 1.120] P-value = 0.177
A	G	0.307	0.276	0.843 [0.755 – 0.940] P-value = 0.002	0.846 [0.757 – 0.946] P-value = 0.003
A	C	0.464	0.497	reference	reference
Test of haplotypic association				X^2 = 9.07 with 3 df P-value = 0.028	X^2 = 9.08 with 3 df P-value = 0.028

Reference: Reference group for the computation of the odds ratios; df: degree of freedom; X^2: Chi-square statistic for the test of haplotypic association. Model1 was unadjusted and model 2 was adjusted for age, sex and the first two principal components. Haplotypic odds ratios represent the odds ratio for haplotypes GG, GC and AC versus the most common reference group AC.

Table A4.5: Characteristics of studies included in the current literature based meta-analyses

Author	Consortium and studies names	Type	Design	Source controls	Location	N cases / N controls	Diagnostic of CAD	Genotyping platform	SNP included in the literature based meta-analysis
Abdullah	GeneQuest	Rep.	PC*	Population	USA	310/560	CAD/MI	TaqMan Assay, Applied Biosystems	rs2383206, rs10757274
Assimes	ADVANCE: White, Black, East Asian, Hispanic, Mixed (Black) and Mixed (non Black)	Rep.	CC	Health care system	USA	3618/3468	MI/PTCA/CABG/angina	TaqMan Assay, Applied Biosystems	rs2383206, rs10757274
Broadbent	PROCARDIS: Germany, Italy, Sweden, UK	Rep.	CC	Hospital	Europe	4251/4443	MI/PTCA/CABG/angina	Sequenom iPLEX™	rs518394, rs1333040,rs1333042, rs1333048, rs1333049, rs2383206, rs2891168, rs4977574, rs10116277, rs10757272, rs10757274, rs10757283
CAD Consortium	AMC-PAS, ECTIM, EPIC-Norfolk, GerMIFS[$], KORA/GOC[$], LURIC, MORGAM[$], UKMI[$], PopGen[$]	Rep.	CC	Mixed	Europe	11550/11205	MI/Stenosis (>50%)/PTCA/ CABG/Angina	Sequenom iPLEX™	rs1333049
Chen	China Study	Rep.	CC	Health care system	China	232/212	MI/Stenosis (≥50%)	TaqMan Assay, Applied Biosystems	rs10757278, rs2383207
Dehghan	Rotterdam study	Rep.	PC	Population	Netherlands	588/5251	MI/CABG/PTCA	TaqMan Assay, Applied Biosystems	rs10757274
Ghazouani	Tunisia study	Rep.	CC	Population	Tunisia	296/324	Stenosis/MI	TaqMan Assay, Applied Biosystems	rs1333049
Helgadottir	Iceland A	GWA	CC	Mixed	Europe, USA	4587/12767	MI	Infinium HumanHap300 (Illumina, USA)	rs1333040, rs10116277
	Iceland B, Atlanta, Durham, Philadelphia (PennCath)	Rep.	CC	Hospital			MI/CABG/PTCA	Centaurus (Nanogen) platform	rs1333040, rs1333048, rs4977574, rs10116277
Hinohara	SMC2 & Tokyo study	Rep.	CC	NF	Japan	1283/1857	MI/Stenosis (≥50%)	TaqMan Assay, Applied Biosystems	rs1333049
Hiura	Suita study	Rep.	CC	Population	Japan	589/2475	MI	TaqMan Assay, Applied Biosystems	rs1333049
Karvanen	MORGAM[$]	Rep.	NCC	Population	Europe	1050/1878	MI/angina/CABG/PCI	Sequenom iPLEX™	rs1333049, rs10757283
Lemmens	Caregene	Ref	CC	Population	Belgium	926/648	MI/angina/CABG/PCI	TaqMan Assay, Applied Biosystems	rs10757278
McPherson	OHS-1	GWA	CC	Hospital	Canada, USA, Europe	4306/20119	MI/PTCA/CABG/CHD death	Perlegen Sciences	rs2383206, rs10757274
	ARIC, CCHS, DHS, OHS-2, OHS-3	Rep.	CC**	Population and hospital					
MIGC	deCODE, MIGen, AMI Gene-VHP-MAHI-IFS£-GerMIFS£ II-INTERHEART£ and PennCATH£-MedSTAR£	Rep.	CC	Population, blood donor	Europe, USA	NF	MI	Affymetrix GeneChip Human Mapping 500KArray Set or Affymetrix 6.0 GeneChip	rs4977574
Paynter	Women's Genome Health Study	Rep.	PC	Population	USA	469/10660	MI/PTCA/CABG/CHD death	Luminex100 xMAP microspheres (Luminex, USA)	rs10757274

Samani	German MI study	Rep.	CC	Population	Germany	844/1605	early MI and family history of MI	Affymetrix GeneChip Human Mapping 500K Array Set	rs518394, rs1333042, rs1333048, rs1333049, rs2891168, rs10757272
Schaefer	PopGen 1, PopGen 2	Rep.	CC	Blood donors	Germany		Stenosis (>70%)/MI/PTCA/ CABG	SNPlex and TaqMan GenotypingSystem (Applied Biosystems, USA)	rs1333042, rs1333048, rs2891168
Schunkert	CARDIOGENICS: AtheroGene, German MI II[$], LMDS, MONICA/KORA[$], PopGen[$], PRIME, UKMI[$]	Rep.	CC	Mixed	Europe	8912/9828	MI/angina	TaqMan Assay, Applied Biosystems	rs1333049
Shen	SMC2	Rep.	CC	Hospital	Korea	611/294	MI/stenosis/CABG/PCI	TaqMan Assay, Applied Biosystems	rs2383206, rs10757274
Shen	VHP	Rep.	CC	Hospital	Italy	416/308	MI	TaqMan Assay, Applied Biosystems	rs2383206, rs10757274
Talmud	NPHS2	Rep.	PC	Population	UK	264/2430	MI/PTCA/CABG	TaqMan Assay, Applied Biosystems	rs2383206, rs10757274
WTCCC	WTCCC	GWA	CC	Population and blood donors	UK	1926/2938	MI or CABG/PTCA	Affymetric GeneChip Human Mapping 500K Array Set	rs518394, rs1333042, rs1333048, rs1333049, rs2891168, rs10757272, rs10757283
Ye	Bruneck study	Rep.	PC	Population	Italy	56/713	MI/CABG/PTCA	TaqMan Assay, Applied Biosystems	rs1333049
Zee	Physician's Health Study	Rep.	NCC	Population (physicians)	USA	335/335	MI	NF	rs2383206, rs10757274
Zhang	Chaoyang study	Rep.	CC	Hospital	China	432/430	MI	Orchid BioSciences (GenomLab SNPstream genotyping platform)	rs2383206, rs10757274
Zhou	Wuhan study	Rep.	CC	Population	China	1360/1360	stenosis (>= 50%)/MI/CABG/PCI	TaqMan Assay, Applied Biosystems	rs2383206

Rep: Replication study; GWA: genome wide association study; PC: prospective cohort study; CC: Case-control study; NCC: Nested case-control study; MI: Myocardial Infarction; NF: Not found. *: analysed as a case-control study; **: except ARIC, a prospective cohort study; [£] The following studies from MIGC were not included in the analysis for rs4977574 because they overlapped partially with previously published results. For prospective cohorts, the number of controls corresponds to the number of individuals who had not developed the disease at the end of follow-up. Study acronyms: ADVANCE: Atherosclerosis Disease Vascular FuNction & GenetiC Epidemiology; PROCARDIS; AMC-PAS: Academic Medical Centre Amsterdam Premature Atherosclerosis Study; ECTIM: Etude Cas-Temoin sur l'Infarctus du Myocarde; EPIC: European Prospective Investigation into Cancer and Nutrition Study; GerMIFS: German MI Family Study; LURIC: Ludwigshafen Risk and Cardiovascular Heatlh; PopGen: PopGen biobank; UKMI: UK MI; SMC: Samsung Medical Centre; OHS: Ottawa Heart Study,; ARIC: Atherosclerosis Risk in Community; CCHS: Copenhagen City Heart Study; DHS: Dallas Heart Study; AMI Gene: Acute MI Gene; VHS: Verona Heart Study; MAHI: Mid America Heart Institute IFS: Irish Family Study; PennCATH: University of Pennsylvania Catheterization Study; MedSTAR: Washington based Study; VHP: Verona Heart Project; NPHS: Northwick Park Health Study; WTCCC: Wellcome Trust Case Control Consortium.

References

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

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

3. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC. A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007;316:1488-1491.

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

5. Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, Konig IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;117:1675-1684.

6. Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, Konig IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;117:1675-1684.

7. Jarinova O, Stewart AF, Roberts R, Wells G, Lau P, Naing T, Buerki C, McLean BW, Cook RC, Parker JS, McPherson R. Functional Analysis of the Chromosome 9p21.3 Coronary Artery Disease Risk Locus. *Arterioscler Thromb Vasc Biol* 2009.

8. Nishtar S. Prevention of coronary heart disease in south Asia. *Lancet* 2002;360:1015-1018.

9. Saleheen D, Zaidi M, Rasheed A, Ahmad U, Hakeem A, Murtaza M, Kayani W, Faruqui A, Kundi A, Zaman KS, Yaqoob Z, Cheema LA, Samad A, Rasheed SZ, Mallick NH, Azhar M, Jooma R, Gardezi AR, Memon N, Ghaffar A, Fazal uR, Khan N, Shah N, Ali SA, Samuel M, Hanif F, Yameen M, Naz S, Sultana A, Nazir A, Raza S, Shazad M, Nasim S, Javed MA, Ali SS, Jafree M, Nisar MI, Daood MS, Hussain A, Sarwar N, Kamal A, Deloukas P, Ishaq M, Frossard P, Danesh J. The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia. *Eur J Epidemiol* 2009;24:329-338.

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

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

12. Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, Konig IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;117:1675-1684.

13. Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von EE, Khoury MJ, Cohen B, vey-Smith G, Grimshaw J, Scheet P, Gwinn M, Williamson RE, Zou GY, Hutchings K, Johnson CY, Tait V, Wiens M, Golding J, van DC, McLaughlin J, Paterson A, Wells G, Fortier I, Freedman M, Zecevic M, King R, Infante-Rivard C, Stewart A, Birkett N. STrengthening the REporting of Genetic Association studies (STREGA)--an extension of the STROBE statement. *Eur J Clin Invest* 2009;39:247-266.

14. Keating BJ, Tischfield S, Murray SS, Bhangale T, Price TS, Glessner JT, Galver L, Barrett JC, Grant SF, Farlow DN, Chandrupatla HR, Hansen M, Ajmal S, Papanicolaou GJ, Guo Y, Li M, Derohannessian S, de Bakker PI, Bailey SD, Montpetit A, Edmondson AC, Taylor K, Gai X, Wang SS, Fornage M, Shaikh T, Groop L, Boehnke M, Hall AS, Hattersley AT, Frackelton E, Patterson N, Chiang CW, Kim CE, Fabsitz RR, Ouwehand W, Price AL, Munroe P, Caulfield M, Drake T, Boerwinkle E, Reich D, Whitehead AS, Cappola TP, Samani NJ, Lusis AJ, Schadt E, Wilson JG, Koenig W, McCarthy MI, Kathiresan S, Gabriel SB, Hakonarson H, Anand SS, Reilly M, Engert JC, Nickerson DA, Rader DJ, Hirschhorn JN, Fitzgerald GA. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS One* 2008;3:e3583.

15. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559-575.

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

17. Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, Brodsky J, Jones CG, Zaitlen NA, Varilo T, Kaakinen M, Sovio U, Ruokonen A, Laitinen J, Jakkula E, Coin L, Hoggart C, Collins A, Turunen H, Gabriel S, Elliot P, McCarthy MI, Daly MJ, Jarvelin MR, Freimer NB, Peltonen L. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet* 2009;41:35-46.

18. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263-265.

19. Tregouet DA, Garelle V. A new JAVA interface implementation of THESIAS: testing haplotype effects in association studies. *Bioinformatics* 2007;23:1038-1039.

20. Tregouet DA, Ricard S, Nicaud V, Arnould I, Soubigou S, Rosier M, Duverger N, Poirier O, Mace S, Kee F, Morrison C, Denefle P, Tiret L, Evans A, Deleuze JF, Cambien F. In-depth haplotype analysis of ABCA1 gene polymorphisms in relation to plasma ApoA1 levels and myocardial infarction. *Arterioscler Thromb Vasc Biol* 2004;24:775-781.

21. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263-265.

22. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539-1558.

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

24. Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, Konig IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;117:1675-1684.

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

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

27. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC. A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007;316:1488-1491.

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

29. Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, Konig IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;117:1675-1684.

30. Abdullah KG, Li L, Shen GQ, Hu Y, Yang Y, MacKinlay KG, Topol EJ, Wang QK. Four SNPS on chromosome 9p21 confer risk to premature, familial CAD and MI in an American Caucasian population (GeneQuest). *Ann Hum Genet* 2008;72:654-657.

31. Assimes TL, Knowles JW, Basu A, Iribarren C, Southwick A, Tang H, Absher D, Li J, Fair JM, Rubin GD, Sidney S, Fortmann SP, Go AS, Hlatky MA, Myers RM, Risch N, Quertermous T. Susceptibility locus for clinical and subclinical coronary artery disease at chromosome 9p21 in the multi-ethnic ADVANCE study. *Hum Mol Genet* 2008;17:2320-2328.

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

33. Samani NJ, Deloukas P, Erdmann J, Hengstenberg C, Kuulasmaa K, McGinnis R, Schunkert H, Soranzo N, Thompson J, Tiret L, Ziegler A. Large scale association analysis of novel genetic loci for coronary artery disease. *Arterioscler Thromb Vasc Biol* 2009;29:774-780.

34. Dehghan A, van HM, Sijbrands EJ, Oostra BA, Hofman A, van Duijn CM, Witteman JC. Lack of association of two common polymorphisms on 9p21 with risk of coronary heart disease and myocardial infarction; results from a prospective cohort study. *BMC Med* 2008;6:30.:30.

35. Karvanen J, Silander K, Kee F, Tiret L, Salomaa V, Kuulasmaa K, Wiklund PG, Virtamo J, Saarela O, Perret C, Perola M, Peltonen L, Cambien F, Erdmann J, Samani NJ, Schunkert H, Evans A. The impact of newly identified loci on coronary heart disease, stroke and total mortality in the MORGAM prospective cohorts. *Genet Epidemiol* 2009;33:237-246.

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

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

38. Paynter NP, Chasman DI, Buring JE, Shiffman D, Cook NR, Ridker PM. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. *Ann Intern Med* 2009;%20;150:65-72.

39. Schaefer AS, Richter GM, Groessner-Schreiber B, Noack B, Nothnagel M, El Mokhtari NE, Loos BG, Jepsen S, Schreiber S. Identification of a shared genetic susceptibility locus for coronary heart disease and periodontitis. *PLoS Genet* 2009;5:e1000378.

40. Shen GQ, Rao S, Martinelli N, Li L, Olivieri O, Corrocher R, Abdullah KG, Hazen SL, Smith J, Barnard J, Plow EF, Girelli D, Wang QK. Association between four SNPs on chromosome 9p21 and myocardial infarction is replicated in an Italian population. *J Hum Genet* 2008;53:144-150.

41. Talmud PJ, Cooper JA, Palmen J, Lovering R, Drenos F, Hingorani AD, Humphries SE. Chromosome 9p21.3 coronary heart disease locus genotype and prospective risk of CHD in healthy middle-aged men. *Clin Chem* 2008;54:467-474.

42. Zee RY, Ridker PM. Two common gene variants on chromosome 9 and risk of atherothrombosis. *Stroke* 2007;38:e111.

43. Ye S, Willeit J, Kronenberg F, Xu Q, Kiechl S. Association of genetic variation on chromosome 9p21 with susceptibility and progression of atherosclerosis: a population-based, prospective study. *J Am Coll Cardiol* 2008;52:378-384.

44. Chen Z, Qian Q, Ma G, Wang J, Zhang X, Feng Y, Shen C, Yao Y. A common variant on chromosome 9p21 affects the risk of early-onset coronary artery disease. *Mol Biol Rep* 2009;36:889-893.

45. Hinohara K, Nakajima T, Takahashi M, Hohda S, Sasaoka T, Nakahara K, Chida K, Sawabe M, Arimura T, Sato A, Lee BS, Ban JM, Yasunami M, Park JE, Izumi T, Kimura A. Replication of the association between a chromosome 9p21 polymorphism and coronary artery disease in Japanese and Korean populations. *J Hum Genet* 2008;53:357-359.

46. Hiura Y, Fukushima Y, Yuno M, Sawamura H, Kokubo Y, Okamura T, Tomoike H, Goto Y, Nonogi H, Takahashi R, Iwai N. Validation of the association of genetic variants on chromosome 9p21 and 1q41 with myocardial infarction in a Japanese population. *Circ J* 2008;72:1213-1217.

47. Shen GQ, Li L, Rao S, Abdullah KG, Ban JM, Lee BS, Park JE, Wang QK. Four SNPs on chromosome 9p21 in a South Korean population implicate a genetic locus that confers high cross-race risk for development of coronary artery disease. *Arterioscler Thromb Vasc Biol* 2008;28:360-365.

48. Zhang Q, Wang XF, Cheng SS, Wan XH, Cao FF, Li L, Chen XD, Liu WJ, Yang XC, Jin L. Three SNPs on chromosome 9p21 confer increased risk of myocardial infarction in Chinese subjects. *Atherosclerosis* 2009.

49. Zhou L, Zhang X, He M, Cheng L, Chen Y, Hu FB, Wu T. Associations between single nucleotide polymorphisms on chromosome 9p21 and risk of coronary heart disease in Chinese Han population. *Arterioscler Thromb Vasc Biol* 2008;28:2085-2089.

50. Ghazouani L, Khalifa SB, Abboud N, Perret C, Nicaud V, Khalfallah AB, Alamawi WY, Cambien F, Mahjoub T. Association of three polymorphisms selected from a genome-wide association study with coronary heart disease in the Tunisian population. *J Thromb Thrombolysis* 2009.

51. Abdullah KG, Li L, Shen GQ, Hu Y, Yang Y, MacKinlay KG, Topol EJ, Wang QK. Four SNPS on chromosome 9p21 confer risk to premature, familial CAD and MI in an American Caucasian population (GeneQuest). *Ann Hum Genet* 2008;72:654-657.

52. Dehghan A, van HM, Sijbrands EJ, Oostra BA, Hofman A, van Duijn CM, Witteman JC. Lack of association of two common polymorphisms on 9p21 with risk of coronary heart disease and myocardial infarction; results from a prospective cohort study. *BMC Med* 2008;6:30.:30.

53. Paynter NP, Chasman DI, Buring JE, Shiffman D, Cook NR, Ridker PM. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. *Ann Intern Med* 2009;%20;150:65-72.

54. Talmud PJ, Cooper JA, Palmen J, Lovering R, Drenos F, Hingorani AD, Humphries SE. Chromosome 9p21.3 coronary heart disease locus genotype and prospective risk of CHD in healthy middle-aged men. *Clin Chem* 2008;54:467-474.

55. Ye S, Willeit J, Kronenberg F, Xu Q, Kiechl S. Association of genetic variation on chromosome 9p21 with susceptibility and progression of atherosclerosis: a population-based, prospective study. *J Am Coll Cardiol* 2008;52:378-384.

56. Karvanen J, Silander K, Kee F, Tiret L, Salomaa V, Kuulasmaa K, Wiklund PG, Virtamo J, Saarela O, Perret C, Perola M, Peltonen L, Cambien F, Erdmann J, Samani NJ, Schunkert H, Evans A. The impact of newly identified loci on coronary heart disease, stroke and total mortality in the MORGAM prospective cohorts. *Genet Epidemiol* 2009;33:237-246.

57. Zee RY, Ridker PM. Two common gene variants on chromosome 9 and risk of atherothrombosis. *Stroke* 2007;38:e111.

58. Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, Konig IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;117:1675-1684.

59. Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, Konig IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;117:1675-1684.

60. Schunkert H, Gotz A, Braund P, McGinnis R, Tregouet DA, Mangino M, Linsel-Nitschke P, Cambien F, Hengstenberg C, Stark K, Blankenberg S, Tiret L, Ducimetiere P, Keniry A, Ghori MJ, Schreiber S, El Mokhtari NE, Hall AS, Dixon RJ, Goodall AH, Liptau H, Pollard H, Schwarz DF, Hothorn LA, Wichmann HE, Konig IR, Fischer M, Meisinger C, Ouwehand W, Deloukas P, Thompson JR, Erdmann J, Ziegler A, Samani NJ. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 2008;117:1675-1684.

61. Jarinova O, Stewart AF, Roberts R, Wells G, Lau P, Naing T, Buerki C, McLean BW, Cook RC, Parker JS, McPherson R. Functional Analysis of the Chromosome 9p21.3 Coronary Artery Disease Risk Locus. *Arterioscler Thromb Vasc Biol* 2009.

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

63. Amaral PP, Dinger ME, Mercer TR, Mattick JS. The eukaryotic genome as an RNA machine. *Science* 2008;319:1787-1789.

64. Jarinova O, Stewart AF, Roberts R, Wells G, Lau P, Naing T, Buerki C, McLean BW, Cook RC, Parker JS, McPherson R. Functional Analysis of the Chromosome 9p21.3 Coronary Artery Disease Risk Locus. *Arterioscler Thromb Vasc Biol* 2009.

65. Hannon GJ, Beach D. p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. *Nature* 1994;371:257-261.

66. Pomerantz J, Schreiber-Agus N, Liegeois NJ, Silverman A, Alland L, Chin L, Potes J, Chen K, Orlow I, Lee HW, Cordon-Cardo C, DePinho RA. The Ink4a tumor suppressor gene product, p19Arf, interacts with MDM2 and neutralizes MDM2's inhibition of p53. *Cell* 1998;%20;92:713-723.

67. Helgadottir A, Manolescu A, Helgason A, Thorleifsson G, Thorsteinsdottir U, Gudbjartsson DF, Gretarsdottir S, Magnusson KP, Gudmundsson G, Hicks A, Jonsson T, Grant SF, Sainz J, O'Brien SJ, Sveinbjornsdottir S, Valdimarsson EM, Matthiasson SE, Levey AI, Abramson JL, Reilly MP, Vaccarino V, Wolfe ML, Gudnason V, Quyyumi AA, Topol EJ, Rader

DJ, Thorgeirsson G, Gulcher JR, Hakonarson H, Kong A, Stefansson K. A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction. *Nat Genet* 2006;38:68-74.

68. Helgason A, Palsson S, Thorleifsson G, Grant SF, Emilsson V, Gunnarsdottir S, Adeyemo A, Chen Y, Chen G, Reynisdottir I, Benediktsson R, Hinney A, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Schafer H, Faruque M, Doumatey A, Zhou J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Sigurdsson G, Hebebrand J, Pedersen O, Thorsteinsdottir U, Gulcher JR, Kong A, Rotimi C, Stefansson K. Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nat Genet* 2007;39:218-225.