Inherited smoking behaviour and Human Epididymis Protein 4 predict smoking-related morbidity and mortality

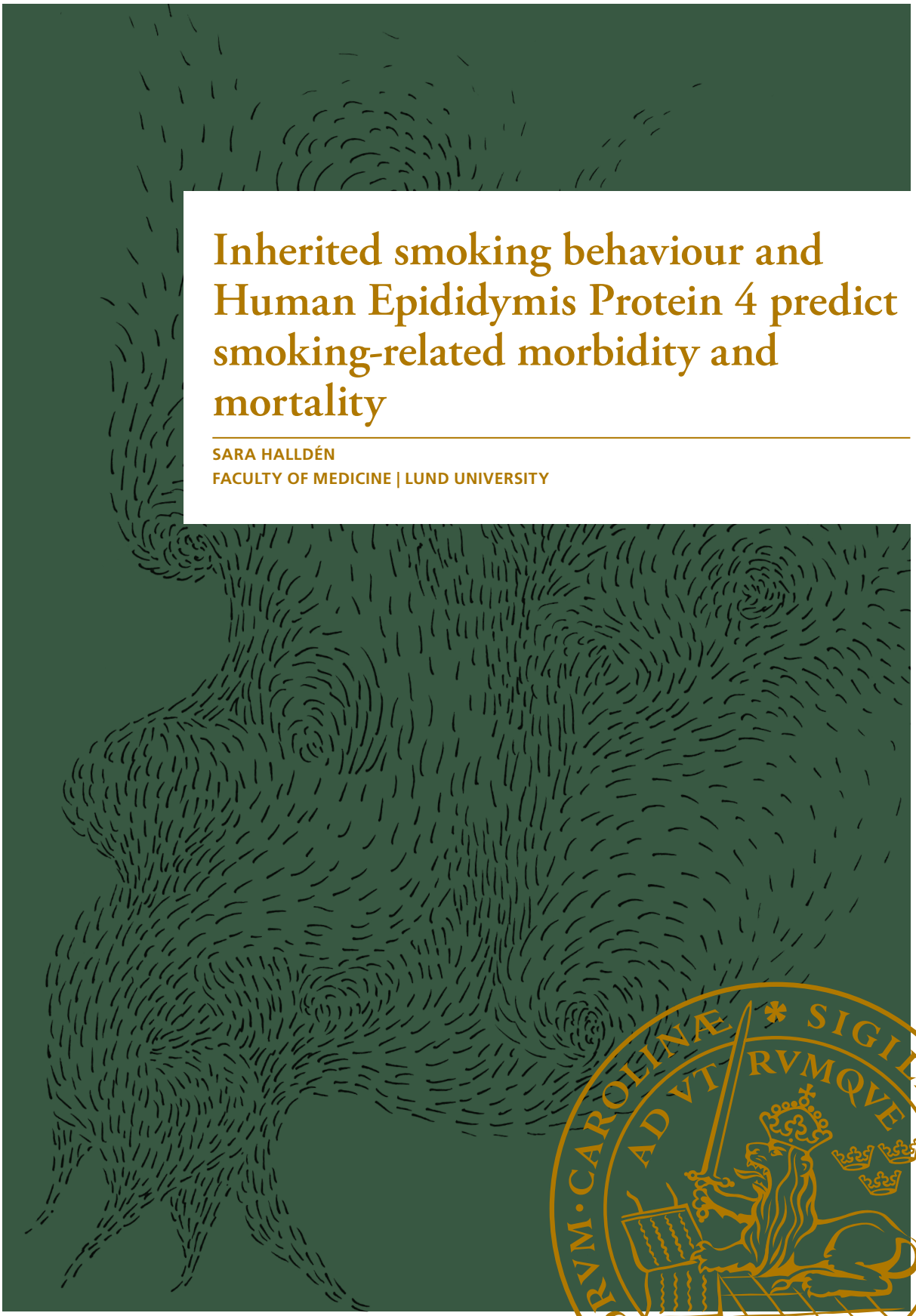
Halldén, Sara

2017

Document Version:
Publisher's PDF, also known as Version of record

Inherited smoking behaviour and Human Epididymis Protein 4 predict smoking-related morbidity and mortality

SARA HALLDÉN
FACULTY OF MEDICINE | LUND UNIVERSITY



Inherited smoking behaviour and Human Epididymis Protein 4 predict smoking-related morbidity and mortality

Sara Halldén

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at Kvinnoklinikens aula, Skånes Universitetssjukhus Malmö.
Date 8th of June at 13:00 p.m.

Faculty opponent
Professor Stig E Bojesen Department of Clinical Chemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Faculty of Health Sciences, University of Copenhagen

Organization	Document name
LUND UNIVERSITY	DOCTORAL DISSERTATION
Department of Clinical Sciences Malmö Faculty of Medicine	Date of issue 8th of June 2017
Author Sara Halldén	Sponsoring organization None

Title

Inherited smoking behaviour and Human Epididymis Protein 4 predict smoking-related morbidity and mortality

Abstract

GENERAL AIM: To investigate prognostic markers identifying smokers with increased risk of smoking-related diseases in the population.

BACKGROUND: Gene variance in the BDNF- respectively CHRNA-gene has been implicated in different smoking behaviours and the risk alleles have also demonstrated an additional risk increase of smoking-related diseases. The Human Epididymis Protein 4 (HE4) is a cancer biomarker that is also affected by active smoking.

AIMS: To replicate the associations between the genotypes of BDNF and CHRNA with smoking phenotypes and to test if the genotypes predicted events from tobacco-related diseases. To investigate the association between smokers and HE4 and to test if altered HE4 could predict future events of smoking-related mortality and morbidity in the population (Study III) or 90-day mortality (Study IV) in an acute setting.

SUBJECTS: Study I and II was based on the prospective, population-based cohort study of Malmö Diet and Cancer Study (MDCS)(n=30 447), and study III, on the sub-study MDC-Cardiovascular cohort (MDC-CC) (n=6102). In paper IV, a population presenting with acute dyspne at the emergency department of Malmö was investigated, ADYS (n= 963).

METHODS: Genotyping of the two polymorphisms, rs4923461 (BDNF) and rs1051730 (CHRNA) was performed in MDCS and correlated to smoking behaviour. In all four studies, participants were stratified according to smoking status and Cox-proportional hazard models were used to determine the correlations of the polymorphisms or the levels of HE4 to outcomes during follow-up.

RESULTS: The associations with smoking behaviour were confirmed for both genotypes. In current smokers, the risk alleles of BDNF and CHRNA significantly predicted all-cause mortality. CHRNA also predicted a higher risk of incident smoking-related diseases. No associations were seen in never smokers. In MDC-CC, HE4-levels were distinctly elevated in current smokers but not in the ADYS-population. In MDC-CC, elevated HE4 predicted all-cause mortality irrespective of smoking status and in ADYS, HE4 strongly predicted 90-day mortality regardless of underlying disease or smoking status.

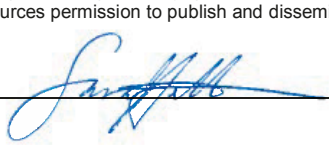
CONCLUSION: Gene variance of BDNF and CHRNA have impact on smoking behaviour and predicts an increased risk of smoking-related complications in smokers. Plasma levels of HE4 predicts mortality in a long- and short-term perspective and may be used as a disease risk marker in smokers and possibly also in non-smokers. In the future, genotype of BDNF and CHRNA polymorphisms and HE4 levels may be helpful in identifying patients with a higher risk of complications.

Key words

Smoking, nicotine dependence, GWAS, SNP, prediction, CVD, cancer, COPD, prevention, biomarkers, HE4, dyspnea

Classification system and/or index terms (if any)

Supplementary bibliographical information		Language English
ISSN and key title 1652-8220		ISBN 978-91-7619-483-6
Recipient's notes	Number of pages 91	Price
	Security classification	

Signature  Date: May 3rd 2017

Inherited smoking behaviour and Human Epididymis Protein 4 predict smoking-related morbidity and mortality

Sara Halldén, MD

Department of Clinical Sciences in Malmö
Faculty of Medicine
Lund University, Sweden

“We are the ones we´ve been waiting for. We are the change that we seek.”

Barack Obama 1961-

Cover by: all the way to paris®

Faculty of Medicine
Department of Clinical Sciences

ISBN 978-91-7619-483-6
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2017

To my mother

Content

Abbreviations

ADYS acute dyspnea study

BDNF brain-derived neurotrophic factor

BMI body mass index

BP blood pressure

CAD coronary artery disease

CHRNA nicotinic cholinergic receptor alpha

CI confidence interval

COPD chronic obstructive pulmonary disease

CPD cigarettes per day

CRP c-reactive protein

CVD cardiovascular disease

DBP diastolic blood pressure

DM diabetes mellitus

eGFR estimated glomerular filtration rate

GWAS genome-wide association study

HE4 human epididymis protein 4

HR hazard ratio

HDL high density lipoprotein

ICD international classification of diseases

IDI integrated discrimination improvement

LDL low density lipoprotein

LLT lipid lowering therapy

LUB lower, upper and bronchial airways (including the pleura)

MDCS Malmö diet and cancer study

MDC-CC Malmö diet and cancer study-cardiovascular cohort

MDRD modification of diet in renal disease formula

ND nicotine dependence

NRI net reclassification improvement

OC other cancer

OR odds ratio

PAD peripheral artery disease

SBP systolic blood pressure

SD standard deviation

SNP single nucleotide polymorphism

TRC tobacco-related cancer

WHO world health organization

Figures

Figure 4, page 22, is reprinted with permission from the American Society of Clinical Oncology.

Figure 5, page 22, is reprinted with permission from BMJ Publishing Group Ltd.

List of papers

This thesis is based on four studies presented in two published papers and two manuscripts, referred to in text by their corresponding Roman numerals. Permission for reprinting has been granted by the publishers.

I. Halldén S, Sjögren M, Hedblad B, Engström G, Narkiewics K, Hoffmann M, Wahlstrand B, Hedner T, Melander O. *Smoking and obesity associated BDNF gene variance predicts total and cardiovascular mortality in smokers*. Heart 2013 July; 99(13): 949-953.

II. Halldén S, Sjögren M, Hedblad B, Engström G, Hamrefors V, Manjer J, Melander O. *Gene variance in the nicotinic receptor cluster (CHRNA5-CHRNA3-CHRNB4) predicts death from cardiopulmonary disease and cancer in smokers.* Journal of Internal Medicine 2016 April; 279(4):388-398.

III. Halldén S, Nilsson J, Orho-Melander M, Engström G, Hamrefors V, Melander O. *Human Epididymis protein 4, a biomarker associated with cancer and smoking, predicts mortality and morbidity in the population regardless of smoking status.* Submitted spring 2017.

IV. Halldén S, Almgren P, Hamrefors V, Melander O. *Human Epididymis Protein 4 predicts 90-day mortality in patients with acute dyspnea.* Manuscript

Introduction

Smoking is an indisputable and avoidable risk factor for several lethal diseases and tobacco is the only legal substance that kills many of its users when used exactly as the manufacturer intends.

Worldwide, smoking is estimated to be responsible for 6 million deaths each year, many of them prematurely, which constitutes not only a social misfortune but also an economic burden for the nations´ health care systems[1,2].

Nicotine is a highly addictive and psychoactive agent[3-6] and an individual susceptibility[7] to nicotine as well a hereditary component on smoking behaviour[8-10] has been known for many years, now put into light again with genome-wide association studies (GWAS)[11]. Several genetic variants have been associated with nicotine dependency[12-18] and smoking behaviour[11,19-23] and interestingly, in some cases the same genetic variants are also associated with diseases related to tobacco[11,14,16,21,24-28]. In addition to gene variance, we investigated whether analysis of a protein named Human Epididymis Protein 4, known to be implicated in cancer prediction[29-33] and also affected by smoking[34-38], could serve as a marker for future smoking-related consequences.

Epidemiology is the study of the distribution and determinants of disease, and the aims of this thesis were to investigate whether smoking is equally dangerous to all and if not, how can we identify the smokers with the most benefit of smoking interventions?

Tobacco update

The prevalence of worldwide smoking fell during the last decade, 2000-10, which sends optimistic signals at first. Unfortunately, the rates are not as low as desired, and in addition, rapid progress of tobacco users are expected in Africa and eastern Mediterranean countries[39].

In 2003, The World Health Assembly adopted the World Health Organization Framework Convention on Tobacco Control (WHO FCTC), where the aims are to protect present and future generations from the deleterious consequences from

tobacco on health, environmental, social and economic factors. But so far, only a small fraction of the participating countries is on track to achieve their targets[39].

WHO have introduced MPOWER, which is short for six evidence based measures to reduce tobacco use and refers to:

M: Monitoring tobacco use and prevention policies.

P: Protecting people from tobacco smoke.

O: Offering help to quit tobacco use.

W: Warning about the dangers of tobacco.

E: Enforcing bans on tobacco advertising, promotion and sponsorship.

R: Raising taxes on tobacco.

According data from the OECD in 2014, 12% of the adults (+15 years) in Sweden smoked every day, in Denmark the proportion was 17% and in Norway 13%[40].

In the Western world, the harmful effects of tobacco are widely known for several years. The causal relationship with lung cancer was detected in the 1950s[41,42] and although convincing data, it was not an unchallenged fact until 20-25 years later, due to tough resistance from the tobacco industry.

But as mentioned, the highly addictive effects of nicotine probably explain the inability to quit smoking even when aware of the risks. Yet, an individual susceptibility for nicotine and a genotype effect on smoking behaviour has been known for many years, exemplified by twin studies[8,10]. Recently, the ability to perform genome-wide association studies (GWAS) have enabled identification of several genetic changes (polymorphisms) which associate with different smoking behaviours, supporting the idea of an individual susceptibility to nicotine[11,17,20-22]. In addition, some of the genetic alterations also associate with a higher risk of the harmful effects of smoking such as mortality and morbidity in tobacco-related diseases[11,21,27,28,43,44]. Assuming these associations are true, the scenario is intricate since smokers face different risks of smoking-related complications.

In addition to heritable factors, gender, occupation, education and anxiety disorders are also known to influence smoking status [45,46].

Second hand smoke is considered responsible for approximately 10% of the 6 million deaths[1]. There are also various sources of tobacco and the WHO-numbers on the different diseases are calculated on all tobacco, including "bidis" (a filter-less hand-rolled cigarette), cigars, chewed tobaccos and hookahs (vaporized tobacco). Yet, in this thesis, active cigarette smokers are in focus.

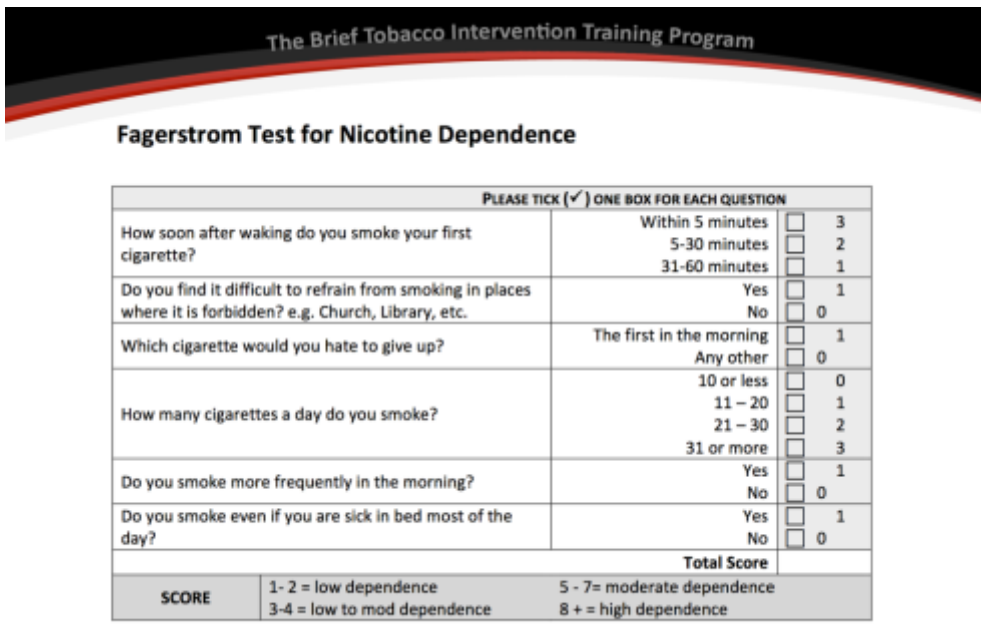
Smoking phenotypes

In epidemiological studies, smoking behaviour is often classified into phenotypes, for example tendency of smoking initiation (age at smoking start), former vs current smoker (ability to quit smoking), and smoking quantity (ND).

The data is usually self-reported: smoker yes/no/previous, and smoking quantity is self-reported as cigarettes per day (CPD). However, there are three other ways to measure levels of smoking [47]:

- Carbon monoxide (CO) in expired air, higher levels in smokers. Alternatively, carboxyhaemoglobin (COHb) measured in blood.
- Nicotine levels can be measured in urine, saliva and blood.
- Cotinine, a rest product of nicotine. Also measured in urine, saliva and blood[48,49].

Moreover, even though smoking quantity can be a surrogate measure of nicotine dependency (ND), there are a number of questionnaires developed for the purpose. One example is the Fagerströms test of nicotine dependency (FTND) (Figure 1) which classifies the dependency from 1-10 and is used in research as well as a guide for suitable nicotine replacement therapy (NRT). Another test is the Heaviness of Smoking Index (HSI) (Figure 2).



The Brief Tobacco Intervention Training Program

Fagerstrom Test for Nicotine Dependence

	PLEASE TICK (✓) ONE BOX FOR EACH QUESTION	
How soon after waking do you smoke your first cigarette?	Within 5 minutes 5-30 minutes 31-60 minutes	☐ 3 ☐ 2 ☐ 1
Do you find it difficult to refrain from smoking in places where it is forbidden? e.g. Church, Library, etc.	Yes No	☐ 1 ☐ 0
Which cigarette would you hate to give up?	The first in the morning Any other	☐ 1 ☐ 0
How many cigarettes a day do you smoke?	10 or less 11 – 20 21 – 30 31 or more	☐ 0 ☐ 1 ☐ 2 ☐ 3
Do you smoke more frequently in the morning?	Yes No	☐ 1 ☐ 0
Do you smoke even if you are sick in bed most of the day?	Yes No	☐ 1 ☐ 0
	Total Score	

SCORE		
	1- 2 = low dependence 3-4 = low to mod dependence	5 - 7= moderate dependence 8 + = high dependence

Figure 1. Fagerströms test for nicotine dependence questionnaire

Heaviness of Smoking Index

Use the following test to score a patient's level of nicotine dependence once they have been identified as a current or recent smoker

Please tick (✓) one box for each question		
How soon after waking do you smoke your first cigarette?	Within 5 minutes 5-30 minutes 31-60 minutes 60+ minutes	☐ 3 ☐ 2 ☐ 1 ☐ 0
How many cigarettes a day do you smoke?	10 or less 11 – 20 21 – 30 31 or more	☐ 0 ☐ 1 ☐ 2 ☐ 3
		Total Score
SCORE	1- 2 = very low dependence 3 = low to mod dependence	4 = moderate dependence 5 + = high dependence

Figure 2. Heaviness of smoking index questionnaire

Smoking cessation

Smokers that quit before developing tobacco-related diseases can to a large extent avoid the increased morbidity and mortality risk[50-53].Unfortunately, most of the smokers (~80%) who attempts to quit on their own, relapse within the first month[54] exposing the need for better guidance or treatments. The majority of medications available for smoking cessation today are all acting by modulating the nicotinic receptor pathways. The existing, approved medications in Sweden are nicotine replacement therapy (gums, patches etc.), vareniclin and bupropion in addition to therapy such as motivating interviewing [55]. A review of the efficacy of NRT: s was published in 2012 and concluded that NRTs are effective and increase the likelihood to quit smoking between 50-70%. Bupropion alone was not more effective than NRT but a combination was more effective than bupropion alone [56].

Smoking-related diseases

In Figure 3, the global distribution of non-communicable diseases (NCD) by cause of death is demonstrated. The attributable risk from smoking on cardiovascular diseases (CVD), cancer and respiratory diseases is more thoroughly discussed below.

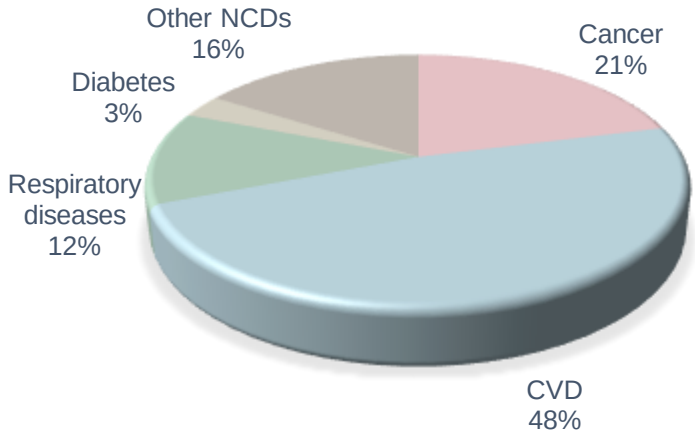


Figure 3. Distribution of global non-communicable disease (NCD) by cause of death, both sexes. (WHO)[57,58]

Cardiovascular disease

CVD constitutes one of our most complex diseases where lifestyle, genes and environment interplays to decide the onset and location of the cardiovascular pathology. Atherosclerosis is a pathology of primarily, the medium-and large-sized vessels in the cardiovascular blood system caused by a complex process involving the vascular wall[59].

In short, the process of atherosclerosis begins when the endothelium is exposed to elevated levels of LDL and other metabolites such as free radicals, which cause an injury and compensatory responses from the endothelial wall. LDL-particles are attracted and engulfed by monocytes which thereafter transform into macrophages (foam cells) and forms the atheromatous plaque covered by a fibrous cap. The plaque may proceed and rupture, exposing the leaking contents of the plaque to thrombogenic agents on the endothelium surface, and potentially a thrombus is formed[60].

Atherosclerosis is the result of various factors, including hypertension, diabetes, unhealthy diet, tobacco use, low physical activity as well as ageing and genetics. Obviously, some of these conditions are possible to influence thus, a large percentage of these diseases are preventable.

Tobacco use is one of the behavioural cardiovascular risk factors and is considered as attributable to 9% of all deaths globally, on second place after the “leading” cause, hypertension, attributable to 13% of all global deaths. Moreover, death from

myocardial infarction in young adults between the ages 30-44-years old is estimated to be attributable to smoking to 38%[61].

Smoking and atherosclerosis

The mechanisms behind smoking and CVD are also complex and not fully understood. What is known is that smokers have an impaired flow mediated dilatation (FMD) in the brachial and coronary arteries, which is considered to be a marker for atherogenesis and vascular dysfunction caused by reduced levels of nitric oxide[62,63]. Tobacco use also enhances the atherosclerotic process by a number of pro-atherogenic mechanisms[64]:

- increased total serum cholesterol, but lower LDL,
- physical damage to the endothelium,
- contains free radicals and oxidants which creates a pro-oxidative environment, promotes conditions for oxidatively modified LDL: s (the only LDL particles taken up by the macrophages and later form foam cells),
- increases plasma concentration of endothelial adhesion molecules attracting leukocytes,
- elevates inflammatory markers (leukocytes, CRP, TNF-α, IL-4, IL-6) and,
- creating a prothrombotic and procoagulative environment.

Cancer

The transformation of a normal cell into a cancer cell is a multistage process. Typically, there is a progress from a pre-cancerous lesion to a malignant tumour and interactions between genetic disposition and external agents such as UV-or ionizing radiation, asbestos, tobacco smoke, and infections from certain viruses, bacteria's or parasites, leads to the development of a tumour. As seen in Figure 3, cancer is the second most common cause of death from NCD: s in the world and the prevalence of cancer is rising due to the increasing amount of elderly in the population. The cancer forms causing most deaths (per year) world-wide are[65]:

- lung (1.69 million deaths)
- liver (788 000 deaths)
- colorectal (774 000 deaths)
- stomach (754 000)
- breast (571 000)

Smoking and cancer

Smoking tobacco is the single largest cause of cancer worldwide[66]. The complexity of the cigarette and the cigarette smoke can be manifested by its diverse compound where 3066 constituents of tobacco and 3996 from tobacco smoke have been identified[67]. By the year 2000, 69 of the constitutes of tobacco smoke was classified as carcinogens[68].

As mentioned earlier, the shift from "whether" tobacco smoking caused cancer to "how" and "what do about it" was already announced in 1954 by cancer authorities in the USA, the Netherlands and the Nordic countries. The change in action was based on lung cancer studies. Simultaneously, researchers from the tobacco industry drew the same conclusions but never admitted it publicly[69]. Today we know that lung cancer is the most common cancer in the world and 71% of all lung cancer death is attributable to tobacco use[70].

Through the years of frequent smoking it has become evident that smoking is linked to other tumours as well, and in the last update in 2004 the International Agency for Research on Cancer (IARC) listed up to 14 different cancers associated with smoking.[70] In 2009, colon cancer and ovarian cancer was added to the list[71].

However, not all cancers are associated with smoking, see Table 1.

Table 1. Evidence for carcinogeneity of tobacco smoking. Adapted from IARC[71]

	Tumour site
Sufficient evidence	Oral cavity, oropharynx, nasopharynx, and hypopharynx, oesophagus (adenocarcinoma and squamous-cell carcinoma), stomach, colo-rectum,* liver, pancreas, nasal cavity and paranasal sinuses, larynx, lung, uterine cervix, ovary (mucinous)*, urinary bladder, kidney (body and pelvis), ureter, bone marrow (myeloid leukaemia)
Limited evidence	Female breast
No evidence	Endometrial (postmenopausal), thyroid, skin (including malignant melanoma), nervous system, malign lymphoma, multipel myeloma, testis, soft-tissue sarcoma

Tobacco-related cancer

The tumours differ in to what extent they depend on tobacco smoking. By calculating the different population attributable fraction (PAF) of separate tumours in a European population, Agudo et al concluded that around 270 000 new cancer diagnoses per year are attributable to tobacco smoking, see Figure 4[72]. The PAF method is a measure of the health burden of a risk factor (here smoking) in a population and estimates the proportion of disease that would not occur in absence of exposure to the risk factor.

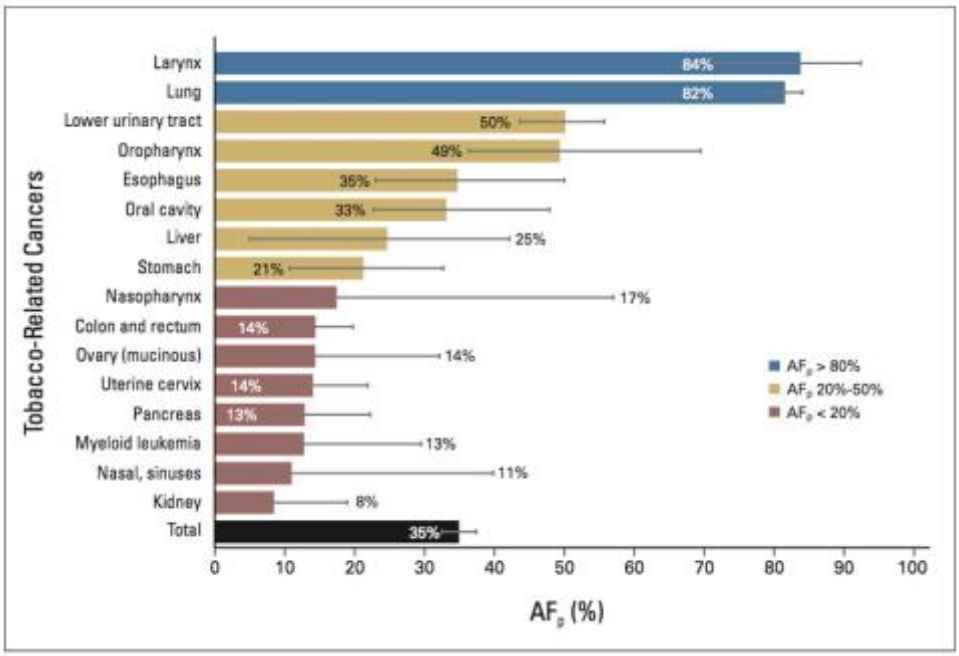


Figure 4. The population attributable fraction (PAF) och tobacco-related cancers based on the EPIC[72]

In figure 5, the causal relationship between death from lung cancer and smoking is shown in addition to the benefits of quitting smoking.

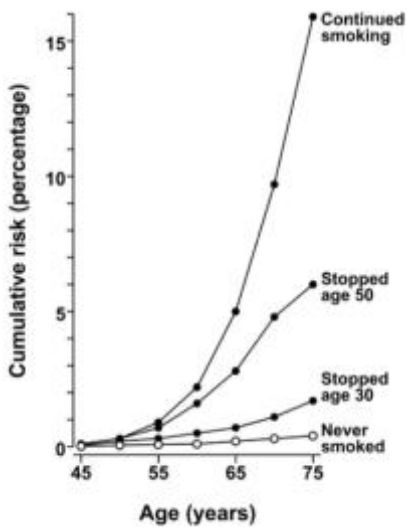


Figure 5. Cumulative risk (%) of death from lung cancer in men at ages 45-75 years: in continuing cigarette smokers,former smokers (stopped at age 50 or 30) and in never smokers in the UK 1990.[51]

Respiratory diseases

The group of respiratory diseases consist of various diseases including chronic and acute diseases as well as both NCD: s and Communicable diseases (CD: s) with alternate associations with tobacco. A more thorough list of the included diseases (ICD codes) is available in the Methods section, p. 33.

However, in the group of respiratory diseases, the proportion of death rates attributable to tobacco worldwide in adults aged 30 and older is 36% , mainly driven by death from COPD where the proportion is approximated to 42%[73]. Moreover, within the group of respiratory diseases, COPD is the major cause of morbidity and mortality worldwide[74]. Smoking is considered to be the most important single causal factor for developing COPD but also factors like outdoor pollution can have impact[75].

Basic genetics and genetic variance

A gene is a section of the DNA and codes for traits, all genes come in pairs and one gene in a pair is called an *allele*, i. e there are two alleles per gene.

The variation in DNA sequence from one individual to another is called *genetic variance* and the most common mutation (change in the DNA sequence) is a single base pair substitution called single nucleotide polymorphism (SNP). More than 38 million SNPs with a frequency of more than 1% were identified in the 1000 Genomes Project[76]. SNPs that occurs in at least 1% of the alleles in a population are referred to as common variants, rare variants occurs in less than 1%[77]. Most of the SNPs are in linkage disequilibrium (LD), meaning that they are plausible to be inherited together. Importantly, the variants discovered in a GWAS are not sufficient to capture a total heritability of polygenetic traits [78,79].

The Hap Map and GWAS

The DNA sequence between any two human beings is 99.5% identical, but the variations may have a large impact on the persons individualized risk of disease. The Hap Map project[80,81] is a genome-wide map of SNPs based on ancestral chromosome segments from four populations. This data base of common variations enabled a new form of studies on genetic associations, GWAS.

In short, GWAS is a hypothesis-free approach that compares the genotype between two populations with different phenotypes and investigates the association between SNPs and the phenotype, identifying loci with SNPs that differs. Furthermore, the

odds ratios (OR) of the probability of phenotype for each genotype relative to the whole population can be calculated. This may be used for "personal genomics" since presence of the SNP and the OR can be used to predict the probability (or risk) of a person to have the phenotype.

A GWAS can also identify candidate genes for a phenotype (which may be a disease or for example a smoking behaviour) by the position of the SNP and provide hypothesis-driven studies. This type of research requires large series of cases and controls to obtain statistical power, and clinical studies are of interest to further investigate the findings.

Genetics and disease

Most diseases in the population as well as the ones studied in this thesis, are common, complex diseases. The underlying causes are multifactorial meaning that they depend on several genetic variants at various loci, in combination with environmental factors. In comparison to rare, mendelian diseases, where rare gene mutations have great impact on disease susceptibility, the impact of one gene in the development of a common disease is to be considered small[82].

BDNF

rs4923461 and rs6265

The SNP investigated in paper I is the rs4923461 on the gene coding for brain-derived neurotrophic factor (BDNF), located on chromosome 11. The rs4923461 had been the lead SNP in a GWAS for increased BMI[83] and short thereafter a GWAS from the Tobacco and Genetic consortium reported that BDNF rs6265 was the SNP with the strongest association to smoking initiation[20]. The rs6265 and rs4923461 are both common variants and are in almost complete LD, meaning that either SNP can be analysed and function as a tag-SNP for the phenotype reported.

Even though the molecular pathway is not fully understood, a plausible theory is that genetic BDNF variation alter the rewarding effects of nicotine and promote continued use. Likewise, the precise mechanism behind higher BMI and BDNF is not entirely clear but low signalling of BDNF leads to hyperphagi and obesity. For example, heterozygous BDNF knockout mice consumed almost 50% more food than their wild-type littermates and were obese. After infusion of BDNF in the wildtype mice their body weight and food intake decreased[84].

Background of BDNF

BDNF is part of a neurotrophin family and plays a critical role in regulating protective mechanisms of neurons, such as survival, function, development, and plasticity of the cell.[85-88] BDNF is distributed in key regions of the central nervous system regulating mood and behaviour, resulting in extensive studies related to several psychiatric disorders[89]. Associations with substance-related disorders, eating disorders, and schizophrenia have been confirmed in a meta-analysis of case–control studies [90] and by acting upon the reward system of the brain BDNF has been suggested to play a role in drug addiction[91]. BDNF is also expressed in vascular endothelial cells[92], macrophages and smooth muscle cells in the atherosclerotic coronary arteries[93].

BDNF in plasma or serum

The protein BDNF can also be measured in plasma or serum and the associations with smoking, CVD and circulating BDNF are not fully understood with conflicting data reported. Jamal et al found that smokers had increased serum levels of BDNF compared to non-smokers but no interaction was seen between serum levels and genotype (rs6265)[19]. Kim et al measured plasma levels of BDNF and found lower concentrations in smokers than in non-smokers. Repeated analysis after 2 months of unaided smoking cessation showed an increase in the levels of BDNF[94]. Kaess et al reported lower risk of CVD and total mortality in patients with higher levels of BDNF in a prospective study, smoking did not affect the results at all[95]. In contrast, Ejiri et al reported higher BDNF levels in patients with unstable angina compared to those with stable angina, suggesting a role of BDNF in plaque instability[93].

CHRNA

rs1051730, rs16969968 and the cluster of nicotine receptors

The polymorphism investigated in paper II is located in a cluster of nicotine receptors (CHRNA) on chromosome 15, 15q25. In 2007, the rs16969968 on CHRNA5 was identified in a candidate gene study as associated with nicotine dependence[12,13]. The SNP correlated to ND measured with Fagerströms score in a study by Chen at al[15] but the finding was not replicated in a GWAS by Hung et al who on the other hand found a strong association with lung cancer irrespective of smoking status [24]. Amos et al[44] reported a strong association with lung cancer but just a weak effect on smoking behaviours. Moreover, other GWAS identified the rs1051730 on CHRNA3 to associate with smoking quantity, peripheral artery disease (PAD) and lung cancer[21,25]or with COPD[43], respectively. In later studies the cluster is also reported to be associated with a higher risk of schizophrenia and bipolar disorders[96,97].

ND, smoking phenotype and functional variant

The SNPs on 15q25 are all in strong LDs (in populations of European ancestry) and other SNPs in the receptor cluster have also reported to be associated with heavy smoking and ND[98]. The rs16969968 is a missense mutation coding for a functional variant where aspartate substitutes to asparagine in the α5 subunit protein. In vitro studies have demonstrated that the switch of the amino acid confers a change in maximal response to nicotine. The subunits with asparagine (the risk variant associated with ND and smoking quantity) were less responsive to nicotine than the receptor complexes with aspartic acid[99]. The rs1051730 SNP is not likely to be of any functional importance since it is a coding, synonymous variant that does not confer any change in amino acids. Nevertheless, the rs1051730 seems to be associated with lower levels of expressed CHRNA5 in the brain and in peripheral blood mononuclear cells[96].

Background of CHRNA

There are two kinds of acetylcholinic receptors (AChR), the nicotinic (NAChR) and the muscarinic (MAChR) receptor. The nicotinic receptors are divided into neuronal and muscle receptors, although they can be present in both non-neural and non-muscle-tissues. The receptors consist of different subunits coded by genes named CHRNA1-7, CHRNA9-10 and CHRNB1-4. The nAChRs are present throughout the nervous system, in bronchial and alveolar epithelium, vascular tissues, lymphocytes, pulmonary neuroendothelial cells as well as in small cell lung cancer (SCLC) cell lines. The receptors seem to play a role both in nicotine dependence and in the pathogenesis of disease/carcinogenic pathways, and the expression of receptor subunits differed between normal tissue compared to tumour tissues, interestingly the distribution also varied with smoking status[100].

In addition to the distribution of receptors, they may also respond differently when binding nicotine. In short, when nicotine (and other nitrosamines from tobacco smoke) binds to the certain nicotine receptor it has been found to stimulate malign lung cell lines[101,102] and block apoptosis in both cancer and normal cells, suggesting that nicotinic receptors are modulating pathways and could function as targets for future cancer therapy[102].

Knowledge of the receptor mechanisms is already in therapeutic use as he target receptor of the partial agonist vareniclin[103] is α4ß2 since it is thought to be responsible for nicotine addiction[104,105].

HE4

The role of the biomarker Human epididymis protein 4 (HE4) and smoking is investigated in paper III and IV. HE4 is a protein named after the tissue where it first was discovered[106]. Since then, the protein has been discovered in numerous tissues including male and female reproductive tracts, respiratory tract, kidney and in several tumour lines such as ovarian, lung, breast, colon and renal cell lines[33,34,37,38,107,108]. Elevated levels of HE4 have also been detected in patients with fibrotic kidney disease[109] and heart failure[110,111], suggesting a role as a mediator of fibrosis. In healthy populations, HE4 increases with age, male sex, smoking and decreased kidney function[34-38].

HE4 is also termed WFDC-2 (whey acidic protein-four-disulphide core-2), and belongs to a protein family with a highly conserved WAP domain[112], containing a protein structure which has suggested anti-protease activity[106]. Other functions linked to this protein family include immunomodulatory and anti-microbial properties, but the function of the HE4 protein is yet unknown.

Clinically, well-founded HE4 research has been done on ovarian cancer, and in the US, the Food and Drug Administration (FDA) has recently approved HE4 to be analysed as an aid in monitoring the occurrence or progression of disease in patients with epithelial ovarian cancer. HE4 is also used together with the protein CA125 in the Risk of Malignancy Algorithm (ROMA), a test to aid in assessing whether a patient presenting with an adnexal mass is at high or low risk of finding malignancy on surgery[30-32,113,114].

The mechanism behind elevated HE4 levels in smokers is not known.

Dyspnea

The symptom of dyspnea can be defined as “shortness of breath” and/or “laboured or difficult breathing” and is a common symptom at the emergency ward. The underlying disease varies from myocardial infarction, pulmonary embolism, COPD, heart failure, infection, anaemia, myasthenia gravis, asthma or anxiety[115-117]. Consequently, the character of dyspnea can be more or less severe and the underlying condition can range from high mortality risk to very low where the extent of dyspnea not always correlates with the mortality risk conferred by the responsible disease. Several studies have investigated biomarkers with prognostic qualities when the underlying cause of dyspnea is known[116,118-121].

In paper IV, the biomarker HE4 was measured in patients presenting with acute dyspnea at the emergency department (ED) in Malmö, Sweden.

Aims

- Given the known association with smoking behaviour and BMI, the purpose was to test the hypothesis that genetic variation in the BDNF locus alter the risk of smoking-related complications such as total and cancer mortality, among smokers in the Malmö Diet and Cancer Study (MDCS), a population based prospective cohort study. (Study I)
- The CHRNA polymorphism had been associated with nicotine dependency and smoking quantity as well as with lung cancer, COPD and peripheral artery disease. We aimed to replicate the associations but also extend to analysis between the SNP and mortality outcomes in the MDCS. (Study II)
- To test whether the diagnostic cancer marker Human Epididymis Protein 4 (HE4) is associated with smoking and whether it predicts risk of tobacco-related diseases and mortality in the population (MDC-CC). (Study III)
- In this study, we aimed to investigate whether HE4 associated with smoking and could predict short-term mortality in patients presenting symptoms of acute dyspnea. (Study IV)

Methods

Study populations

The study protocols of Malmö Diet and Cancer Study (MDCS) and Acute Dyspnea Study (ADYS) were approved by the ethical committee at the University of Lund. All participants provided written consent.

Malmö Diet and Cancer Study (Paper I and II)

The MDCS was designed as a 10-year prospective case-control study and there were two initial purposes. One was to investigate the association between western diet and certain cancer forms taking other life-style factors into account, and another whether oxidative stress and DNA-repairing systems had influence on the impact of diet on the development of cancer forms. Moreover, the collected data was also to be used as a resource for testing future hypothesis[122].

Recruitment was carried out between 1st January 1991 and 25th September 1996. Since the main interest was cancer, the recruitment focused on obtaining sufficient number of cancer cases. Initially, invitation letters were sent to all subjects living in Malmö the 1st January 1991, born between 1926-1945, which was a total of 53 325 individuals. In 1995, the recruitment was extended to invite female inhabitants born between 1923-1950 and to male inhabitants between 1923-1945. The point of inviting somewhat younger women was to obtain adequate cases of breast cancer in premenopausal women.

Invitations were also published in local newspapers and in public areas such as primary care centres, resulting in a recruitment of 5505 participants who had not received an invitation letter. A total of 30 447 participants completed the anthropometric measurements and is the cohort referred to in this thesis. Furthermore, they filled in a self-administered questionnaire about education, occupation, social network, physical activity, use of tobacco and alcohol, current health, current medication and disease in close relatives. Women were asked about reproductive history. Blood samples were collected and stored in -80°C. Exclusion criteria were language problems or mental retardation.

The participants in the MDCS and the non-participants (n=40 807) were later compared with regards to cancer incidence and mortality, to declare potential selection bias when interpreting the results in cohort studies. Participants and non-participants did not differ according to sociodemographic structure and smoking status was similar, but non-participants had a higher mortality during the recruitment period as well as during follow-up, than the participants[123].

MDC Cardiovascular Cohort (MDC-CC) (Study III)

Every second subject who entered the MDCS from November 1991 to February 1994 (n=12 445) was invited to participate in a sub-study on the relationship between carotid atherosclerosis and cardiovascular risk factors. A total of 6103 subjects accepted and formed the Malmö Diet and Cancer Cardiovascular cohort (MDC-CC). The subjects from MDC-CC had an ultrasound examination of the right carotid and provided additional fasting blood samples[124].

ADYS (Study IV)

The study of patients with acute dyspnea (ADYS) was performed at the Emergency Department of Skåne University Hospital of Malmö, Sweden. The catchment-area is approximately 400 000 and the emergency ward takes around 85 000 visits per year, where dyspnea is the main cause of the visit in around 7%[125] .

During the years 2013-2016, adult patients presenting with acute dyspnea were asked to participate. The enrolment took place during office hours (06.45 AM to 4.30 PM) when between 1 and 3 research nurses were present. Clinically relevant information about health background, medication and parameters such as saturation, Medical Emergency Triage and Treatment System-Adult Score (METTS-A)[126], CRP, heart rate etc. were documented and plasma samples were secured and frozen. The research nurses also searched relevant patient data from the medical records from the University Hospital of Skåne.

The research examination process of ADYS was divided into two sessions. Between 6th of March 2013 and 25th of May 2014, 439 patients were enrolled. Next registration process started 15th of January 2015 to 20th of January 2016 and 524 patients were enrolled.

The exclusion criterias were deficient Swedish civil registration number or not being able to give consent.

End points

The assessment of end points was similar in all four studies. Information of mortality end points during follow-up was retrieved by linking the national civil digit number of each participant to the Swedish Cause of Death Register[127], information on the various incident diseases during follow-up was retrieved from Swedish Hospital Discharge Register, the Stroke in Malmö register, the Swedish Cancer Registry and the Swedish Coronary Angiography and Angioplasty Registry. These registers have been described in detail and validated for classification of outcome[128-130].

Coronary Artery Bypass Graft (CABG) was identified from the national Swedish classification systems of surgical procedures, the Op6 system from 1963 until 1997 and the KKÅ system since then. End points from all four studies are defined below and defined on the basis of International Classification of Diseases (ICD) codes.

Study I

Four end points were examined:

- **Total mortality**: Death from any cause.
- **CVD mortality**: when the main ICD code was 390–459 (ICD 9) or I00-I99 (ICD 10).
- **Cancer mortality**: ICD codes 140–239 (ICD 9) or C00-C99 (ICD 10).
- **First incidence of CVD**: CVD was defined as fatal or non-fatal myocardial infarction (MI) or stroke or death due to ischemic heart disease from the Swedish Hospital Discharge Register or SNCDR. MI was defined as codes 410 (ICD9) or I21 (ICD10), death due to ischemic heart disease as codes 412 and 414 (ICD9) or I22-I23 and I25 (ICD10), and stroke as codes 430, 431, 434, and 436 (ICD9) or I60-I61, I63, and I64 according to ICD10.

Study II

Eight end points were examined:

- **Incident COPD**: 490–496 (ICD9) or J40–44 (ICD10).
- **Incident tobacco-related cancer (TRC)**: cancer of the oral cavity [ICD seventh revision (ICD7) code 140–144], oropharynx (ICD7 145, 147 and 148), nasopharynx (ICD7 146), oesophagus (ICD7 150), stomach (ICD7 151), colon (ICD7 153), rectum (ICD7 154), liver (ICD7 155 and 156), pancreas (ICD7 157), nose and sinuses (ICD7 160), larynx (ICD7 161), lung (ICD7 162 and 163), uterine cervix (ICD7 171), ovary (ICD7 175), kidney (ICD7 180) and lower urinary tract (ICD7 181) and myeloid leukaemia (ICD7 205).

- **Incident other cancer (OC)**: included cancers of the breast (ICD7 170), prostate (ICD7 177), skin including malignant melanoma (ICD7 190–191) and nervous system (ICD7 193) and malignant lymphoma (ICD7 200–201).
- **Incident CVD:** (see Study I).
- **Total mortality:** (see Study I).
- **CVD mortality**:(see Study I).
- **Cancer mortality**: (see Study I).
- **Respiratory disease mortality**: codes 460–519 (ICD9) or J00-99 (ICD10).

Study III

Seven end points were examined:

- **Incident coronary artery disease (CAD)**: first incident coronary event of either fatal or nonfatal myocardial infarction [codes 410 (ICD9) or I21 (ICD10)], death due to ischemic heart disease, [codes 412 and 414 (ICD9) or I22**–**I23 and I25 (ICD10)], percutaneous coronary intervention or coronary artery bypass grafting (CABG). CABG was defined as a procedure code of 3065, 3066, 3068, 3080, 3092, 3105, 3127, 3158 (Op6) or FN (KKÅ 97). PCI was defined based on the operation codes FNG05 and FNG02.
- **Incident TRC**: (see Study II).
- **Incident OC**: (see Study II).
- **Incident lower, upper and bronchial airways, including the pleura (LUB):** ICD7 code 162.
- **Total mortality**: (see Study I).
- **CVD mortality**: (see Study I).
- **Cancer mortality**: (see Study I).

Study IV

One end point was examined.

- **90-day mortality**: Date of death was registered from a national population register.

Analysis and Statistics

SPSS versions 19-23 were used for all calculations (IBM Corporation, New York, NY, USA). C-statistics and net reclassification improvement in study IV were performed using STATA (vs 13.1 corp. Texas) and R 3.3.1.

Genotyping in study I and II was performed using TaqMan (Applied Biosystems) with primers and conditions according to the manufacturer's recommendation.

In Study III and IV frozen plasma samples of the biomarker HE4 were analyzed by Proseek Multiplex Oncology II panel (Olink Bioscience, Uppsala, Sweden). The method is a multiplex immunoassay based on a Proximity Extension Assay and the data is presented as normalized protein expression (NPX) where a high value corresponds to a high protein concentration, but not to an absolute value. Intra-and inter-assay coefficient of variation was 7% and 13% respectively.

Dichotomous variables are reported as numbers (%) and continuous variables as mean (SD) in all papers and a two-sided p value of <0.05 was considered statistically significant.

Study I

Genotyping of rs4923461 was successfully performed in 27 508 out of 28 564 subjects, success rate 96%. Complete data for age, sex, BMI, and smoking status was available in 25 789 individuals. This cohort was used in analysis of ever smoker and the polymorphism.

Since the SNP was associated with smoking initiation we decided to analyse current, ever and non-smokers respectively. Current smokers (n=7225) included participants smoking regularly (n=6057) or sometimes (n=1168) within the past year. Non-smokers (n=18 564) were former (n=8791) and never smokers (n=9773). Furthermore, the subgroup ever smokers (n=16 016) was formed and included former and current smokers. Complete data for age, sex, BMI, smoking status and cigarettes per day (current smokers) was available in 25 071 participants and this cohort was used in the analysis between the genotype and the different end points.

Cross-sectional relationships between genotype and smoking status were evaluated with crude and multivariate (age and sex) adjusted logistic regression. Relationship between genotype and BMI was tested with crude and multivariate (age and sex) adjusted linear regression models. The SNP was coded additively (GG=0, AG=1, AA=2) in all analyses.

Stratified by smoking status, we calculated crude and multivariate adjusted hazard ratios (HR) and 95% confidence intervals (95% CI) for the genotype in relation to the incidence of the four different end points during follow-up using the Cox proportional hazards model.

The multivariate models were:

1. Age and sex
2. Model 1+BMI
3. Model 1+2 +CPD (only in analysis of current smokers)

Study II

Genotyping of rs1051730 was performed in 26 471 out of 28 564 (93%). Complete data for age, sex, BMI, smoking status, hypertension (SBP ≥140 mmHg or DBP≥90 mmHg and/or ongoing antihypertensive therapy), LLT and previous diabetes was available for 24 794 subjects (see Figure 1 in Paper II).

This particular SNP was associated with smoking quantity, and we decided to separately analyse current (n=6951), former (n=8426) and never (n=9417) smokers respectively, as well as ever smokers.

The SNP rs1051730 was coded additively (CC=0, CT=1, TT=2) in all analyses. Cross-sectional relationships between the genotype and smoking status were evaluated using logistic regression. Cox-proportional hazard models were used to calculate crude and multivariate adjusted HR: s and 95% CI: s for rs1051730 in relation to first event of each end point during follow-up.

The multivariate models were:

1. Age and sex
2. Model 1+BMI, hypertension, previous DM, LLT
3. Model 1+2+ CPD (only in analysis of current smokers)

Study III

The distribution of the biomarker HE4 was skewed and the natural logarithms were derived. The log-transformed values were expressed on a standardized scale (zscore), per 1-SD increment in the cohort and in each smoking strata respectively.

Cross-sectional analyses were performed with baseline data to investigate associations between HE4, smoking and relevant variables. Linear regression models were used with standardized logarithmic HE4 as the dependent variable. Cox-proportional hazard models were used to calculate HR: s and 95% CI: s for HE4 in relation to first incident event of each endpoint during the follow-up.

Since the end points studied demanded different adjustments we created two different analysis arms. The **"CVD-arm"** for CVD-related outcomes also including total mortality and the "**Cancer-arm**" for cancer-related outcomes.

Pack-years was used as a measure for tobacco burden, calculated as (number of cigarettes per day/20) x years of smoking.

Three models were created in each arm:

CVD-arm (n= 4614):

1. Age and sex
2. Model 1+BMI, hypertension, diabetes (fasting b-glucose>6.0 and/or diagnosed diabetes), HDL, LDL, and eGFR (MDRD-formula)
3. Model 1+2+current smoking (in all) or pack-years (in current smokers).

Cancer-arm (n = 4671):

1. Age and sex
2. Model 1+ BMI
3. Model 1+2 +current smoking (in all) or pack-years (in current smokers).

Study IV

The research examination process of ADYS (see Study population, ADYS on p. 32) was divided into two sessions, we have named the first ADYS-Discovery (**ADYS-D**) and the later ADYS Replication (**ADYS-R**). The ADYS-D enrolled 439 patients, 60 was lost due to missing covariates in the current study, leaving 379 patients for analysis. ADYS-R enrolled 524 patients where 104 lacked data on all covariates, leaving 420 patients for analysis.

Blood biomarkers were first measured in ADYS-D followed by replication attempts in ADYS-R. As HE4 was measured both in ADYS-D and ADYS-R, the two data bases were subsequently merged, resulting in 799 patients remaining for analysis with data on all covariates, referred to as ADYS-Pooled (**ADYS-P**). To investigate the accuracy of the laboratory measurements, 49 plasma samples from participants in ADYS-D were re-analyzed with the samples of ADYS-R. The correlation was tested by using Spearman's rank correlation test with a correlation of 0,9.

The distributions of the two biomarkers in this paper, HE4 and CRP, were skewed, why the natural logarithms were derived to achieve statistical normality. The log transformed values were expressed on a standardized scale in the included participants respectively stratified by smoking status and also ranked and ordered into quartiles.

To test associations between HE4, smoking and covariates, cross-sectional analyses were performed with baseline data, using linear regression models with standardized logarithmic HE4 as the dependent variable.

Cox proportional hazard models were created to obtain HR: s and 95% CI: s for standardized HE4 towards the end point 90-day mortality. Day 1 of follow-up was the date of the emergency visit that led to study participation, end of follow-up was

occurrence of death or end of the study, i. e. 90 days after presentation. The Cox regressions were analyzed in the cohort (“all”), ever smokers and in never smokers in ADYS-D, ADYS-R and ADYS-P, respectively. The models were adjusted for age, sex, respiratory rate (breath/min), oxygen saturation (%), METTS-A and CRP(mg/L).

To evaluate the potential of HE4 in risk prediction at the ED, we assessed discrimination using C-index[131] and reclassification with continuous and categorical NRI[131] comparing the basic model with the traditional risk factors (age, sex, oxygen saturation, respiratory rate, CRP, and METTS-A) with the same model adding quartiles of HE4.

In categorical NRI, we defined ”very low risk” as a predicted risk of death within 90 days of <1%, and assessed movements across this clinical threshold when adding HE4 to traditional risk factors.

Results

Study I.

Carriers of a BDNF risk allele (A) are more likely to be smokers, and if they smoke they are at higher risk of dying than other smokers.

Follow-up extended until 1 January 2007. Mean ± SD follow-up in analysis of mortality and analysis of first incident CVD was 12±3 years. Baseline data for study participants are presented in Table 2 and the genotype distribution in Table 3. In Tables 4 and 5, multivariable adjusted HRs *per genotype* is presented using the GG genotype as a reference (HR: 1.0) and the p-value is for trend (from the additive models).

There was no significant deviation from Hardy-Weinberg equilibrium in any of the groups studied, P>0.05.

Table 2. Baseline characteristics study cohort

	Female	Male	All
Participants, n(%)	15 665(61)	10 124(39)	25 789(100)
Age, years	57±8	59±7	58±7
BMI, kg/m^2	25±4	26±4	26±4
Current smokers	4450	2865	7225
CPD	13±7	16±9	14±8

Table 3. Distribution of rs4923461 in the study population and stratified by smoking status

Genotype	G/G	G/A	A/A
All[1] (n=25 789)	16 414 (63.5%)	8375 (32.5%)	1000 (4.0%)
Ever smokers (n=16 016)	10 256 (64.0%)	5171 (32.5%)	589 (3.5%)
Non-smokers (n=9773)	6158 (63.0%)	3204 (33.0%)	411 (4.0%)
All[2] (n=25 071)	15 959 (63.5%)	8146 (32.5%)	966 (4.0%)
Current smokers (n=6507)	4192 (64.5%)	2074 (32.0%)	241 (3.5%)
Non-smokers (n=18 564)	11 767 (63.5%)	6072 (32.5%)	725 (4.0%)

1. In cohort with complete data for age, sex, rs4923461, BMI and smoking status. 2. In cohort with complete data for age, sex, rs4923461,BMI, smoking status and smoking quantity.

rs4923461, smoking initiation and BMI

The risk allele (A) of rs4923461 had been associated with smoking initiation[20] and high BMI[83,132] and in the MDCS we found a significantly increased odds ratio between the SNP and ever having smoked (OR:1.05, 95%CI:1.00-1.10; p=0.03). The association was still significant after adjustment for age, sex and BMI (1.05; 1.01-1.10;0.03). The BMI-association was also replicated in the MDCS (β=0.15±0.04; p=0.001).

Association with smoking-related diseases in current smokers

Mortality end points

During follow-up, 1049 (16.1%) deaths from all-cause occurred. We used additive models with the risk-allele (A) coded, which showed that each copy of the risk allele associated with a significantly increased risk of **total mortality** in the three models. Model 1 was adjusted for age and sex (HR1.13, 95%CI 1.01-1.26) and model 2 additionally for BMI (1.13, 1.01-1.26). In current smokers, a model 3 was analysed with additional adjustment of smoking quantity and the association was still significant (1.12, 1.00-1.25).

Death from CVD occurred in 346 (5.3%) participants during follow-up. Presence of an A-allele resulted in significantly higher HRs (same adjustments as above):

- model 1: 1.25,1.02-1.52
- model 2: 1.24,1.01-1.50
- model 3: 1.23,1.01-1.49

During follow-up, 492 (7.6%) **deaths from cancer** were reported but no significant association was seen in any of the models (model 1:1.14, 0.97-1.34). In the genotype analysis, a significant association was seen among carriers with two risk alleles (AA) compared to the reference with no risk allele (GG), see Table 4.

First incidence in CVD

A total of 802 (12.7%) first **events of CVD** were registered during follow-up. No association was present in the additive models (model 1: 1.09, 0.96-1.24)

Association with smoking-related diseases in non-smokers

Mortality end points

In this group reporting never or previous smoking, 1871 (10.1%) **deaths from all cause** occurred, 603 (3.2%) **CVD deaths** and 877 (4.7%) **cancer deaths** during follow-up. No significant association with rs4923461 was seen in any end point, Table 5.

The group of previous smokers (n=8791) was analysed separately but no associations were seen in this subgroup either.

First incidence in CVD

No significant association was seen in non-smokers or in previous smokers, 1606 (8.9%) events occurred during follow-up (non-smokers).

Additional analyses

Since smoking and its adverse effects are associated with socioeconomic status (SES)[45,133,134] we also adjusted for level of education. The variable was defined by the highest level of education and the participants were divided into three groups. The low SES group had not completed elementary school, corresponding to a maximum of 6–8 years of education; the middle SES group had 9–12 years of education; and the high SES group reported a university degree or studies at least 1 year after GCE (General Certificate of Education).

The variable had little impact on the results why it was not included in the text or tables.

Data on the end point incident tobacco-related diseases was available for analysis after the publication of this paper. A complete analysis in the present study population with identical models stratified by smoking status was calculated, without any significant association between rs4923461 and incident TRC.

Table 4. Multivariable-adjusted HRs (95%CI) per BDNF genotype in current smokers

Genotype	GG	AG	AA	
End point **Events (%)**	HRs 95% CI			*P*trend
Total mortality **1049 (16.1)**				
Model 1	1.0 (ref)	1.09 (.78-1.53)	1.22 (.89-1.73)	<.05
Model 2	1.0(ref)	1.09 (.78-1.53)	1.22 (.89-1.73)	<.05
Model 3	1.0 (ref)	1.10 (.78-1.55)	1.24 (.89-1.72)	<.05
CVD mortality **346 (5.3)**				
Model 1	1.0 (ref)	1.06 (.58-1.93)	1.37 (.77-2.45)	<.05
Model 2	1.0 (ref)	1.06 (.58-1.93)	1.35 (.76-2.42)	<.05
Model 3	1.0 (ref)	1.06 (.58-1.94)	1.34 (.75-2.40)	<.05
Cancer mortality **492 (7.6)**				
Model 1	1.0 (ref)	1.79 (.97-3.30)	1.88 (1.03-3.25)*	.11
Model 2	1.0 (ref)	1.79 (.97-3.30)	1.89 (1.04-3.44)*	.10
Model 3	1.0 (ref)	1.82 (.98-3.35)	1.87 (1.03-3.42)*	.15
Incident CVD **802 (12.7)**				
Model 1	1.0 (ref)	.94 (.65-1.37)	1.06 (.74-1.53)	.17
Model 2	1.0 (ref)	.94 (.64-1.36)	1.05 (.73-1.50)	.23
Model 3	1.0 (ref)	.94 (.64-1.36)	1.03 (.72-1.49)	.29

*p<0.05 Adjustments: Model 1) age and sex. Model 2) +BMI. Model 3) +CPD

Table 5. Multivariable-adjusted HRs (95%CI) per BDNF genotype in non-smokers

Genotype	GG	AG	AA	
End point **Events (%)**		HRs 95% CI		*P*trend
Total mortality **1871 (10.1)**				
Model 1	1.0 (ref)	1.09 (.85-1.39)	1.02 (.81-1.30)	.41
Model 2	1.0 (ref)	1.08 (.85-1.38)	1.02 (.80-1.29)	.39
CVD mortality **603 (3.2)**				
Model 1	1.0 (ref)	.86 (.57-1.30)	.93 (.63-1.39)	.65
Model 2	1.0 (ref)	.85 (.56-1.28)	0.92 (0.62-1.37)	.66
Cancer Mortality **877 (4.7)**				
Model 1	1.0 (ref)	1.24 (.86-1.78)	1.03 (0.72-1.47)	.06
Model 2	1.0 (ref)	1.23 (.85-1.77)	1.02 (0.71-1.46)	.06
Incident CVD **1606 (8.9)**				
Model 1	1.0 (ref)	.88 (.69-1.13)	.90 (.71-1.15)	.85
Model 2	1.0 (ref)	.87 (.68-1.12)	.89 (.70-1.13)	.78

Adjustments: Model 1) age and sex. Model 2) +BMI.

Study II.

Smokers carrying a CHRNA risk allele (T) smoke more cigarettes and have a higher risk of death, incident COPD and tobacco-related cancer, compared to other smokers

Follow-up differed between the end points:

- Total, CVD, respiratory disease and cancer mortality, incident COPD and CVD, extended until 31 December 2009.
- Incident TRC and OC extended until 31 December 2010.

Mean follow-up was calculated stratified by smoking status, see Supplementary Table 1 in Paper II. Baseline data for study participants are presented in Table 6 and the genotype distribution in Table 7. In the tables 8-10, multivariable adjusted HRs *per genotype* is presented using the GG genotype as a reference (HR: 1.0) and the p value is for trend (from the additive models).

There was no significant deviation from Hardy-Weinberg equilibrium in any of the groups studied, P>0.05.

Table 6. Baseline data for study participants

	Female	Male	All
Study participants (%)	15 094(61)	9700(39)	24 794(100)
Age (mean±SE)	57(±8)	59(±7)	58(±8)
BMI (mean±SE)	25(±4)	26(±4)	26(±4)
Hypertension(%)	8569(57)	6655(69)	15 224(61)
Lipid Lowering Therapy (%)	320(2)	475(5)	795(3)
Previous diabetes (%)	480(3)	558(6)	1038(4)
CPD (mean±SE)	13±7	16(±9)	14±8

Table 7. Distribution of rs1051730 in the study population and stratified by smoking status

Genotype	CC	CT	TT
All genotyped subjects (n=26 471)	12 092 (45.7%)	11 442 (43.2%)	2937 (11.1%)
Analyzed subjects[1](n=24 794)	11 311 (45.6%)	10 732 (43.3%)	2751 (11.1%)
Current smokers (n=6951)	3109 (44.7%)	3041 (43.8%)	801 (11.5%)
Previous smokers (n=8426)	3916 (46.5%)	3611 (42.8%)	899 (10.7%)
Never smokers (n=9417)	4286 (45.5%)	4080 (43.3%)	1051 (11.2%)

1)With complete data for age, sex, BMI, smoking status, hypertension, LLT and previous diabetes

rs1051730 and smoking behaviour

The risk allele (T) showed an association with being a current smoker (compared to being a non-smoker) in an additive model adjusted for age and sex (OR: 1.04, 95% CI 1.00-1.09; p=0.05). Conversely, the risk allele was associated with a lower probability of being a former smoker (0.96,0.92-1.00; p=0.05). No association was seen with never smoker status. As seen before, the risk allele showed a strong linear association with smoking quantity (cigarettes smoked per day), β=1.14 CPD per allele; $p=9x10^{-5}$.

Association with smoking-related diseases in current smokers

Mortality end points

A total of 1508 (22%) events occurred of **all-cause mortality** during follow-up, and in additive models with the risk-allele (T) coded, there was a significant association with the risk allele in model 1 and 2 (HR 1.10, 95% CI, 1.02-1.10 and 1.10, 1.03-1.19, respectively). In model 3 the number of study participants was lower due to lacking data on CPD, and the association was nearly significant: 1.08, 1.00-1.17).

No significant association was seen with **CVD mortality** (model 1: 1.03, 0.90-1.17) where 500 (7.2%) events were reported during follow-up. Model 1 and 2 showed significant associations in **mortality from respiratory diseases** (1.38,1.05-1.83; and 1.38, 1.04-1.82) where 102 (1.5%) events occurred during follow-up. No association was seen with **cancer mortality** model 1: 1.08, 0.97-1.21.

Incident disease end points

All three models were significant for **incident COPD** during follow-up, with a total of 852 (12.3%) events (model 1-3: 1.29, 1.13-1.46; 1.29, 1.13-1.46; and 1.25, 1.09-1.43) as well as for **incident TRC** where 852 (13.5%), events occurred (model 1-3:

1.17, 1.06-1.29; 1.12, 1.06-1.29; and 1.11, 1.00-1.24). For **incident OC**, 810 (12.2%) events occurred but no association was seen, model 1: 1.05, 0.95-1.16 nor in **incident CVD** 500 (15.1%), model 1: 0.97, 0.89-1.07).

Association with smoking-related diseases in previous smokers

Mortality end points

In line with results from current smokers, significant results were seen in both models for **total mortality** where 1425(16.9%) events were reported, model 1+2: 1.12, 1.04-1.21 and 1.12, 1.04-1.21. For **CVD mortality**, 488 (5.8%) events, there was a significant association in both models: 1.20, 1.06-1.37 and 1.21, 1.06-1.37). **Respiratory disease mortality,** 70 (0.8%) events, was not significantly associated with the polymorphism, model 1: 1.28, 0.91-1.80, nor **cancer mortality** with 614 (7.3%) events reported, model 1: 1.08, 0.96-1.21).

Incident disease end points

Interestingly, **incident COPD** was no longer significantly associated with the risk allele, where only 211 (2.5%) events occurred during follow-up (model 1: 1.11-0.91-1.36). The associations with **incident TRC** were nearly significant (model 1: 1.11-0.99-1.24), a total of 693 (8.8%) events were reported. No significant associations were seen with **incident OC**, 1218 (15%) events, model 1: 0.95, 0.87-1.03, nor with **incident CVD**, 1047 (13%ents) ev, model 1: 1.00, 0.91-1.10.

Association with smoking-related diseases in never smokers

Mortality end points

No significant association was seen in any endpoint. In **all-cause mortality**, 1143 (12.1%) events were reported, model 1: 0.96, 0.88-1.04. In **CVD mortality**, 350 (3.7%) events occurred, model 1: 0.96, 0.88-1.04. Only 33 (0.4%) events occurred in **respiratory disease mortality**, model 1: 0.82, 0.48-1.39 and in **cancer mortality** 520 (5.5%) events were reported, model 1: 0.95, 0.83-1.08.

Incident disease end points

Equally, no significant association was seen in any endpoint. In **incident COPD**, 79 (0.8%) events occurred, model 1: 1.0, 0.72-1.39. A total of 559 (6.2%) events occurred in **incident TRC**, model 1:1.02, 0.90-1.15 and 1233 (13.8%) events of **incident OC** were reported, model 1: 0.95, 0.87-1.03. Model 1 for **incident CVD** was: 0.94, 0.85-1.03, and 955 (10.3%) events occurred during follow-up.

Table 8. Multivariable-adjusted HRs (95%CI) per CHRNA genotype in current smokers

Genotype		CC	CT	TT			
End point	Events	HR (95%CI)[1]			p_{trend1}	p_{trend2}	p_{trend3}
Inc COPD	480	1.0(ref)	1.28 1.05-1.56*)	1.66 (1.27-2.17*)	<.05	<.05	<.05
Inc TRC	852	1.0(ref)	1.14 (.99-1.33)	1.39 (1.13-1.71*)	<.05	<.05	<.05
Inc OC	810	1.0(ref)	1.10 (0.98-1.27)	1.06 (.84-1.33)	.35	.36	.77
Inc CVD	1022	1.0(ref)	.96 (.84-1.09)	.96 (.78-1.18)	.53	.68	.62
Total mortality	1508	1.0(ref)	1.04 (.93-1.16)	1.26 (1.08- 1.47)*	<.05	<.05	.07
CVD mortality	500	1.0(ref)	.89 (.74-1.08)	1.18 (.90-1.54)	.67	.57	.67
Resp. disease mortality	102	1.0(ref)	1.53 (.10-2.36)	1.83 (1.01-3.31)*	<.05	<.05	.06
Cancer mortality	677	1.0(ref)	1.11 (.95-1.31)	1.15 (.90-1.46)	.16	.16	.50

Adjustments: 1) Age and sex. 2) +BMI, hypertension, previous DM and LLT. 3) +CPD

Table 9. Multivariable-adjusted HRs (95%CI) per CHRNA genotype in previous smokers

		CC	CT	TT		
End point	Events		HRs (95%CI)[1]		*P*trend[1]	*P*trend[2]
Inc COPD	211	1.0(ref)	1.09 (.81-1.45)	1.27 (.82-1.96)	.29	.29
Inc TRC	693	1.0(ref)	1.10 (.94-1.29)	1.24 (0.98-1.58)	.06	.06
Inc OC	1218	1.0(ref)	.92 (.81-1.03)	.93 (.77-1.14)	.21	.21
Inc CVD	1047	1.0(ref)	1.01 (.89-1.14)	.99 (.81-1.22)	.99	.87
Total mortality	1425	1.0(ref)	1.17 (1.04-1.30)*	1.21 (1.02-1.44)*	<.05	<.05
CVD mortality	488	1.0(ref)	1.37 (1.14-1.66)*	1.28 (0.95-1.74)	<.05	<.05
Resp. disease mortality	70	1.0(ref)	1.38 (.83-2.29)	1.543 (.72-3.28)	.16	.17
Cancer mortality	614	1.0(ref)	1.04 (.88-1.24)	1.20 (.92-1.55)	.22	.22

Adjustments: 1) Age and sex. 2) +BMI, hypertension, previous DM and LLT.

Table 10. Multivariable-adjusted HRs (95%CI) per CHRNA genotype in never smokers

Genotype		**CC**	**CT**	**TT**		
End point	Events		HRs (95%CI)[1]		*P*trend[1]	*P*trend[2]
Inc COPD	79	1.0(ref)	1.27 (.80-2.01)	.73 (.31-1.74)	.99	.99
Inc TRC	559	1.0(ref)	1.08(.91-1.29)	.97(.73-1.29)	.79	.79
Inc OC	1175	1.0(ref)	.98 (.87-1.11)	.86 (0.71-1.04)	.20	.21
Inc CVD	955	1.0(ref)	.94 (.82-1.07)	.88 (0.71-1.09)	.18	.18
Total mortality	1143	1.0(ref)	.97 (.85-1.09)	.91 (0.75-1.11)	.33	.31
CVD mortality	350	1.0(ref)	.81 (.65-1.02)	.89 (0.63-1.26)	.18	.15
Resp. disease mortality	33	1.0(ref)	.79 (.39-1.63)	.57 (0.21-2.39)	.46	.45
Cancer mortality	520	1.0(ref)	1.03 (.86-1.24)	.81 (.60-1.11)	.42	.41

Adjustments: 1) Age and sex. 2) +BMI, hypertension, previous DM and LLT.

Additional analyses

Bladder and lung cancer

Since the rs1051730 had been associated with incident bladder and lung cancer[27] we analysed these end points separately. The association with bladder cancer was not confirmed in our cohort but a significant risk increase was seen for lung cancer in current smokers with 277 events (4.4%). Model 1 and 2 were significant (1.29, 1.09-1.53 and 1.28, 1.08-1.52) but not model 3 where we adjusted for CPD, (1.18, 0.98-1.41). As we excluded lung cancer from the end point incident TRC, the associations with current smoking were no longer significant.

Genotype and median age of death

The median age of death stratified by genotype was also analysed, with focus on potential differences between subjects with a low- (CC) and high-risk (TT) genotypes. In current smokers, the median age of death was 1.4 years lower in the TT carriers compared to CC carriers and in previous smokers 0.2 years lower amongst TT carriers compared to CC carriers. In the never smoking group, the median age of death was 2.1 years higher in TT carriers compared to those with the CC genotype.

Study III.

HE4 is higher in current smokers but predicts mortality and morbidity irrespective of smoking status

Characteristics from the baseline examinations of the participants in the two ”analysis arms” are shown in Table 11.

Follow-up for end points:

- total, CVD and cancer mortality extended until 31st December 2013.
- LUB until 31st December 2011.
- incident TRC and OC until 31st December 2010.

Table 11. Baseline characteristics of the two analysis arms in study MDC-CC

	All	Women	Men
Cardiovascular arm[1]			
N (%)	4614 (100)	2776 (60)	1838 (40)
Age mean (SD)	58 (6)	57(6)	56(6)
BMI mean (SD)	26(4)	25(4)	26(3)
Hypertension (%)	2909(63)	1654(60)	1255(68)
DM (%)	351(8)	152(6)	199(11)
HDL mean (SD)	1(0.4)	1.5(0.4)	1.2(0.3)
LDL mean (SD)	4(1)	4(1)	4(0.9)
eGFR mean (SD)	75(15)	71(13)	80(15)
Current smokers (%)	1214(26)	711(26)	503(27)
Former smokers (%)	1536(33)	755(27)	781(43)
Never smokers (%)	1864(41)	1310(47)	554(30)
Cancer arm[2]			
N (%)	4671(100)	2804(60)	1867(40)
Age mean (± SD)	57(6)	57(6)	58(6)
BMI (mean ±)	26(4)	25(4)	26(3)
Current smokers (%)	1227(26)	717(26)	510(27)
Former smokers (%)	1558(34)	764(27)	794(43)
Never smokers (%)	1886(40)	1323(47)	563(30)

End points analysed: 1) Total and CVD mortality and incident CAD. 2) Cancer mortality, incident LUB/TRC/OC

HE4 in relation to smoking status

Earlier studies had observed higher HE4 in smokers and we investigated the relationship between plasma levels of HE4 and smoking with a linear regression with age, sex and current smoking as independent variables. The dependent was levels of HE4 expressed as per SD increment of log-transformed HE4 and the analysis was made in the ”cancer-arm” with 4671 participants. The levels of HE4 was distinctly higher in current smokers (β=0.89, p=2.8x10^{-178}), also demonstrated in Figure 5. Levels of HE4 were consistently lower in never smokers (β=-0.46, p=6.0x10$^{-55)}$ compared to ever smokers (former+current smokers).

Furthermore, the OR for being a current smoker versus being a non-smoker in the top versus bottom quartile of HE4 was (14.7, 95% CI: 11.6-18.6, p=3.9x10^{-111}).

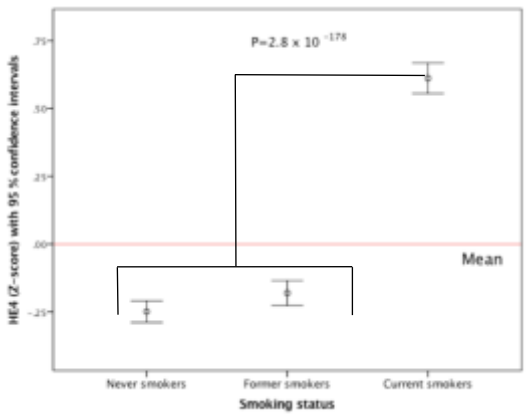


Figure 5: Levels of HE4 are significantly higher in smokers compared to former and never smokers

HE4 in relation to covariates

Each covariate was tested in relation to HE4. In the present study, each SD increment of log transformed HE4 was positively related to: age (β=0.38, 4.2x10^{-55}), male sex (β=0.11, 2.8x10^{-4}), hypertension, (β=0.10, p=0.01), and per mmol increment of LDL (β=0.07, p=3.6x10^{-5}). Negative linear associations were seen between HE4 and: HDL (β= -0.29, p=6.9x10^{-14}), per ml/min increment of eGFR calculated with the MDRD formula (β=-0.01, p=3.7x10^{-8}), and with per kg/m^2 increment of BMI (β=-0.02, p=3.6x10^{-5}). No significant association was seen with diabetes status (β=0.08, p=0.15).

Association with smoking-related diseases in the "CVD-arm"

Mortality end points

The mean follow-up time for **total mortality** was 19.2 (4) years and HE4 strongly predicted total mortality in the entire cohort. The effect estimate (HR) was only slightly attenuated when stratified by smoking status, from current to previous to never smoker. Adjusting for pack-years (model 3 in current smokers) modestly attenuated the effect estimate, see Table 12. The mean follow-up time for **CVD mortality** was 19.2 (4) years and HE4 predicted the outcome in the whole cohort. However, the association was no longer significant in former smokers after model 2 adjustment, see Table 12.

Incident CAD

For **incident coronary artery disease**, mean follow-up was 18.5 (5) years. HE4 predicted CAD in the entire cohort as well as in current and never smokers, but not in former smokers.

Association with smoking-related diseases in the "Cancer-arm"

Mortality end point

For **cancer mortality**, the mean follow-up time was 19.2(4) years and HE4 predicted the outcome in the cohort as a whole. When stratified to smoking status the significance remained among current and former smokers. (Table 13)

Incident disease end points

The mean follow-up time for **incident LUB** was 17.6 (4) years and HE4 predicted inc LUB in the entire cohort, remaining significant in current and former smokers when stratified by smoking status. **Incident TRC** had a mean follow-up of 14.5 (9) years and HE4 predicted TRC among current smokers. For **incident OC**, with a mean follow-up of 14.8(7) years, HE4 predicted OC among never smokers(Table 13).

Table 12. HRs from multivariable adjusted models per 1 SD HE4 in the "Cardiovascular arm"

End point	Cases model 1/3[a]	Events model 1/3	Model 1 HR (95%CI)[1]	p_{value1}	Model 2 HR (95%CI)[2]	p_{value2}	Model 3 HR (95%CI)[3]	p_{value3}
Total mortality								
All	4613	1123	1.41(1.32-1.50)	<.05	1.41(1.33-1.51)	<.05	1.25(1.17-1.34)	<.05
Current smokers	1214/1057	410/354	1.27(1.15-1.41)	<.05	1.30(1.16-1.44)	<.05	1.25(1.11-1.41)	<.05
Former smokers	1536	372	1.21(1.08-1.35)	<.05	1.18(1.05-1.32)	<.05	NA	
Never smokers	1863	341	1.17(1.05-1.32)	<.05	1.18(1.05-1.32)	<.05	NA	
CVD mortality								
All	4611	327	1.44(1.28-1.61)	<.05	1.45(1.29-1.63)	<.05	1.27(1.12-1.45)	<.05
Current smokers	1214/1057	117/97	1.24(1.02-1.51)	<.05	1.28(1.04-1.57)	<.05	1.27(1.01-1.60)	<.05
Former smokers	1532	110	1.23(1.00-1.50)	.05	1.18(0.96-1.45)	.11	NA	
Never smokers	1863	100	1.28(1.04-1.58)	<.05	1.25(1.02-1.54)	<.05	NA	
Inc coronary artery disease								
All	4516	515	1.24(1.13-1.36)	<.05	1.24(1.13-1.36)	<.05	1.16(1.05-1.28)	<.05
Current smokers	1193/1040	164/144	1.25(1.06-1.47)	<.05	1.23(1.04-1.46)	<.05	1.23(1.02-1.49)	<.05
Former smokers	1478	186	1.02(0.87-1.19)	.85	1.01(0.87-1.18)	.89	NA	
Never smokers	1845	165	1.19(1.01-1.40)	<.05	1.18(1.00-1.40)	<.05	NA	

a: The variations in the number of cases are due to censored cases and excluded prevalent events. In model 3 of current smokers the lower number of cases are due to lacking data on smoking quantity.

1) Age and sex 2) 1+ BMI, hypertension, diabetes, HDL, LDL, eGFR, 3) 1+2+ in All: + current smoking. Model 3 in current smokers: +pack years. NA: Not analyzed

Table 13: HRs from multivariable adjusted models per 1 SD HE4 in the "Cancer arm"

End point	Cases model 1/3[a]	Events model 1/3	Model 1 HR (95%CI)[1]	p_{value1}	Model 2 HR (95%CI)[2]	p_{value2}	Model 3 HR (95%CI)[3]	p_{value3}
Cancer mortality								
All	4666	502	1.38(1.26-1.51)	<.05	1.39(1.26-1.52)	<.05	1.22(1.10-1.35)	<.05
Current smokers	1226/1067	192/166	1.26(1.08-1.47)	<.05	1.26(1.08-1.47)	<.05	1.19(1.01-1.41)	<.05
Former smokers	1556	166	1.19(1.00-1.40)	<.05	1.19(1.00-1.40)	<.05	NA	
Never smokers	1884	144	1.10(0.92-1.31)	.29	1.11(0.92-1.31)	.29	NA	
Incident LUB (trachea, bronchus, lung or pleural cancer)								
All	4665	104	2.06(1.71-2.48)	<.05	1.99(1.65-2.40)	<.05	1.52(1.22-1.88)	<.05
Current smokers	1211/1055	65/59	1.59(1.22-2.08)	<.05	1.50(1.15-1.97)	<.05	1.33(0.99-1.78)	.06
Former smokers	1554	28	1.69(1.14-2.51)	<.05	1.69(1.14-2.51)	<.05	NA	
Never smokers	1881	11	.89(0.49-1.62)	.71	.89(0.49-1.62)	.71	NA	
Incident tobacco-related cancer								
All	4393	434	1.27(1.15-1.40)	<.05	1.26(1.14-1.39)	<.05	1.12(1.01-1.25)	<.05
Current smokers	1111/960	161/140	1.27(1.07-1.50)	<.05	1.27(1.07-1.50)	<.05	1.19(.99-1.44)	0.07
Former smokers	1479	157	1.08(0.91-1.28)	.38	1.08(.91-1.28)	.38	NA	
Never smokers	1803	116	.94(0.78-1.14)	.54	.94(.78-1.14)	.54	NA	
Inc other cancer								
All	4458	640	1.10(1.02-1.19)	<.05	1.11(1.03-1.21)	<.05	1.10(1.00-1.20)	<.05
Current smokers	1173/1020	171/146	.97(.83-1.12)	.64	.97(.83-1.13)	.68	.97(.82-1.15)	.69
Former smokers	1492	220	1.07(.93-1.24)	.35	1.07(.93-1.24)	.34	NA	
Never smokers	1793	249	1.19(1.04-1.35)	<.05	1.19(1.05-1.36)	<.05	NA	

a)The variations in the number of cases are due to censored cases and excluded prevalent events. In model 3 of current smokers the lower number of cases are due to lacking data on smoking quantity.

Adjustments: 1) Age and sex. 2) Model 1 + BMI. 3) Model 3 in *All* + current smoking/Model 3 in current smokers: +pack years

Additional analyses

HE4 and ovarian cancer

As mentioned, HE4 is used in the clinic to support the diagnosis of ovarian cancer (OVC). Among the 2804 women in the ”Cancer arm”, 14 had a prevalent diagnosis of ovarian cancer at baseline, and we tested whether there were any associations between HE4 and the women with a prevalent OVC diagnosis at baseline as well as between the women who would be affected by OVC during follow-up (n=29).

In a linear regression with previous OVC and age as independent variables and HE4 as the dependent variable, a significant association was seen, (β=0.57 SDs higher HE4, p=0.03). A logistic regression model with prevalent OVC as dependent and HE4, age and BMI as independent variables also showed a significant association between HE4 and prevalent OVC (OR per HE4 SD: 1.79, 95%CI 1.06-3.02, p=0.029)

No association was seen in a linear analysis with incident OVC and age as independent, and HE4 as dependent variable, (β=-0.02, p=0.93). In a logistic regression model including HE4, age and BMI as independent variables, HE4 was not associated with incident OVC (OR: 0.97, 95%CI: 0.67-1.43, p=0.89).

In conclusion, the 14 women with a previous OVC had elevated HE4 at baseline, and the OR for an elevated HE4 and a diagnosed OVC was significant. The 29 women with a diagnosis of OVC during follow-up were not shown to have higher HE4 at baseline. Neither was there a significant association between baseline HE4 and a future diagnose of OVC in a logistic regression model.

The results from the Cox regressions presented in Tables 12–13 (Results p. 53 and 54) were not substantially affected when excluding the 14 women with a prevalent OVC.

Study IV.

HE4 predicts 90-day mortality in patients presenting with acute dyspnea in the Malmö Emergency Department

In Table 14 baseline characteristics and comorbidities of the study population of ADYS are presented.

Table 14.

	ADYS-Discovery (N=379)	ADYS-Replication (N=420)
Age mean ±SD	70 ±18	70 ±18
Male sex (%)	182 (48)	187(45)
Respiratory rate mean ±SD	24±7	25±7
Saturation mean ±SD	94±6	93±7
CRP (mg/L) median (IQR) [a]	8(31)	10(32)
METTS-A (%) (green, yellow, orange, red)	33(9)/149(47)/124(33)/43(11)	19(4)/207(49)/137(33)/57(14)
Smoking status CS/FS/NS [c] **(%)**	77(20)/198(52)/104(28)	77(18)/207(49)/136(33)
Diseases[d] **n (%)**		
Coronary Artery Disease	120 (32)	135(32)
Chronic Heart Failure	134 (35)	136(32)
COPD	105(28)	139(33)
Asthma	48(13)	46(11)
Restrictive lung disease	20(5)	24(6)
Other lung disease	4(1)	9(2)
Pulmonary thromboembolism	33(9)	59(14)
Cancer	62(16)	76(18)
Diabetes mellitus	73(19)	71(17)
Hypertension	177(47)	171(41)
Renal disease[e]	43(11)	30(7)

a: Interquartile range b: Categories of METTS-A (least to most critical) c: Current smoker/former smoker/never smoker
d: Self-reported at baseline examination e: Self-reported or eGFR<30 at baseline

HE4 in relation to smoking

The association between HE4 and smoking status were examined with a linear regression with age, sex and current smoking as independents and log-transformed HE4, expressed as z-scores within the 379/420/799 study participants in the analysis of ADYS-D/ADYS-R and ADYS-P respectively. HE4 was not associated with current smoking in any of the studies, ADYS-D (β per 1-SD: 0.89; p=0.37), ADYS-R (0.12; p= 0.21) nor in ADYS-P (0.12; p = 0.11).

HE4 in relation to covariates

The analysis of associations between log-transformed HE4 and covariates presented below are from ADYS-P. There were no major differences in the associations between the variables and HE4 in ADYS-D or ADYS-R, why covariate data is not presented.

Each SD-increment of log transformed HE4 was positively and significantly associated with age (ß=0.03, p-=<0.00), log-transformed CRP (mg/L): [0.01; <0.001], METTS (low priority to high priority) [0.31; <0.001], and for respiratory rate (per breath/min) [0.04; <0.001]. Males had higher HE4 (0.30 SD higher, p=<0.001).

A negative association was seen with saturation (%) [-0.04, p=<0.001].

HE4 and 90-day mortality

See Table 15 for HR: s and 95%CI from multi adjusted Cox regressions stratified by smoking status. In Table 16, HR: s between quartiles of log-transformed HE4 and p-values for trend are presented. Quartile 1 (Q1) represents the lowest values of HE4 and Q4 the highest. Figures 6-8 display crude survival rates in Kaplan-Meier curves in ADYS-D, ADYS-R and ADYS-P.

ADYS-D

A total of 46 events occurred during follow-up in the 379 participants. HE4 was significantly associated with the outcome in all (n=379) respectively ever smokers (n=275) but only borderline significant in never smokers (n=104). The effect estimates were similar but somewhat higher in never smokers.

ADYS-R

In the 420 participants, 53 events were reported during follow-up. HE4 was significantly associated with 90-day mortality in all (n=420), ever smokers (n=284)

respectively never smokers (n=136). The effect estimate was highest in never smokers.

ADYS-P

In the pooled analysis, HE4 was significantly associated with 90-day mortality in all smoking categories, with the highest effect estimate in never smokers.

HE4 and risk improvement

In ADYS-P, we calculated C-statistic index for 90-day mortality for a model with traditional risk factors (age, sex, respiratory rate, saturation, CRP and METTS) to 0.77 (95 % CI 0.73-0.81). After addition of HE4 the index increased to 0.79 (0.76-0.83) compared with the model with only traditional risk factors.

When analysing the net reclassification index (NRI), adding HE4 to traditional risk factors resulted in a highly significant *continuous* (which does not consider crossing any particular clinical threshold of risk) NRI of 62% (95% CI=26-83%) (p=<0.001).

For analysis of *categorical* NRI, we tested a cut-off of predicted risk of 90-day mortality <1%, (which we considered a very low-risk category) or >1% respectively. When HE4 was added to traditional risk factors, it resulted in a significant improvement of risk reclassification across this border, with an NRI of 5,7% ($p < 0.001$).

When HE4 was added to traditional risk factors it conveyed a correct down-classification of 6,6% of the patients that survived during the 90-days period into the very low-risk category (< 1% predicted risk of 90-day mortality). Only 0,9% of the patients that survived were incorrectly reclassified from the very low risk category into the higher risk category when adding HE4 to the traditional risk factors.

No patients that died during the 90-days follow-up period were reclassified to the very low risk category when HE4 was added to traditional risk factors.

Table 15: Multi adjusted Cox regression models of HE4, end point 90-day mortality

	All	P_{value}	Ever smokers	P_{value}	Never smokers	P_{value}
ADYS-Discovery						
N [a]/n [b]	379/46	<.05	275/32	<.05	104/14	.07
HR (95%CI) [c]	2.48(1.46 -4.21)		2.50(1.36-4.59)		2.90(.92-9.13-)	
ADYS-Replication						
N/n	420/53	<.05	284/37	<.05	136/16	<.05
HR (95%CI)	3.36(2.07 -5.45)		2.70(1.61-4.51)		6.70(1.94-23.15)	
ADYS-Pooled						
N/n	799/99	<.05	559/69	<.05	240/30	<.05
HR (95%CI)	1.86(1.38 -2.52)		1.77(1.27-2.52)		2.24(1.15-4.37)	

a=Participants b=Events c=Adjusted for age, sex, respiratory rate, oxygen saturation, METTS-A and CRP.

Table 16. HR:s between quartiles of HE4 and p-value for trend

ADYS-P	All participants	P_{value}	Quartile 1	Quartile 2	Quartile 3	Quartile 4	P_{trend}
N [a]/n [b] events	799/99	$5x10^{-5}$	199/2	200/9	200/36	200/52	$1.1x10^{-6}$
HR (95%CI) [c]	1.82 (1.40-2.38)		ref (1.0)	2.19 (0.47-10.3)	6.43 (1.48-27.9)	8.66 (1.99-37.7)	

a=Participants b=Events c=Adjusted for age, sex, respiratory rate, oxygen saturation, METTS-A and CRP.

Figure 6-8. Kaplan-Meier curves of cumulative mortality during 90-day follow-up in ADYS-D, ADYS-R and ADYS-P. Quartile 1 represents the lowest values of HE4, quartile 4 the highest.

Figure 6. ADYS-Discovery

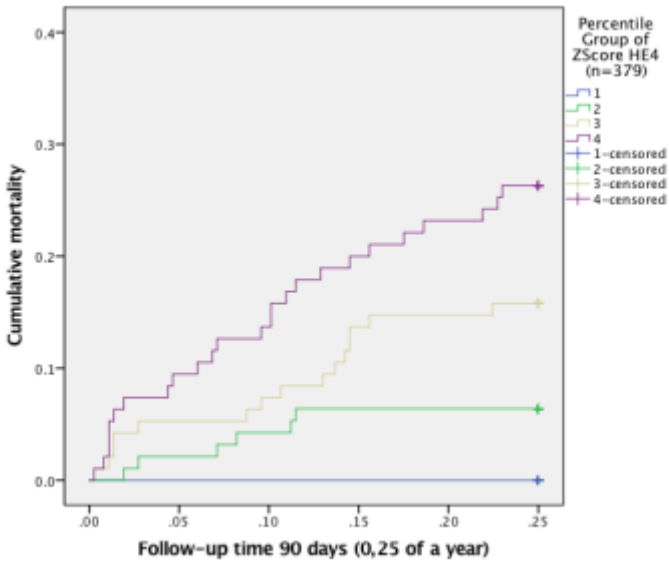


Figure 7. ADYS-Replication

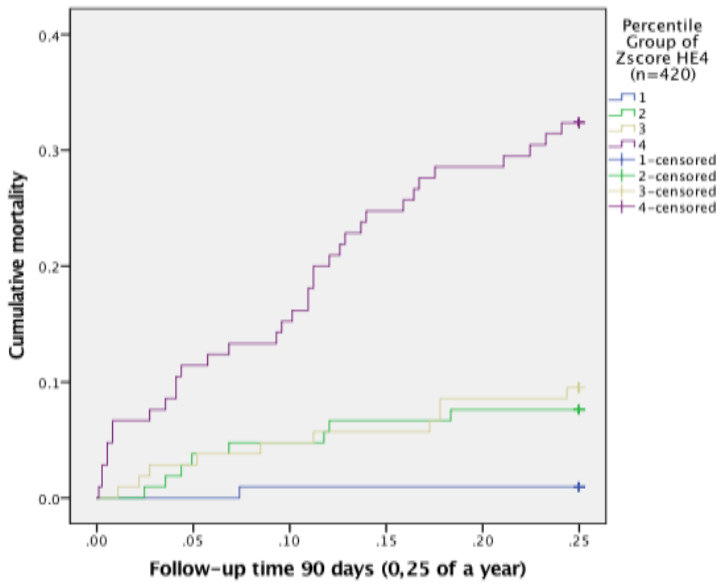
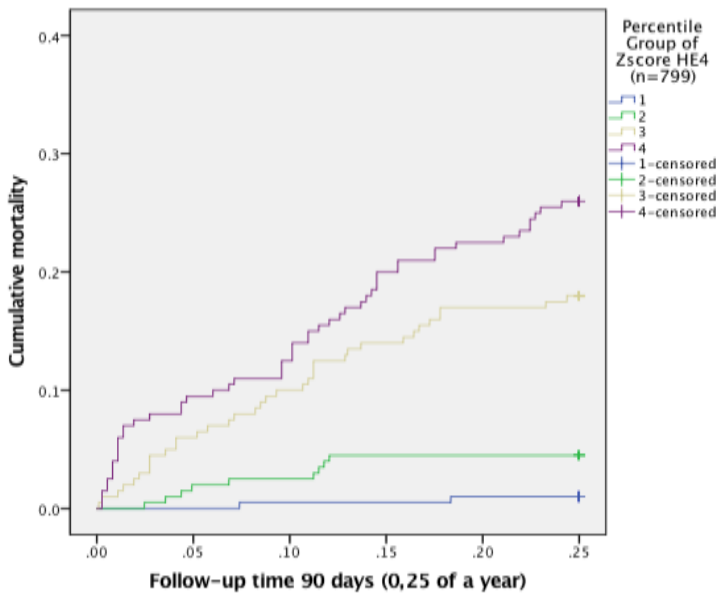


Figure 8. ADYS-Pooled



Discussion

Studies I and II

BDNF polymorphism, smoking initiation and mortality

In Paper I, we concluded that a genetic variation of BDNF, known to be associated with both smoking and high BMI[132], predicts risk of death in smokers but not in non-smokers. The risk of death is especially from CVDs and the associations with the polymorphism are independent of traditional risk factors, such as age, sex, BMI and smoking quantity. These findings indicate that smokers with the polymorphism are facing a higher mortality risk that is not mediated by obesity and its complications, but to the prolonged exposure to smoking, suggesting that the polymorphism affects their ability to quit smoking during follow-up. Supporting this interpretation is also the lack of associations among previous and never smokers carrying the risk allele, even though the number of deaths in non-smokers is somewhat lower than in current smokers.

However, there are links between the risk allele and abuse of other substances as well as with psychiatric diseases[89,90], which can affect the lower life expectancies. Unfortunately, we cannot exclude these patients since there are no data reported on these circumstances in the MDCS.

The SNP did not predict mortality from cancer or incident CVD events. An explanation could be that smokers already have had their first CVD event (before baseline), and therefore were excluded from analysis, or that the first event had a more severe outcome in smokers with the risk allele.

The CI: s were wide, especially for CVD mortality and demands a need for caution when drawing conclusions. The effect estimates (HRs) for total mortality were additive, suggesting an estimated risk increase of around 10% per risk allele of having the event during follow-up, the trend was significant but in genotype analysis (GG as reference) the associations were not significant, with the exception of CC vs GG in cancer mortality model 1 and 2, a finding of unknown importance since the trend was not significant.

CHRNA polymorphism, smoking quantity and tobacco-related diseases

The novel finding in Paper II was that the CHRNA polymorphism rs1051730 predicted an increased risk of death among current and previous smokers in the MDCS. We were also able to confirm the associations between the SNP and incident lung cancer[11,14,16,24,25,27,28,44,135], TRC[14,27,136] and COPD[11,26,43,135] as well as smoking quantity[11,20,22,27,98]. Furthermore, despite a similar number of events, the gene variance could not predict any incident diseases or deaths in never smokers.

In addition to all-cause mortality, there was also a significant association with death from respiratory diseases in current smokers, which was not seen in former smokers. Even though it could be interpreted as a lower risk due to smoking cessation, this lack of association may also be explained by low power since the number of events was more than 30% lower in the former smoking group.

Furthermore, one might consider the non-significant association reported in former smokers for incident COPD compared to the quite strong association for current smokers, as a highlight for the benefits of smoking cessation. But again, the incidence of COPD is lower in previous smokers.

Unlike the current smokers, there was an association between the SNP and CVD mortality in former smokers which might be a bit surprising. One explanation could be that smoking in itself is such a strong risk factor for CVD mortality, that genetic influences may affect the risk of CVD to a lesser extent and therefore not detectable in this type of assays.

Reflections on overlapping genetic influences on smoking behaviour and smoking-related diseases

In analysis of both SNPs, the associations in current smokers were significant even after adjustment for smoking quantity (borderline significance in total mortality for CHRNA) suggesting another mechanism than tobacco burden.

As mentioned before, these common, complex diseases are dependent of several mechanisms. Yet, there are no obvious pathways that link any of the studied polymorphisms to the increased risk of death and smoking-related diseases but the receptors are widely distributed in the body and could theoretically have impact on various, pathophysiological processes. Up to this date, it seems more plausible that the increased risk is caused by the altered smoking behaviours conferred by the SNPs since there are doubts on the measurements of self-reported smoking burden [23,27,137,138]. The fact that the SNPs in our studies associate with most end points for currents smokers, followed by fewer associations for former smokers but none at all for never smokers is striking and may be an argument for the fact that it really is the smoking burden that confers the higher risks.

There are concerns of the reliability of self-reported data of smoking quantity[139]. For example cotinine levels in saliva can differ hugely in smokers who report the same CPD[49] and self-reported smoking amount over time may contain many errors[140]. Bloom et al suggested that exhaled CO would be a better biomarker for cigarette exposure [141], but the method is not suitable for large cohort studies.

In MDC-CC we did analyse plasma cotinine levels in 517 participants (unpublished data). Current smokers (n=168) had significantly higher values compared to non-smokers (n=349), which could be regarded as a validation of the reported smoking status in this cohort. Still, cotinine is not a measure of smoking quantity and the number of cigarettes reported by the current smokers at baseline should be interpreted with caution. Another way of measuring smoking burden is *smoking topography*, which is the measuring of number of puffs and depth of inhalations which can differ and could have consequences since it may affect the amount of toxins inhaled[142,143]. Data on these parameters could have been of value to strengthen the genetic association demonstrated in Papers I and II.

One of the limitations when interpreting the results is that the genetic associations reported are found in a population consisting to a large extent of Nordic origin, and one must be aware that the associations may be different in other populations. The low autopsy rates must result in possible low power for interpretation of cause-specific mortality and incident diseases. Even though the burden of nicotine during follow-up is not known in the MDCS, the increased risk is based on the smoking status at baseline. Yet, a re-investigation between 2007-2012 of study participants of the MDC-CC, showed that more than half (54%) of the smokers who reported daily smoking at baseline had stopped smoking and nearly two thirds of the intermittent smokers. Former smokers stayed former to the most part (91%) and never smokers stayed never smokers at re-examination to 99%[144].

Strengths of the study are the large cohorts, duration of follow-up, and to large extents validated end points[130,145].

Study III

HE4 as a marker for smoking and prediction of disease

The novel approach in this paper was the investigation of the relationship between plasma levels of HE4 and smoking in a population of middle-aged subjects (MDC-CC). We detected an extremely strong association between HE4 and current smokers compared to former and never smokers (Figure 5, p. 51). Moreover, in the search for a biomarker predicting disease in smokers, we could show that HE4

predicted future events of death and smoking-related diseases in the MDC-CC, not only in smokers but also in non-smokers.

The effect estimates of HE4 on total mortality was highest among current smokers, but the attenuation was only slightly and gradually affected in former and never smokers. In current smokers, high HE4 identified smokers with a higher risk despite smoking quantity.

As mentioned, the mechanism behind elevated HE4 is not fully understood and one can only speculate that HE4 is a marker of an ongoing disease process, which may be accelerated by the organ damage of tobacco in smokers. In non-smokers, elevated HE4 may reflect a similar process, presumably caused by other agents. However, adjustments for traditional risk factors for CVD in our models did not weaken the risk increase by HE4 on CVD outcomes, with the exception of the subgroup of previous smokers.

Malign tissues have been found to express HE4[33,34,37,38,107,108] and HE4 is also suggested as a marker or mediator of fibrosis in the heart and kidney, regardless of smoking status[109-111] which also may correspond to the association with age. An approach and possible key finding in understanding the mechanism behind reduced kidney function and high HE4 was presented by Le Bleu et al in 2013[109]. In an attempt to rule out the functional contribution of myofibroblasts in fibrosis, mouse models were used to perform a gene expression profiling, aiming to identify candidate genes mediating fibrosis. Unexpectedly, HE4 was the most upregulated gene in fibrosis-related fibroblasts. The proposed mechanism is an inhibition of protease activity by HE4 and thereby reducing the degradation of collagen type 1, leading to fibrosis. To further confirm the hypothesis, mice were treated with HE4-specific neutralizing antibodies resulting in less progress of fibrosis. A similar upregulation of the HE4 gene was seen in human myofibroblasts from fibrotic kidneys and the fibroblasts secreted HE4. Consequently, patients with kidney disease had elevated HE4 compared to healthy patients, correlating with the grade of fibrosis.

A clinical, prognostic value of HE4 was also investigated by deBoer[110] and Piek[111] in relation to heart failure (HF), since elevated plasma levels of HE4 had been detected in patients with heart failure (unpublished data). In 567 patients hospitalized and treated for HF, HE4 was collected at time of discharge. After 18 months of follow-up, significant associations between high HE4 and 1) HF severity 2) New York Heart Association Classification (NYHA) 3) all cause-mortality 4) rehospitalisation were seen. These associations were later confirmed in a study of 101 patients with chronic HF[111]. Smoking was not considered in any of the two studies.

In summary, the role of HE4 as a mediator and/or marker of fibrosis and whether this is the causal relationship between HE4 and various diseases need to be further investigated.

The associations presented in the MDC-CC may be different in other populations and relations between HE4 and risk of future mortality and morbidity in populations with chronic diseases or acutely ill patients is not possible to assess in this setting. Apart from the general limitation on interpretation of associations in other populations, underlying, not diagnosed malignancies could potentially be responsible for the elevated levels of HE4 in smokers at baseline.

The fact that HE4 could be interpreted differently in a patient with acute illness was investigated by Nagy[37], who measured HE4 levels in men with and without lung cancer. HE4 was positively associated with age and smoking in healthy participants but no such association was observed in the cancer patients.

Study IV

Elevated HE4 predicts short-term mortality in an acute setting

When we investigated the properties of HE4 in an acute setting with patients presenting at an emergency ward (ADYS), we could not replicate the association between HE4 and smoking. This may be due to low power or to the assumption that smokers with acute dyspnea symptoms abstain from smoking and thereby blurring a potential relationship.

Nevertheless, HE4 predicted 90-day mortality in ADYS-D (n=379) in the entire study population as well in ever smokers irrespective of underlying disease. A borderline significant relation was seen for never smokers. When repeating the analysis in the ADYS-R (n=420), the findings were replicated and also seen in never smokers. When pooling the data in ADYS-P (n=799), the significant effect estimates seen in ADYS-R, were consistent and with less wide CI: s. Additionally, in quartile analysis of HE4 and 90-day mortality in ADYS-P, with the lowest levels (Q1) of HE4 as reference, the highest levels (Q4) exhibited an effect estimate of 8.7.

These results support the idea that HE4 not only reflects the amount of smoking-related substances in the body but rather a generic tissue damage not only induced by smoking.

HE4 and improvement of risk prediction

Dyspnea can be demanding since the symptom can be caused of various diseases and the presentation does not always correlate with the medical interventions or investigations needed. In the acute setting of the ED, the clinician commonly has to decide whether a patient is to be admitted or not without knowing the underlying

cause of dyspnea. In addition to patient history, cause-specific biomarkers and clinical parameters can be helpful but also challenging, for example in patients with multiple underlying diseases.

Given the high risk of 90-day mortality in patients with elevated HE4, we tested the categorical reclassification properties, to guide in decisions about safe discharging of the dyspnea patient. In this cohort, 6.6% of the patients that were admitted could potentially have been sent home, according to a risk of 90-day mortality of <1% if HE4 levels were taken into account. Very few patients were reclassified from the low risk category to a higher risk and importantly, no patients that died during follow-up were reclassified into the category of low risk. However, we cannot be sure that the surviving patients survived due to the fact that they were admitted and received treatment.

Based on these findings from the ED, we suggest that HE4 could be a valuable complement to clinical judgement, when deciding whether the patient with dyspnea should be admitted or sent home.

Conclusions

- Understanding the genetics of dependence could be a way to optimize targeted treatments and preventing adverse health effects of smoking. Further genetic variants need to be identified to improve risk stratification which possibly could help motivating high-risk individuals to smoking cessation, or perhaps, never initiate smoking. New molecular pathways related to genetic variance should be investigated to enable development of a more individualized intervention therapy.
- Carriers of a BDNF risk allele (A) are more likely to be smokers, and if they smoke they are at higher risk of dying than other smokers.
- Smokers carrying a CHRNA risk allele (T) smoke more cigarettes and have a higher risk of death, incident COPD and tobacco-related cancer, compared to other smokers.
- Plasma levels of HE4 are elevated in smokers in the general population but not in a population with acute dyspnea. HE4 is an accurate marker of smoking exposure but also reflects subclinical disease or pathology susceptibility.
- HE4 may be used as a mortality and disease risk marker in smokers and possibly also in non-smokers, both in the general population and in patients with acute dyspnea.
- Smokers with high values of HE4 could be regarded as high risk, which may motivate a more aggressive smoking cessation treatment even when accompanied with negative side effects.

Perspectives

It is well established that smoking causes serious health consequences but it is also apparent that there is a variation of susceptibility to the harmful effects of tobacco smoke. Is smoking equally dangerous to all and if not, how can we identify and motivate smokers with the higher risk?

The aims of this thesis are about finding smoking individuals with a higher risk where the benefits of smoking cessation are undisputable, for the individual as for society. Again, inhaling tobacco smoke is always an unhealthy behaviour and cessation should always be promoted, but if the treatment recommended by the clinician better meets the needs of the smoker, the chances of successful results would be more likely, at least in theory. Today the work against tobacco use is mostly about education and other actions such as media campaigns, increasing taxes, banning tobacco advertising and providing smoke-free indoor environments which are all expensive efforts. In the future, knowledge of unfortunate genetic variance instead of basing risk profiles on family history (which is not a valid information and sometimes induces a feeling of false security) could be used in the individual patient for motivating smoking cessation or even better, preventing smoking initiation.

One should also be aware that other mechanisms are involved in smoking behaviour as shown in Johnson et al[45] where gender, occupation, anxiety disorder and substance use disorders independently predicted daily onset of smoking in an American population. Thus, genetic information may be one, but not the only factor that should be taken into account when directing and planning interventions of smoking cessation.

As mentioned in the Introduction, the diseases investigated in this thesis are all common complex diseases and the underlying causes are mixtures of genetic predisposition and environmental factors. The interaction between environmental factors, such as tobacco use, is suggested to be important when determining susceptibility to disease. Consequently, the associations between the BDNF and CHRNA SNPs and smoking behaviour needs to be established in other populations and in other study designs.

Further information on HE4 such as a GWAS to explore further properties, associations with other end points in other populations and in short- and long-term

settings as well as assessing a cut-off value for plasma HE4 to be used in the clinic, are examples of potential, future projects.

Maybe, in the future, a polygenetic risk score and/or biomarkers such as HE4 could add sufficient individual risk information and benefit to our global health. Moreover, the development of e-cigarettes offers new challenges, and even though they are likely to be less lethal than the traditional cigarette[146], they seem to induce epigenetic changes in an antifibrotic mediator of cardiac and renal tissue, which may result in fibrosis formation[147].

In conclusion, the intentions of this thesis were to investigate new ways to reduce the burden of tobacco-related diseases which is still a challenge, but sometimes "forgotten and solved" in an era of obesity and diabetes.

Summary in Swedish

Populärvetenskaplig sammanfattning

Bakgrund och mål

Är det lika farligt för alla att röka? Behöver vissa mer hjälp än andra att sluta eller handlar det bara om att ”bestämma sig”? Löper vissa rökare dessutom större risk än andra att drabbas av rökningsrelaterad sjuklighet?

Kring år 2007 presenterades ny information kring nedärvda rökningsmönster (förklaras nedan) där det även framgick att en genetisk ofördelaktig förändring som påverkade rökningsbeteende, dessutom verkade påverka risken att drabbas av sjukdomar relaterade till rökningen.

Målet med denna avhandling var att undersöka huruvida nedärvda rökningsbeteenden kan identifieras och om det går att förutspå framtida insjuknanden och för tidig död hos de som har en ogynnsam genetisk profil och därmed kunna propagera för en mer riktad och aktiv behandling av rökstopp. Ytterligare mål var att undersöka huruvida det finns någon mätbar markör i ett vanligt blodprov som påverkas av rökning och eventuellt kan identifiera och förutse sjukdomsrisken hos rökare, både på lång och kort sikt.

Studie I och II

Genetiska förändringar i BDNF och CHRNA kopplas till rökning

Det är sedan länge känt att rökning är förknippat med sjukdomar som orsakar påtagligt stort lidande samt för tidig död. Några välkända exempel är cancer, kroniskt obstruktiv lungsjukdom (KOL) samt hjärtkärl-sjukdomar. Enligt WHO orsakas 6 miljoner dödsfall per år runt om i världen direkt av rökningens skadliga effekter och ett världsomfattande initiativ har tagits för att säkra den globala hälsan.

Eftersom nikotin är en beroendeframkallande substans har många som en gång börjat röka svårt att sluta, trots att det finns läkemedel som ersätter nikotinet eller dämpar suget efter nikotinets effekter. Genom bl. a. tvillingstudier har det under lång tid varit känt att nikotinbehovet har en ärftlig komponent men det finns även

faktorer i omgivningen som styr rökningsbeteendet såsom sociala miljöaspekter och att nikotinet påverkar det psykiska välbefinnandet för vissa rökare.

Under 2000-talet har det börjat bli enklare att söka igenom stora mängder DNA i jakten på att hitta förändringar i DNA-molekylens struktur av kvävebaser som sedan kan relateras till olika genuttryck. Genom att det sker en ändring (mutation) som medför att en specifik kvävebas blir annorlunda, kan det leda till ett förändrat genuttryck, vilket för individen ibland medför exempelvis en ökad känslighet för en viss sjukdom eller att ett specifikt beteende uppkommer. Förändringen i kvävebasstrukturen kallas för enbaspolymorfism, på engelska *single nucleotide polymorphism* (SNP*)*.

När det gäller rökning letade forskarna efter genetiska förändringar i kvävebasparen som förekom i signifikant högre utsträckning hos de som uppgav ett visst rökningsbeteende jämfört med de som uppgav att de *inte* hade beteendet. Exempel på olika rökningsbeteenden kan vara att man har större tendens att börja röka än andra, att man röker mer jämfört med andra rökare (dvs ökat nikotinbehov) eller uppger att man har svårare att sluta röka än vad andra rökare uppger.

Det gick att identifiera kopplingar mellan många olika enbaspolymorfismer och rökningsbeteenden men två särskilt starka kopplingar uppmärksammades. Den ena var lokaliserad på kromosom 11, *brain-derived neurotropic factor* (BDNF) och kunde relateras till en högre tendens att börja röka. Den andra enbaspolymorfismen var belägen på kromosom 15, *nicotinic acetylcholine receptor A* (CHRNA), och kunde förknippas med ökad mängd rökta cigaretter, alternativt ett ökat behov av nikotin. Intressant var även att BDNF-förändringen också kunde kopplas samman med ett högre BMI samt att CHRNA-förändringen även kopplades samman med insjuknande i lungcancer, perifer kärlsjukdom respektive KOL.

Metod

För att kunna bekräfta sambanden mellan genförändring och rökningsbeteende har vi använt en prospektiv (framåtblickande) befolkningsstudie ifrån Malmö, *Malmö Kost- och Cancer-studien*. Studiens ursprungliga plan var egentligen att se om det fanns samband mellan kostintag och insjuknande i de olika cancersjukdomarna.

Rekryteringen av Malmöbor startade under tidigt 90-tal och de som var skrivna i Malmö och födda mellan 1926–1946 fick ett brev där de informerades och tillfrågades om deltagande. Det sattes även upp information kring deltagande på allmänna platser i staden såsom vårdcentraler. Totalt anmälde sig drygt 30 000 (av totalt drygt 50 000 personer) som därmed blev kallade till ett mottagningsbesök där bl. a. vikt, längd och blodtryck registrerades. De fyllde i frågeformulär gällande kost, läkemedel, sjukdomar, socioekonomiskt status, rökstatus m.m. samt lämnade blodprov som analyserades men även frystes ned för framtida analyser. Samtliga deltagare avidentifierades och tilldelades ett deltagarnummer. Med hjälp av

personnummer och dödsorsaksregister samt andra register som förs över sjukdomsregistrering kan vi senare med hjälp av statistiska beräkningar, koppla samman uppgifter från baslinjen, dvs vid besökstillfället, och händelser framåt i tiden, s.k. prediktion.

De genetiska förändringarna, inklusive BDNF-polymorfismen och CHRNA-polymorfismen, analyserades i de ca 30 000 deltagarna och därefter kunde vi gå vidare med att studera om det gick att återfinna de samband som upptäckts mellan polymorfismerna och rökningsbeteende. Vidare kunde vi utforska mer okända områden, såsom huruvida det gick att förutse insjuknanden och dödsfall hos studiepopulationen. Med andra ord, vid baslinjesundersökningen delgavs vissa uppgifter samt information om vilken typ av BDNF- eller CHRNA-uppsättning individen hade. Kunde vi hitta associationer mellan de faktorerna och framtida insjuknande eller död?

Bekräftande av tidigare fynd samt nya samband i MKC-studien

För BDNF- och CHRNA-polymorfismerna kunde vi bekräfta kopplingarna till ökad tendens att röka, till ett högre BMI (BDNF) samt ökad mängd rökta cigaretter (CHRNA). Under uppföljningstiden som var >12 år kunde vi se signifikant säkerställda kopplingar mellan BDNF-polymorfismen och utfallet död, speciellt död från kardiovaskulära sjukdomar, hos de som uppgav att de var rökare. Sambanden gällde även om man bortsåg från antalet rökta cigaretter.

Det gick inte att förutspå dödsfall ifrån cancer eller insjuknande i kardiovaskulära sjukdomar under uppföljningen. Inga samband hittades mellan icke rökare och något av de studerade utfallen.

Avseende CHRNA förutspådde polymorfismen insjuknande i KOL, tobaksrelaterade cancrar samt död ifrån alla orsaker samt från luftvägarnas sjukdomar hos rökare. Inga samband hittades hos icke-rökare.

Sammanfattningsvis visar resultaten från befolkningsstudien i Malmö att nedärvt rökningsbeteende medför en högre sannolikhet både att röka och att drabbas av rökningens skadliga effekter *om man röker*. Det skulle därför kunna vara av betydande värde för individen, sjukvården och samhället att identifiera de rökarna, för att kunna intensifiera och individanpassa behandling mot ett rökstopp.

Studie III

Identifiering av en biomarkör som är förhöjd hos rökare

Human epididymis protein 4 (HE4) är ett protein som först identifierades i bitestikeln (epididymis) men som sedan kunnat påvisas i en mängd organ. I nuläget används den som en markör för äggstockscancer, där den är förhöjd p.g.a. att

specifika cancerceller utsöndrar HE4. Tidigare studier hade också påvisat att markören bl.a. steg med åldern, var högre hos män samt hos rökare, men orsaken till stegringen eller betydelsen av ett förhöjt HE4 var inte klarlagd.

Eftersom vi önskade hitta en markör som identifierade rökare med förhöjd risk att drabbas av rökningens komplikationer, undersökte vi till att börja med om HE4 var förhöjt hos rökare i en del av MKC-studien som benämns som Kardiovaskulära kohorten med drygt 6000 deltagare. Vi kunde också beräkna huruvida ett förhöjt värde av HE4 vid baslinje-undersökningen kunde förutspå dödsfall eller insjuknande i tobaksrelaterad sjuklighet under uppföljningstiden som var >14 år.

HE4 som prediktiv biomarkör oavsett rökstatus

Vi fann att rökare hade klart högre HE4-nivåer jämfört med icke-rökare (innefattar f.d. rökare samt de som aldrig rökt). Vi kunde också dra slutsatsen att höga värden förutspådde död av alla orsaker samt insjuknande i tobaksrelaterade sjukdomar både hos rökare och, lite oväntat, icke-rökare. Vår studie ger inte svar på vad som orsakar det funna sambandet men det förefaller som att HE4 är markör för generell vävnadsskada som kan vara ett mått på rökningens skadliga effekter men som även andra sjukdomsprocesser medför, oavsett rökning.

Studie IV

HE4 som riskmarkör i en befolkning med pågående sjukdom?

Populationen i studie III var vid insamlingen av blodproverna till största delen friska, varför vi nu valde att undersöka HE4 som riskmarkör i en befolkning med pågående sjukdomssymptom. Under 2013–2016 erbjöds patienter som sökte akutmottagningen i Malmö med symptomet dyspné (andningssvårigheter) att delta i en studie där utfallsmåttet var 90-dagars mortalitet. Studien kallas ADYS, en förkortning för Akut dyspné. Vid deltagarens akutbesök, "dag 0", gjordes mätningar och blodprovsanalyser enligt ett protokoll. Liksom i MKC studien analyserades blod men frystes även ned för senare analyser, deltagarna avidentifierades och ev. död inom 90 dagar efter akutbesöket registrerades i efterhand med hjälp av journalsystem och dödsregister.

HE4 kan förbättra uppskattningen av patientens risk för 90-dagars mortalitet

I den aktuella studien kunde inget samband mellan rökning och HE4 ses i de 799 deltagarna. Dock kunde vi se att det fanns ett statistiskt säkerställt samband mellan höga HE4-värden och död inom 90-dagar, oavsett rökstatus.

Andra mätvärden kontrolleras redan på akutmottagningen för att riskvärdera patienten och för att kunna ta ställning till allvarlighetsgrad, underliggande orsak till symptomet, behandling samt vilken grad av övervakning eller uppföljning som är

nödvändig. Exempel på mätvärden utöver blodprovsanalyser är syrgashalten i blodet, blodtryck, andningsfrekvens och puls. Genom att räkna fram ett statistiskt index kunde vi se att HE4 förbättrade säkerheten i bedömningen av risk för 90-dagars mortalitet om man adderade den till övriga, kända mätvärden. Vidare, om det funnits tillgång till HE4 hos de patienter som undersökts i ADYS så hade riskbedömningen kring 90-dagars död kunnat förbättras och bidragit till att omkring 7% hade kunnat skickas hem istället för att läggas in.

I framtiden skulle analys av HE4 hos patienter som söker akut för dyspné kunna vara värdefullt för personalen på akutmottagningen, där det på kort tid och med begränsade utredningsmöjligheter, kan vara svårt att värdera symptom och framtida risk.

Erratum

In Paper I, Table 1, the genotypes are switched, indicating that G/G is the genotype with the lowest frequency which is not correct. The correct distribution is as shown in Table 3, page 40.

Acknowledgements

Det finns mycket och många att uppmärksamma när ett projekt som det här håller på att få ett avslut. Framför allt måste jag tacka alla deltagare, forskningspersonal och inte minst initiativtagarna till de välgjorda befolkningsstudierna som ligger till grund för de studier som ingår i denna avhandling.

Olle Melander, min huvudhandledare, som på många sätt är vad som kan betecknas som motsägelsefull. I ett samarbete med dig slås jag av ett enormt driv men ändå avvaktan och tålmodighet, mänsklig i mycket men samtidigt en närmast omänsklig kapacitet. Alltid genomtänkta svar men också spontanitet, forskningstiden med dig har varit extremt utvecklande och arbetsam men lika ofta festlig, rolig och social! Jag är oerhört tacksam över att ha varit din doktorand.

Viktor Hamrefors, min bihandledare, som på grund av en sällsynt kombination av hög teoretisk begåvning och pedagogisk kompetens, varit helt oersättlig i den här processen. Din noggrannhet och strävan att alltid kunna göra en lite bättre formulering smittar av sig! Jag är väldigt glad och lite stolt över att ha fått dra nytta av din kompetens, det kommer att gå strålande för dig både i forskningsvärlden *och* i kliniskt arbete där jag hoppas att medicinkliniken även i framtiden lyckas behålla dig.

Dessutom vill jag tacka samtliga **medförfattare** som bidragit med med uppmuntrande och kvalitetshöjande kommentarer, och ett särskilt tack till **Peter Almgren** för ditt tålamod att förklara och lyssna på försök till tolkningar av dina statistiska beräkningar. Och tack till **Gunilla Hughes-Wulkan**, för att du alltid är så optimistisk och löser saker innan man hinner oroa sig.

Sofia Enhörning, oavsett om du är släkt med en viss Bonaparte eller inte, så har du många storslagna egenskaper. Det jag uppskattar mycket hos dig är din ödmjukhet inför kunskap, att du alltid har nära till skratt men också att du vågar ifrågasätta det som för de flesta andra tas för givet. Att få ta del av dina forskningserfarenheter har gjort min doktorandresa smidigare, men inte minst så uppskattar jag vår vänskap väldigt mycket.

Lunchgruppen, med lite olika medlemmar men en kärna av **Céline**, **Marketa**, **Filip**, **Widet**, **Nathalie** och **Erik**, många "polletter har trillat ner" under våra luncher och ärlig och konstruktiv återkoppling har givits på presentationsteknik och innehåll. Ovärderligt!

Och alla, utöver redan nämnda, som varit med på forskningsresor runt om i Europa med **Olle** i spetsen både dag och natt: **Jonas, Martin, Artur, Erasmus, Klas, Irina, Jasmina, Marcus, Patrik, Hannes, John, Ayesha, Amra** och säkert ännu fler…!

Från min arbetsplats, Internmedicinska kliniken, vill jag tacka verksamhetschef **Maria Ohlson-Andersson** som värdesätter klinikens forskning vilket gjort att jag kunnat ta ut den forskningstid jag har önskat. Tack även till **Marie Roos** och **Monika Östergren** som genom smidigt samarbete underlättar samordningen mellan forskning, klinik och vardag.

Oskar Hammar och **Linn Kennedy,** nu som orädda sektionschefer på Internmedicin men också mina f.d. handledare. Ni är båda förebilder och ovärderliga för min tid på medicinkliniken med er förmåga att vara lyhörda och intresserade för att förbättra och utveckla samt få mig att sträva efter utmaningar.

Fredrik von Wowern, tack för att du håller medicinklinikens fana högt och gör det kul att arbeta som läkare och alltid ser en ljuspunkt eller lösning trots att det ibland känns lite hopplöst.

Fredrik Buchwald, tack för att du som min handledare på Neurologiska kliniken, gav ovärderlig klinisk handledning med höga men inte orimliga krav.

Bokklubben med de polyglotta donnorna **Parisa Mokarami, Petrea Frid och Teresa Ullberg,** som jag en gång i tiden tog initiativet till, var ju bara en förevändning för att få träffa er oftare! Tack vare att vi läser en massa bra böcker och ser till att träffas så fort vi får ihop det, får jag följa era livspussel, kloka och lärorika och ibland gränslösa tankar om i princip allt. Måttfullheten har ingen plats i den här bokklubben!

Nathalie Lund, vi fann varandra snabbt över en rond på Medicinavd 4 för många år sedan nu. Du har en beundransvärd inställning till både livet och arbetslivet och jag uppskattar din ärlighet mycket.

Per Bengtsson, numera allmänläkare och pappa istället för vapendragare på medicinkliniken… tack för att du, med väl genomtänkta formuleringar, ser det roliga och ibland dråpliga i det mesta och tipsar om upplevelser och strapatser som gör vardagen guldkantad.

Anna Blomkvist, tack för att du tog befälet i marschen som ledde mot läkarprogrammet, att du alltid uppmuntrat min forskning och att jag får ta del av dina finurliga perspektiv på det mesta, du borde starta en blogg åtminstone!

Jennie Wickenberg, att du och jag träffades var en av de bästa sakerna som hände på Panum. Tack för att du visat mig att i princip allt är möjligt, bara man bestämmer sig.

Andrea, Brita, Hanna, Lina, och Lisa T, tänk vilken tur att vi träffades på gymnasiet!! Tack för lång, sjukt rolig, trygg, inspirerande, hoppingivande och varm vänskap.

Petra Olsson-Gendt, tack för alla stunder med samtal, yoga, naturvin, klokhet, flärd, mod, oro, sorg, havsbad, tångbad, picnic-korgar, nya kläder, inspiration och livsglädje som en vänskap med dig erbjuder. Och…tack för att du gav denna avhandling ett stolt omslag.

Carin, Carolina, Elin och Sanna! Ladies in the hood, jag blir lite gladare och lite klokare varje gång jag vi ses (även om det är på Ica Maxi).

Mamma, mamsen, MLZH, Marie-Louise, Misse, mormor, amaryllis…du går under många namn och lämnar alltid ett avtryck hos de du träffar. Det är tacksamt att vara din dotter! Din beslutsamhet, nyfikenhet, öppenhet och strävan efter att utveckla, inspirera och att själv utvecklas och bli inspirerad, är det få som matchar. Du är anledningen till att jag började på läkarprogrammet och sedermera började forska, du är också anledningen till att jag inte slutar leta efter nya utmaningar och möjligheter, tack…!

Pappa, måhända att du är lite väl generös mot dina barnbarn, annars har jag ingenting att invända. Tack för att du är en pappa som alltid ställer upp, aldrig dömer, alltid respekterar beslut och visar fullt förtroende.

Lisa, rent utseendemässigt kanske vi inte ser ut som tvillingar men på insidan känner jag en nästan utomjordisk känsla av samhörighet med dig (fastän jag inte riktigt tror på sånt). Det lugnet som infann sig i mig när du, **André, Hedda och Jakob** bestämde er för att flytta ner till Malmö var överrumplande, och sen kom ju **Sonja** också… Jag blir faktiskt fortfarande lycklig när jag tänker på att ni bor så nära oss, trots att ni gjort det i nästan två år!

Alex, av dig lär jag mig saker varenda dag, om dig, om livet och om mig själv. Dessutom har du har en förmåga att göra så att allting blir *lite* bättre och *lite* roligare när du är med. Jag kommer fortsätta att hålla mig så nära dig som det bara går.

Mikkel och **Harriet**, tänk att ni finns också! Få överraskar mig så mycket som ni och min glädje över att ni finns hoppas jag att ni märker av i varje sekund.

Mormor Margareta, för alla år av omtanke och kärlek. **Moster Tita**, bättre moster går faktiskt inte att tänka sig, vare sig jag är ett barn eller nu nästan 40 år!

Anders, Brita, Kicke, Maddis, Kim och lilla Isabel, min tredje familj! Tack för alla härliga äventyr som ni drar med oss på, ser fram emot kommande stunder på Österlen, Oslo, Köpenhamn och …….?

References

1. WHO. WHO global report on trends in prevalence of tobacco smoking. 2015.
2. WHO. WHO technical manual on tobacco tax administration. 2010.
3. Caggiula AR, Donny EC, White AR, et al. Cue dependency of nicotine self-administration and smoking. Pharmacol Biochem Behav. 2001;70(4):515-530.
4. Stolerman IP, Shoaib M. The neurobiology of tobacco addiction. Trends Pharmacol Sci. 1991;12(12):467-473.
5. Stolerman IP, Jarvis MJ. The scientific case that nicotine is addictive. Psychopharmacology (Berl). 1995;117(1):2-10; discussion 14-20.
6. Rose JE. Multiple brain pathways and receptors underlying tobacco addiction. Biochem Pharmacol. 2007;74(8):1263-1270.
7. Balfour DJ. Neural mechanisms underlying nicotine dependence. Addiction. 1994;89(11):1419-1423.
8. Carmelli D, Swan GE, Robinette D, Fabsitz R. Genetic influence on smoking--a study of male twins. N Engl J Med. 1992;327(12):829-833.
9. Hughes JR. Genetics of Smoking: A Brief Review. Behaviour Therapy. 1986;17(4):335-345.
10. Swan GE, Carmelli D, Rosenman RH, Fabsitz RR, Christian JC. Smoking and alcohol consumption in adult male twins: genetic heritability and shared environmental influences. J Subst Abuse. 1990;2(1):39-50.
11. Gabrielsen ME, Romundstad P, Langhammer A, Krokan HE, Skorpen F. Association between a 15q25 gene variant, nicotine-related habits, lung cancer and COPD among 56,307 individuals from the HUNT study in Norway. Eur J Hum Genet. 2013;21(11):1293-1299.
12. Saccone SF, Hinrichs AL, Saccone NL, et al. Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. Hum Mol Genet. 2007;16(1):36-49.
13. Bierut LJ, Madden PA, Breslau N, et al. Novel genes identified in a high-density genome wide association study for nicotine dependence. Hum Mol Genet. 2007;16(1):24-35.
14. Lips EH, Gaborieau V, McKay JD, et al. Association between a 15q25 gene variant, smoking quantity and tobacco-related cancers among 17 000 individuals. Int J Epidemiol. 2010;39(2):563-577.
15. Chen X, Chen J, Williamson VS, et al. Variants in nicotinic acetylcholine receptors alpha5 and alpha3 increase risks to nicotine dependence. Am J Med Genet B Neuropsychiatr Genet. 2009;150B(7):926-933.

16. Spitz MR, Amos CI, Dong Q, Lin J, Wu X. The CHRNA5-A3 region on chromosome 15q24-25.1 is a risk factor both for nicotine dependence and for lung cancer. J Natl Cancer Inst. 2008;100(21):1552-1556.
17. Saccone NL, Wang JC, Breslau N, et al. The CHRNA5-CHRNA3-CHRNB4 nicotinic receptor subunit gene cluster affects risk for nicotine dependence in African-Americans and in European-Americans. Cancer Res. 2009;69(17):6848-6856.
18. Kaur-Knudsen D, Nordestgaard BG, Bojesen SE. CHRNA3 genotype, nicotine dependence, lung function and disease in the general population. Eur Respir J. 2012;40(6):1538-1544.
19. Jamal M, Van der Does W, Elzinga BM, Molendijk ML, Penninx BW. Association between smoking, nicotine dependence, and BDNF Val66Met polymorphism with BDNF concentrations in serum. Nicotine Tob Res. 2015;17(3):323-329.
20. Tobacco, Genetics C. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet. 2010;42(5):441-447.
21. Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature. 2008;452(7187):638-642.
22. Thorgeirsson TE, Gudbjartsson DF, Surakka I, et al. Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. Nat Genet. 2010;42(5):448-453.
23. Munafo MR, Timofeeva MN, Morris RW, et al. Association between genetic variants on chromosome 15q25 locus and objective measures of tobacco exposure. J Natl Cancer Inst. 2012;104(10):740-748.
24. Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature. 2008;452(7187):633-637.
25. Rafnar T, Sulem P, Besenbacher S, et al. Genome-wide significant association between a sequence variant at 15q15.2 and lung cancer risk. Cancer Res. 2011;71(4):1356-1361.
26. Kaur-Knudsen D, Bojesen SE, Nordestgaard BG. CHRNA3 and CYP3A5*3 genotype, lung function and chronic obstructive pulmonary disease in the general population. Pharmacogenet Genomics. 2014;24(4):220-229.
27. Kaur-Knudsen D, Bojesen SE, Tybjaerg-Hansen A, Nordestgaard BG. Nicotinic acetylcholine receptor polymorphism, smoking behavior, and tobacco-related cancer and lung and cardiovascular diseases: a cohort study. J Clin Oncol. 2011;29(21):2875-2882.
28. Timofeeva MN, McKay JD, Smith GD, et al. Genetic polymorphisms in 15q25 and 19q13 loci, cotinine levels, and risk of lung cancer in EPIC. Cancer Epidemiol Biomarkers Prev. 2011;20(10):2250-2261.
29. Steffensen KD, Waldstrom M, Brandslund I, et al. Identification of high-risk patients by human epididymis protein 4 levels during follow-up of ovarian cancer. Oncol Lett. 2016;11(6):3967-3974.
30. Hellstrom I, Raycraft J, Hayden-Ledbetter M, et al. The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma. Cancer Res. 2003;63(13):3695-3700.

31. Holcomb K, Vucetic Z, Miller MC, Knapp RC. Human epididymis protein 4 offers superior specificity in the differentiation of benign and malignant adnexal masses in premenopausal women. Am J Obstet Gynecol. 2011;205(4):358 e351-356.
32. Steffensen KD, Waldstrom M, Brandslund I, Petzold M, Jakobsen A. The prognostic and predictive value of combined HE4 and CA-125 in ovarian cancer patients. Int J Gynecol Cancer. 2012;22(9):1474-1482.
33. Kamei M, Yamashita S, Tokuishi K, et al. HE4 expression can be associated with lymph node metastases and disease-free survival in breast cancer. Anticancer Res. 2010;30(11):4779-4783.
34. Hertlein L, Stieber P, Kirschenhofer A, et al. Human epididymis protein 4 (HE4) in benign and malignant diseases. Clin Chem Lab Med. 2012;50(12):2181-2188.
35. Bolstad N, Oijordsbakken M, Nustad K, Bjerner J. Human epididymis protein 4 reference limits and natural variation in a Nordic reference population. Tumour Biol. 2012;33(1):141-148.
36. Urban N, Thorpe J, Karlan BY, et al. Interpretation of single and serial measures of HE4 and CA125 in asymptomatic women at high risk for ovarian cancer. Cancer Epidemiol Biomarkers Prev. 2012;21(11):2087-2094.
37. Nagy B, Jr., Bhattoa HP, Steiber Z, et al. Serum human epididymis protein 4 (HE4) as a tumor marker in men with lung cancer. Clin Chem Lab Med. 2014;52(11):1639-1648.
38. Karlsen NS, Karlsen MA, Hogdall CK, Hogdall EV. HE4 tissue expression and serum HE4 levels in healthy individuals and patients with benign or malignant tumors: a systematic review. Cancer Epidemiol Biomarkers Prev. 2014;23(11):2285-2295.
39. Bilano V, Gilmour S, Moffiet T, et al. Global trends and projections for tobacco use, 1990-2025: an analysis of smoking indicators from the WHO Comprehensive Information Systems for Tobacco Control. Lancet. 2015;385(9972):966-976.
40. Daily smokers. OECD 2017.
41. Hammond EC, Horn D. The relationship between human smoking habits and death rates: a follow-up study of 187,766 men. J Am Med Assoc. 1954;155(15):1316-1328.
42. Doll R, Hill AB. The mortality of doctors in relation to their smoking habits; a preliminary report. Br Med J. 1954;1(4877):1451-1455.
43. Pillai SG, Ge D, Zhu G, et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. PLoS Genet. 2009;5(3):e1000421.
44. Amos CI, Wu X, Broderick P, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet. 2008;40(5):616-622.
45. Johnson EO, Novak SP. Onset and persistence of daily smoking: the interplay of socioeconomic status, gender, and psychiatric disorders. Drug Alcohol Depend. 2009;104 Suppl 1:S50-57.
46. Balfour DJ, Fagerstrom KO. Pharmacology of nicotine and its therapeutic use in smoking cessation and neurodegenerative disorders. Pharmacol Ther. 1996;72(1):51-81.

47. Verification SSoB. Biochemical verification of tobacco use and cessation. Nicotine Tob Res. 2002;4(2):149-159.
48. Benowitz NL. Cotinine as a biomarker of environmental tobacco smoke exposure. Epidemiol Rev. 1996;18(2):188-204.
49. Etter JF, Vu Duc T, Perneger TV. Saliva cotinine levels in smokers and nonsmokers. Am J Epidemiol. 2000;151(3):251-258.
50. Doll R, Peto R, Wheatley K, Gray R, Sutherland I. Mortality in relation to smoking: 40 years' observations on male British doctors. BMJ. 1994;309(6959):901-911.
51. Peto R, Darby S, Deo H, Silcocks P, Whitley E, Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. BMJ. 2000;321(7257):323-329.
52. Kawachi I, Colditz GA, Stampfer MJ, et al. Smoking cessation and decreased risk of stroke in women. JAMA. 1993;269(2):232-236.
53. Kawachi I, Colditz GA, Stampfer MJ, et al. Smoking cessation in relation to total mortality rates in women. A prospective cohort study. Ann Intern Med. 1993;119(10):992-1000.
54. Benowitz NL. Pharmacology of nicotine: addiction, smoking-induced disease, and therapeutics. Annu Rev Pharmacol Toxicol. 2009;49:57-71.
55. Gilljam H. Tobaksavvänjning. Läkemedelsverket;2014.
56. Stead LF, Perera R, Bullen C, et al. Nicotine replacement therapy for smoking cessation. Cochrane Database Syst Rev. 2012;11:CD000146.
57. Causes of death 2008. World Health Organization, Geneva;2008.
58. Global status report on noncommunicable diseases WHO, Geneva;2010.
59. Ross R. Atherosclerosis--an inflammatory disease. N Engl J Med. 1999;340(2):115-126.
60. Davis NE. Atherosclerosis--an inflammatory process. J Insur Med. 2005;37(1):72-75.
61. Global Atlas on cardiovascular disease control and prevention. WHO;2011.
62. Celermajer DS, Sorensen KE, Gooch VM, et al. Non-invasive detection of endothelial dysfunction in children and adults at risk of atherosclerosis. Lancet. 1992;340(8828):1111-1115.
63. Zeiher AM, Schachinger V, Minners J. Long-term cigarette smoking impairs endothelium-dependent coronary arterial vasodilator function. Circulation. 1995;92(5):1094-1100.
64. Messner B, Bernhard D. Smoking and cardiovascular disease: mechanisms of endothelial dysfunction and early atherogenesis. Arterioscler Thromb Vasc Biol. 2014;34(3):509-515.
65. WHO. Fact sheet Cancer. 2017.
66. Collaborators GBDRF. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016;388(10053):1659-1724.
67. Roberts DL. Natural tobacco flavor. Recent Adv Tob Sci. 1988;14:49–81.

68. Hoffmann DH, I. . The changing cigarette: Chemical studies and bioassays. In:Risks Associated with Smoking Cigarettes with Low Machine-Measured Yields of Tar and Nicotine NIH Publ.: National Cancer Institute;2001.
69. Proctor RN. The history of the discovery of the cigarette-lung cancer link: evidentiary traditions, corporate denial, global toll. Tob Control. 2012;21(2):87-91.
70. WHO. IARC Monographs on the evaluation of carcinogenic risks to humans. Tobacco smoke and involuntary smoking.: International institute for research on cancer;2004.
71. Secretan B, Straif K, Baan R, et al. A review of human carcinogens--Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish. Lancet Oncol. 2009;10(11):1033-1034.
72. Agudo A, Bonet C, Travier N, et al. Impact of cigarette smoking on cancer risk in the European prospective investigation into cancer and nutrition study. J Clin Oncol. 2012;30(36):4550-4557.
73. WHO. Mortality attributable to tobacco. 2004.
74. WHO. Global surveillance, prevention and control of chronic respiratory diseases WHO;2007.
75. Eisner MD, Anthonisen N, Coultas D, et al. An official American Thoracic Society public policy statement: Novel risk factors and the global burden of chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2010;182(5):693-718.
76. Genomes Project C, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491(7422):56-65.
77. Kruglyak L, Nickerson DA. Variation is the spice of life. Nat Genet. 2001;27(3):234-236.
78. Zaitlen N, Kraft P, Patterson N, et al. Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. PLoS Genet. 2013;9(5):e1003520.
79. Bjorkegren JL, Kovacic JC, Dudley JT, Schadt EE. Genome-wide significant loci: how important are they? Systems genetics to understand heritability of coronary artery disease and other common complex disorders. J Am Coll Cardiol. 2015;65(8):830-845.
80. International HapMap C. The International HapMap Project. Nature. 2003;426(6968):789-796.
81. International HapMap C. A haplotype map of the human genome. Nature. 2005;437(7063):1299-1320.
82. Strachan TR, A. Human Molecular Genetics. Vol 4th edition: Garland Science; 2011.
83. Thorleifsson G, Walters GB, Gudbjartsson DF, et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nat Genet. 2009;41(1):18-24.
84. Kernie SG, Liebl DJ, Parada LF. BDNF regulates eating behavior and locomotor activity in mice. EMBO J. 2000;19(6):1290-1300.
85. Hashimoto K, Koizumi H, Nakazato M, Shimizu E, Iyo M. Role of brain-derived neurotrophic factor in eating disorders: recent findings and its pathophysiological implications. Prog Neuropsychopharmacol Biol Psychiatry. 2005;29(4):499-504.

86. Lu B, Figurov A. Role of neurotrophins in synapse development and plasticity. Rev Neurosci. 1997;8(1):1-12.
87. Kuczewski N, Porcher C, Lessmann V, Medina I, Gaiarsa JL. Activity-dependent dendritic release of BDNF and biological consequences. Mol Neurobiol. 2009;39(1):37-49.
88. Murer MG, Yan Q, Raisman-Vozari R. Brain-derived neurotrophic factor in the control human brain, and in Alzheimer's disease and Parkinson's disease. Prog Neurobiol. 2001;63(1):71-124.
89. Notaras M, Hill R, van den Buuse M. A role for the BDNF gene Val66Met polymorphism in schizophrenia? A comprehensive review. Neurosci Biobehav Rev. 2015;51:15-30.
90. Gratacos M, Gonzalez JR, Mercader JM, de Cid R, Urretavizcaya M, Estivill X. Brain-derived neurotrophic factor Val66Met and psychiatric disorders: meta-analysis of case-control studies confirm association to substance-related disorders, eating disorders, and schizophrenia. Biol Psychiatry. 2007;61(7):911-922.
91. Grimm JW, Lu L, Hayashi T, Hope BT, Su TP, Shaham Y. Time-dependent increases in brain-derived neurotrophic factor protein levels within the mesolimbic dopamine system after withdrawal from cocaine: implications for incubation of cocaine craving. J Neurosci. 2003;23(3):742-747.
92. Nakahashi T, Fujimura H, Altar CA, et al. Vascular endothelial cells synthesize and secrete brain-derived neurotrophic factor. FEBS Lett. 2000;470(2):113-117.
93. Ejiri J, Inoue N, Kobayashi S, et al. Possible role of brain-derived neurotrophic factor in the pathogenesis of coronary artery disease. Circulation. 2005;112(14):2114-2120.
94. Kim TS, Kim DJ, Lee H, Kim YK. Increased plasma brain-derived neurotrophic factor levels in chronic smokers following unaided smoking cessation. Neurosci Lett. 2007;423(1):53-57.
95. Kaess BM, Preis SR, Lieb W, et al. Circulating brain-derived neurotrophic factor concentrations and the risk of cardiovascular disease in the community. J Am Heart Assoc. 2015;4(3):e001544.
96. Jackson KJ, Fanous AH, Chen J, Kendler KS, Chen X. Variants in the 15q25 gene cluster are associated with risk for schizophrenia and bipolar disorder. Psychiatr Genet. 2013;23(1):20-28.
97. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511(7510):421-427.
98. Berrettini W, Yuan X, Tozzi F, et al. Alpha-5/alpha-3 nicotinic receptor subunit alleles increase risk for heavy smoking. Mol Psychiatry. 2008;13(4):368-373.
99. Bierut LJ, Stitzel JA, Wang JC, et al. Variants in nicotinic receptors and risk for nicotine dependence. Am J Psychiatry. 2008;165(9):1163-1171.
100. Lam DC, Girard L, Ramirez R, et al. Expression of nicotinic acetylcholine receptor subunit genes in non-small-cell lung cancer reveals differences between smokers and nonsmokers. Cancer Res. 2007;67(10):4638-4647.

101. Schuller HM. Cell type specific, receptor-mediated modulation of growth kinetics in human lung cancer cell lines by nicotine and tobacco-related nitrosamines. Biochem Pharmacol. 1989;38(20):3439-3442.

102. Egleton RD, Brown KC, Dasgupta P. Nicotinic acetylcholine receptors in cancer: multiple roles in proliferation and inhibition of apoptosis. Trends Pharmacol Sci. 2008;29(3):151-158.

103. Tonstad S, Tonnesen P, Hajek P, et al. Effect of maintenance therapy with varenicline on smoking cessation: a randomized controlled trial. JAMA. 2006;296(1):64-71.

104. Picciotto MR, Zoli M, Rimondini R, et al. Acetylcholine receptors containing the beta2 subunit are involved in the reinforcing properties of nicotine. Nature. 1998;391(6663):173-177.

105. Tapper AR, McKinney SL, Nashmi R, et al. Nicotine activation of alpha4* receptors: sufficient for reward, tolerance, and sensitization. Science. 2004;306(5698):1029-1032.

106. Kirchhoff C, Habben I, Ivell R, Krull N. A major human epididymis-specific cDNA encodes a protein with sequence homology to extracellular proteinase inhibitors. Biol Reprod. 1991;45(2):350-357.

107. Galgano MT, Hampton GM, Frierson HF, Jr. Comprehensive analysis of HE4 expression in normal and malignant human tissues. Mod Pathol. 2006;19(6):847-853.

108. Lu KH, Patterson AP, Wang L, et al. Selection of potential markers for epithelial ovarian cancer with gene expression arrays and recursive descent partition analysis. Clin Cancer Res. 2004;10(10):3291-3300.

109. LeBleu VS, Teng Y, O'Connell JT, et al. Identification of human epididymis protein-4 as a fibroblast-derived mediator of fibrosis. Nat Med. 2013;19(2):227-231.

110. de Boer RA, Cao Q, Postmus D, et al. The WAP four-disulfide core domain protein HE4: a novel biomarker for heart failure. JACC Heart Fail. 2013;1(2):164-169.

111. Piek A, Meijers WC, Schroten NF, Gansevoort RT, de Boer RA, Sillje HH. HE4 Serum Levels Are Associated with Heart Failure Severity in Patients With Chronic Heart Failure. J Card Fail. 2016.

112. Bingle L, Singleton V, Bingle CD. The putative ovarian tumour marker gene HE4 (WFDC2), is expressed in normal tissues and undergoes complex alternative splicing to yield multiple protein isoforms. Oncogene. 2002;21(17):2768-2773.

113. Escudero JM, Auge JM, Filella X, Torne A, Pahisa J, Molina R. Comparison of serum human epididymis protein 4 with cancer antigen 125 as a tumor marker in patients with malignant and nonmalignant diseases. Clin Chem. 2011;57(11):1534-1544.

114. Yang J, Sa M, Huang M, et al. The reference intervals for HE4, CA125 and ROMA in healthy female with electrochemiluminescence immunoassay. Clin Biochem. 2013;46(16-17):1705-1708.

115. Parshall MB, Schwartzstein RM, Adams L, et al. An official American Thoracic Society statement: update on the mechanisms, assessment, and management of dyspnea. Am J Respir Crit Care Med. 2012;185(4):435-452.

116. Gori CS, Magrini L, Travaglino F, Di Somma S. Role of biomarkers in patients with dyspnea. Eur Rev Med Pharmacol Sci. 2011;15(2):229-240.

117. Michelson E, Hollrah S. Evaluation of the patient with shortness of breath: an evidence based approach. Emerg Med Clin North Am. 1999;17(1):221-237, x.
118. Maisel A, Mueller C, Nowak R, et al. Mid-region pro-hormone markers for diagnosis and prognosis in acute dyspnea: results from the BACH (Biomarkers in Acute Heart Failure) trial. J Am Coll Cardiol. 2010;55(19):2062-2076.
119. Heining L, Giesa C, Ewig S. MR-proANP, MR-proADM, and PCT in Patients Presenting with Acute Dyspnea in a Medical Emergency Unit. Lung. 2016;194(2):185-191.
120. Dieplinger B, Gegenhuber A, Kaar G, Poelz W, Haltmayer M, Mueller T. Prognostic value of established and novel biomarkers in patients with shortness of breath attending an emergency department. Clin Biochem. 2010;43(9):714-719.
121. Villacorta H, Duarte A, Duarte NM, et al. The role of B-type natriuretic peptide in the diagnosis of congestive heart failure in patients presenting to an emergency department with dyspnea. Arq Bras Cardiol. 2002;79(6):569-572, 564-568.
122. Berglund G, Elmstahl S, Janzon L, Larsson SA. The Malmo Diet and Cancer Study. Design and feasibility. J Intern Med. 1993;233(1):45-51.
123. Manjer J, Carlsson S, Elmstahl S, et al. The Malmo Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. Eur J Cancer Prev. 2001;10(6):489-499.
124. Rosvall M, Janzon L, Berglund G, Engstrom G, Hedblad B. Incident coronary events and case fatality in relation to common carotid intima-media thickness. J Intern Med. 2005;257(5):430-437.
125. Wessman TT, R.; Melander, O.; Gränsbo, K. Increased mortality among acute dyspnea patients from immigrant dense urban districts.
126. Widgren BR, Jourak M. Medical Emergency Triage and Treatment System (METTS): a new protocol in primary triage and secondary priority decision in emergency medicine. J Emerg Med. 2011;40(6):623-628.
127. Johansson LA, Westerling R. Comparing Swedish hospital discharge records with death certificates: implications for mortality statistics. Int J Epidemiol. 2000;29(3):495-502.
128. Hammar N, Alfredsson L, Rosen M, Spetz CL, Kahan T, Ysberg AS. A national record linkage to study acute myocardial infarction incidence and case fatality in Sweden. Int J Epidemiol. 2001;30 Suppl 1:S30-34.
129. Jerntorp P, Berglund G. Stroke registry in Malmo, Sweden. Stroke. 1992;23(3):357-361.
130. Ludvigsson JF, Andersson E, Ekbom A, et al. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450.
131. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008;27(2):157-172; discussion 207-112.
132. Gunstad J, Schofield P, Paul RH, et al. BDNF Val66Met polymorphism is associated with body mass index in healthy adults. Neuropsychobiology. 2006;53(3):153-156.

133. Adler NE, Ostrove JM. Socioeconomic status and health: what we know and what we don't. Ann N Y Acad Sci. 1999;896:3-15.
134. Sommer I, Griebler U, Mahlknecht P, et al. Socioeconomic inequalities in non-communicable diseases and their risk factors: an overview of systematic reviews. BMC Public Health. 2015;15:914.
135. Saccone NL, Culverhouse RC, Schwantes-An TH, et al. Multiple independent loci at chromosome 15q25.1 affect smoking quantity: a meta-analysis and comparison with lung cancer and COPD. PLoS Genet. 2010;6(8).
136. Russo P, Cardinale A, Margaritora S, Cesario A. Nicotinic receptor and tobacco-related cancer. Life Sci. 2012;91(21-22):1087-1092.
137. Wang JC, Cruchaga C, Saccone NL, et al. Risk for nicotine dependence and lung cancer is conferred by mRNA expression levels and amino acid change in CHRNA5. Hum Mol Genet. 2009;18(16):3125-3135.
138. Gage SH, Davey Smith G, Ware JJ, Flint J, Munafo MR. G = E: What GWAS Can Tell Us about the Environment. PLoS Genet. 2016;12(2):e1005765.
139. Shipton D, Tappin DM, Vadiveloo T, Crossley JA, Aitken DA, Chalmers J. Reliability of self reported smoking status by pregnant women for estimating smoking prevalence: a retrospective, cross sectional study. BMJ. 2009;339:b4347.
140. Krall EA, Valadian I, Dwyer JT, Gardner J. Accuracy of recalled smoking data. Am J Public Health. 1989;79(2):200-202.
141. Bloom AJ, Hartz SM, Baker TB, et al. Beyond cigarettes per day. A genome-wide association study of the biomarker carbon monoxide. Ann Am Thorac Soc. 2014;11(7):1003-1010.
142. Zacny JP, Stitzer ML, Brown FJ, Yingling JE, Griffiths RR. Human cigarette smoking: effects of puff and inhalation parameters on smoke exposure. J Pharmacol Exp Ther. 1987;240(2):554-564.
143. Perkins KA, Karelitz JL, Giedgowd GE, Conklin CA. The reliability of puff topography and subjective responses during ad lib smoking of a single cigarette. Nicotine Tob Res. 2012;14(4):490-494.
144. Hansen K, Ostling G, Persson M, et al. The effect of smoking on carotid intima-media thickness progression rate and rate of lumen diameter reduction. Eur J Intern Med. 2016;28:74-79.
145. Appelros P, Terent A. Validation of the Swedish inpatient and cause-of-death registers in the context of stroke. Acta Neurol Scand. 2011;123(4):289-293.
146. Goniewicz ML, Knysak J, Gawron M, et al. Levels of selected carcinogens and toxicants in vapour from electronic cigarettes. Tob Control. 2014;23(2):133-139.
147. Drummond CA, Crotty Alexander LE, Haller ST, et al. Cigarette Smoking Causes Epigenetic Changes Associated With Cardiorenal Fibrosis. Physiol Genomics. 2016:physiolgenomics 00070 02016.

Paper I

ORIGINAL ARTICLE

Smoking and obesity associated *BDNF* gene variance predicts total and cardiovascular mortality in smokers

Sara Halldén,[1] Marketa Sjögren,[1] Bo Hedblad,[1] Gunnar Engström,[1] Krzysztof Narkiewicz,[2] Michal Hoffmann,[2] Björn Wahlstrand,[3] Thomas Hedner,[3] Olle Melander[1]

[1]Department of Clinical Sciences, Lund University, Malmö, Sweden
[2]Department of Hypertension and Diabetology, Medical University of Gdansk, Gdansk, Poland
[3]Institute of Medicine at Sahlgrenska Academy, University of Gothenburg, Sweden

Correspondence to
Sara Halldén, Department of Clinical Sciences, Lund University, Clinical Research Center, Entrance 72, Building 91, Floor 12, Malmö University Hospital, Malmö SE 205 02, Sweden;
sara.hallden@med.lu.se

Received 11 January 2013
Revised 4 April 2013
Accepted 5 April 2013
Published Online First
27 April 2013

ABSTRACT

Objective The brain derived neurotrophic factor (BDNF) locus has been implicated in psychiatric and substance related disorders. Recent genome-wide association studies (GWAS) have shown strong associations between single nucleotide polymorphisms in BDNF, smoking behaviour and high body mass index (BMI). Our aim was to test whether genetic BDNF variation alters the risk of smoking related morbidity and mortality.

Design Cox proportional hazards models were used to relate the BDNF rs4923461(A/G) polymorphisms to all-cause, cancer and cardiovascular mortality and cardiovascular disease (CVD) incidence adjusted for age, sex, BMI, and smoking quantity.

Setting The Malmö Diet and Cancer Study (MDCS), a population based prospective cohort study (n=30 447).

Patients We obtained complete data on 25 071 subjects, of whom 6507 were current smokers and 18 564 were non-smokers who underwent a baseline examination from 1991–1996.

Main outcome measures During a mean follow-up time of 12 years, 1049 deaths (346 cardiovascular deaths and 492 cancer deaths) and 802 incident CVD events occurred among current smokers.

Results The major allele (A) of rs4923461 was significantly associated with ever having smoked (p=0.03) and high BMI (p=0.001). The A-allele was associated with risk of all-cause (HR=1.12, 95% CI 1.00 to 1.25; p<0.05) and CVD (HR=1.23, 95% CI 1.01 to 1.49; p=0.04) mortality. There was no significant association between the rs4923461 and cancer mortality or CVD incidence.

Conclusions Our data suggest that smoking- and obesity-associated variation of the *BDNF* gene affects the risk of death, especially due to cardiovascular causes, in smokers. Determination of the BDNF genotype in smokers may guide the need for smoking cessation interventions.

INTRODUCTION

Cigarette smoking accounts for several adverse health effects. Despite increasing awareness, more than one billion people worldwide smoke tobacco daily. Recent reports estimate that smoking accounts for nearly one of every five deaths each year in the USA. In addition to a distinct correlation with a number of cancer diagnoses, smoking is also estimated to augment the risk of coronary heart disease and stroke by two to four times, when compared to non-smokers.[1]

Variance in smoking behaviour is influenced by both psychosocial factors and genetic disposition; however, tools for assessing future smoking related complications are lacking.[2 3] Previous genome-wide association studies (GWAS) have identified multiple loci related to different smoking phenotypes. These include the brain derived neurotrophic factor (BDNF) locus on chromosome 11 which has shown a strong association with smoking initiation.[4] Furthermore, prior studies have showed a strong connection between the BDNF locus and body mass index (BMI).[5 6]

The protein BDNF belongs to a neurotrophin family and plays a critical role in regulating neuron protective mechanisms such as survival, function, development, and plasticity of the cell.[7] The distribution in key regions of the central nervous system regulating mood and behaviour has resulted in extensive studies related to several psychiatric disorders. Associations with substance related disorders, eating disorders, and schizophrenia have been confirmed in a meta-analysis of case–control studies[8] and animal studies,[9] and have suggested an important role in drug addiction by acting upon the reward system of the brain. A hypothesis that BDNF might be associated with nicotine addiction has further been suggested and differences in plasma concentrations of BDNF have been seen in smokers compared to non-smokers, indicating that chronic smoking leads to a downregulation of the protein.[10]

Given the association with smoking behaviour and substance related disorders, the purpose of this study was to test the hypothesis that genetic variations in the BDNF locus alter the risk of smoking related complications among smokers in the Malmö Diet and Cancer Study (MDCS), a population based prospective cohort study.

METHODS

Study population

The population based MDCS included 12 121 men born from 1923 to 1945 and 18 326 women born from 1923 to 1950 from Malmö, Sweden. Participants attended baseline examinations between 1991 and 1996.

To cite: Halldén S, Sjögren M, Hedblad B, *et al*. *Heart* 2013;**99**:949–953.

All participants provided written informed consent and study protocols were approved by the ethical committee at Lund University, Lund, Sweden.

We selected the BDNF rs4923461, the lead single nucleotide polymorphism (SNP) in previous GWAS for BMI,[6] as it is in almost complete linkage disequilibrium (http://www.HapMap.org) with the SNPs shown to be smoking associated in recent GWAS.[4] Genotyping of rs4923461 was successfully performed in 27 508 out of 28 564 subjects (success rate 96.3%). Complete data for age, sex, BMI, and smoking status were available in 25 789 individuals. This cohort was used in the analysis of genotype and ever smoker status (table 1).

Complete data for age, sex, BMI, and smoking status as well as quantity (cigarettes per day (CPD))[11] were available for 25 071 individuals. This cohort was used in the analysis of genotype and prediction of smoking related complications (table 1).

Socioeconomic status (SES) was defined by the highest level of education and the participants were divided into three groups. The low SES group had not completed elementary school, corresponding to a maximum of 6–8 years of education; the middle SES group had 9–12 years of education; and the high SES group reported a university degree or studies at least 1 year after GCE (General Certificate of Education).[12]

Smoking status

Study participants (n=25 789) were classified as current smokers (n=7225) if they reported smoking regularly (n=6057) or sometimes within the past year (n=1168), and as non-smokers (n=18 564) if they reported never having smoked (n=9773) or having quit smoking at least 1 year before interview (n=8791). Further on, a combined group of ever smokers was formed, consisting of the current and previous smokers (n=16 016).

After further adjustment for smoking quantity, complete data on current smokers were registered for 6507 current smokers, of whom 5647 subjects reported as smoking regularly and 860 as smoking occasionally. We have no data on smoking status after the baseline exam.

DNA extraction and genotyping

Genotyping of BDNF rs4923461 was performed using TaqMan (Applied Biosystems) with primers and conditions according to the manufacture's recommendation.

Clinical end points

Four end points were examined: total mortality, cardiovascular mortality, cancer mortality, and first incidence of cardiovascular disease (CVD).

Table 1 Distribution of genotypes in the study population

	A/A	A/G	G/G
All* (n=25 789)	16 414 (63.5%)	8375 (32.5%)	1000 (4.0%)
Ever smokers (n=16 016)	10 256 (64.0%)	5171 (32.5%)	589 (3.5%)
Non-smokers (n=9773)	6158 (63.0%)	3204 (33.0%)	411 (4.0%)
All† (n=25 071)	15 959 (63.5%)	8146 (32.5%)	966 (4.0%)
Current smokers (n=6507)	4192 (64.5%)	2074 (32.0%)	241 (3.5%)
Non-smokers (n=18 564)	11 767 (63.5%)	6072 (32.5%)	725 (4.0%)

*In cohort with complete data for sex, age, rs4923461, BMI, and smoking status.
†In cohort with complete data for sex, age, rs4923461, BMI, smoking status, and smoking quantity.
BMI, body mass index.

Information on total mortality, cardiovascular mortality, and cancer mortality during follow-up was retrieved by linking the 10-digit civil registration number with the Swedish National Cause of Death Register (SNCDR). Mortality was classified as attributable to cardiovascular causes when the main International Classification of Diseases (ICD) code was 390–459 (ICD 9) or I00-I99 (ICD 10) and attributable to cancer when the ICD code was 140–239 (ICD 9) or C00-C99 (ICD 10) on the cause of death certificate.

CVD was defined as fatal or non-fatal myocardial infarction (MI) or stroke or death due to ischaemic heart disease from the Swedish Hospital Discharge Register or SNCDR. MI was defined as codes 410 (ICD9) or I21 (ICD10), death due to ischaemic heart disease as codes 412 and 414 (ICD9) or I22-I23 and I25 (ICD10), and stroke as codes 430, 431, 434, and 436 (ICD9) or I60-I61, I63, and I64 according to ICD10.

Follow-up extended until 1 January 2007. Mean±SD follow-up in analysis of mortality and analysis of first incident CVD was 12±3 years.

Statistical analysis

SPSS V.19.0 (IBM Corp) was used for all calculations.

Continuous variables are reported as means ±SD and dichotomous variables as numbers (%).

Cross-sectional relationships between genotype and smoking status were evaluated with crude and multivariate adjusted logistic regression. Relationships between genotype and BMI were tested with crude and multivariate adjusted linear regression models.

We calculated crude and multivariate adjusted hazard ratios (HR) and 95% confidence intervals (95% CI) for genotype in relation to the incidence of the four different end points during follow-up using the Cox proportional hazards model.

A two-sided p value of <0.05 was considered statistically significant.

A post hoc power analysis showed that the power among current smokers to detect the observed BDNF genetic effect on total mortality was 82%.

RESULTS

Baseline characteristics

Approximately 60% of the MDCS participants were women. The mean age was 57±8 years, the mean BMI was 26±4 kg/m², and the mean number of CPD for current smokers was 13±7. The participating men had a mean age of 59±7 years, mean BMI of 26±4 kg/m², and a mean cigarette intake of 16±9 CPD. The genotype distribution in the population did not deviate from the Hardy–Weinberg equilibrium (p=0.09).

BDNF rs4923461 polymorphism related to ever having smoked and BMI

In an additive model, the major allele (A), which previously has been associated with smoking and with high BMI, was significantly associated with increased odds of ever having smoked (odds ratio (OR) 1.050, 95% CI 1.004 to 1.098; p=0.032). After adjustment for age and sex, the association was still significant (OR 1.050, 95% CI 1.003 to 1.099; p=0.035) as well as after additional adjustment for BMI (OR 1.052, 95% CI 1.006 to 1.101; p=0.028). Also, the major allele was significantly associated with BMI in crude analysis (β=0.145±0.044; p=0.001) as well as after age and sex adjustment (β=0.155 ± 0.044; p<0.001).

Table 2 Multivariable adjusted HRs (95% CI) per genotype in current smokers

Genotype	G/G	A/G	A/A	p trend
Total mortality				
Model 1	1.0 (ref)	1.091 (0.776 to 1.534)	1.224 (0.894 to 1.732)	0.029
Model 2	1.0 (ref)	1.091 (0.776 to1.534)	1.224 (0.893 to 1.732)	0.029
Model 3	1.0 (ref)	1.101 (0.783 to 1.548)	1.236 (0.887 to 1.720)	0.047
CVD mortality				
Model 1	1.0 (ref)	1.061 (0.582 to 1.934)	1.372 (0.768 to 2.451)	0.028
Model 2	1.0 (ref)	1.057 (0.580 to 1.928)	1.352 (0.756 to 2.416)	0.036
Model 3	1.0 (ref)	1.062 (0.582 to 1.936)	1.344 (0.752 to 2.402)	0.047
Cancer mortality				
Model 1	1.0 (ref)	1.788 (0.970to 3.295)	1.880 (1.030 to 3.249)*	0.108
Model 2	1.0 (ref)	1.791 (0.972 to 3.302)	1.888 (1.035 to 3.444)*	0.103
Model 3	1.0 (ref)	1.815 (0.984 to 3.345)	1.872 (1.026 to 3.415)*	0.147
First incident CVD				
Model 1	1.0 (ref)	0.940 (0.647 to 1.366)	1.063 (0.740 to 1.527)	0.171
Model 2	1.0 (ref)	0.939 (0.644 to 1.360)	1.045 (0.727 to 1.501)	0.233
Model 3	1.0 (ref)	0.936 (0.644 to 1.361)	1.034 (0.720 to 1.486)	0.291

*p<0.050.
Adjustments: Model 1: age and sex. Model 2: age, sex, and BMI. Model 3: age, sex, BMI, and cigarettes per day (CPD).
BMI, body mass index; CVD, cardiovascular disease.

BDNF rs4923461 polymorphism and smoking related complications in current smokers

Total mortality

During follow-up 1049 (16.1%) deaths occurred among smokers.

Additive models, with the major allele (A) coded and adjusted for age and sex (model 1), showed that each copy of the smoking associated allele was associated with significantly increased risk of death (HR 1.131, 95% CI 1.013 to 1.263). After further adjustment including age, sex, and BMI (model 2), the increased risk of death remained significant (HR 1.131, 95% CI 1.013 to 1.263), as well as after additional adjustment for smoking quantity (model 3) (HR 1.118, 95% CI 1.002 to 1.249). Multivariable adjusted HRs per genotype with the GG genotype as the reference and p values for trend are shown in table 2.

Death from cardiovascular disease

Among current smokers, 346 (5.3%) cardiovascular deaths occurred. In all three additive models (models 1–3), higher HRs were significantly related to the presence of the A-allele: HR 1.247, 95% CI 1.024 to 1.518; HR 1.235, 95% CI 1.014 to 1.503; and HR 1.225, 95% CI 1.006 to 1.492, respectively. Multivariable adjusted HRs per genotype and p values for trend are shown in table 2.

Death from cancer

A total of 492 (7.6%) deaths from cancer were reported during follow-up. No significant association between the A-allele and risk of cancer mortality was observed in additive models (models 1–3): HR 1.142, 95% CI 0.971 to 1.343; HR 1.145, 95% CI 0.973 to 1.346; and HR 1.127, 95% CI 0.959 to 1.326, respectively. A significant association indicating an increased risk of cancer mortality among the major homozygotes (A/A) compared to the minor homozygotes (G/G) was seen in models 1–3. Multivariable adjusted HRs per genotype and p values for trend are shown in table 2.

First incidence in CVD

When analysing the BDNF polymorphism in relation to first incident CVD event, a total of 6321 cases with complete data

Table 3 Multivariable adjusted HRs (95% CI) per genotype in non-smokers

Genotype	G/G	A/G	A/A	p trend
Total mortality				
Model 1	1.0 (ref)	1.090 (0.853 to 1.393)	1.024 (0.806 to 1.302)	0.414
Model 2	1.0 (ref)	1.080 (0.845 to 1.381)	1.015 (0.799 to 1.290)	0.386
CVD mortality				
Model 1	1.0 (ref)	0.861 (0.571 to 1.299)	0.933 (0.627 to 1.388)	0.646
Model 2	1.0 (ref)	0.851 (0.564 to 1.283)	0.923 (0.621to 1.373)	0.655
Cancer mortality				
Model 1	1.0 (ref)	1.238 (0.860 to 1.780)	1.028 (0.719 to 1.469)	0.064
Model 2	1.0 (ref)	1.228 (0.854 to 1.767)	1.019 (0.713 to 1.457)	0.059
First incident CVD				
Model 1	1.0 (ref)	0.883 (0.688 to 1.134)	0.902 (0.708 to 1.149)	0.850
Model 2	1.0 (ref)	0.872 (0.679 to 1.120)	0.890 (0.698 to 1.133)	0.782

Adjustments: Model 1: age and sex. Model 2: age, sex, and BMI.
BMI, body mass index; CVD, cardiovascular disease.

were registered and included 802 (12.7%) first CVD events. No significant association of BDNF rs4923461 was present in the additive models (models 1–3): HR 1.092, 95% CI 0.963 to 1.238; HR 1.079, 95% CI 0.952 to 1.224; and HR 1.070, 95% CI 0.944 to 1.213, respectively. Multivariable adjusted HRs per genotype and p values for trend are shown in table 2.

Additional adjustments

The exposure for cigarette smoke was in addition to CPD calculated as 'pack years'. The results were unchanged, demonstrated here with model 3 for the end points of total mortality and CVD mortality: HR 1.117, 95% CI 1.000 to 1.248, p=0.049; and HR 1.229, 95% CI 1.009 to 1.498, p=0.040. The variable SES had little impact on the results, which is why it is not included in the tables.

BDNF rs4923461 polymorphism and smoking related complications in non-smokers

We reproduced Cox regression analyses for subjects reporting to be never or previous smokers, excluding the variable CPD. In this group, 1871 (10.1%) deaths occurred, 603 (3.2%) cardiovascular deaths, and 877 (4.7%) cancer deaths. A total of 1606 (8.9%) first CVD events were reported.

There was no association between rs4923461 and any of the four end points in non-smokers. A borderline significant inverse association between the A-allele and cancer mortality (HR 0.898, 95% CI 0.802 to 1.006) was observed (table 3).

To distinguish the group of previous smokers (n=8791) from never smokers, this group was analysed separately, but no associations were seen in this subgroup either (data not shown).

DISCUSSION

We show here that a genetic variation of the previously smoking and high BMI associated BDNF locus predicts an increased risk of dying among smokers, especially the risk of death from CVDs. Our research also confirms the recent associations with regular smoking[4] and increasing BMI.[6] Thus, genetic variation of the BDNF locus not only increases the likelihood of being a smoker, but also confers an increased risk of death among smokers. Assuming that the genetic BDNF association with mortality is attributable to a lesser likelihood of smoking cessation during long term follow-up, our results may have clinical implications warranting more intense smoking cessation interventions in subjects at such increased genetic risk.

The associations between the genetic BDNF variation and the outcomes of total and cardiovascular mortality are independent of traditional risk factors such as age, sex, BMI, and smoking quantity. This indicates that, despite the association between the BDNF locus and BMI, the association with mortality in smokers is not mediated by obesity and its complications, such as diabetes mellitus, but rather with prolonged exposure to smoking due to the lesser likelihood of smoking cessation during follow-up. This interpretation is further supported by a total lack of association between the BDNF locus and mortality among never or previous smokers. However, we cannot exclude other causes of the increased mortality rates among smokers carrying the risk allele. For example, genetic BDNF variation has also been linked to other substance abuse disorders and psychiatric diseases[8] with notable lower life expectancies. Unfortunately, we do not have records of substance abuse or psychiatric disorders and therefore these subjects could not be excluded from analyses.

We could not predict new events of CVD in our cohort. One theory is that smokers have had their first event before baseline exams and therefore were excluded in our analysis of incident CVD. Another explanation, as discussed above, is that the first event more often had a severe outcome in patients who use tobacco.

While the ample size of our cohort provides us with adequate statistical power to detect associations, the wide CIs demand a need for caution when drawing conclusions. In order to improve precision in our results, it is possible that a genetic smoking propensity score based on GWAS identified SNPs may have greater predictive accuracy.[13] Moreover, chronic obstructive pulmonary disease is a frequent consequence of both smoking and CVD which we did not take into account.

Tobacco consumption conveys negative consequences both for the health of the individual as well as increasing national health care costs. Although there has been substantial progress in preventing the spread of tobacco use worldwide, it is still the single most preventable cause of death in the USA.[14] Nicotine replacement therapies have resulted in low quitting rates, pointing out the need for new methods of intervention.[15] The progress of applicable techniques for genetic decoding continuously introduce new possibilities for identifying individuals with a higher risk of being smokers.

In conclusion, genetic BDNF variation predicts the risk of death in smokers. Our results suggest that future treatment may involve the molecular consequences of this genetic variation or guide the need for smoking cessation interventions by determination of the BDNF genotype in smokers.

Contributors Conception and design, or analysis and interpretation of data: SH, MS, BH, GE, M. Drafting the article or revising it critically for important intellectual content: SH, MS, BH, GE, KN, MH, BW, TH, M. Final approval of the version to be published: SH, M.

Funding This work was supported by grants from the European Research Council (StG-282255), Swedish Medical Research Council, the Swedish Heart and Lung Foundation, the Medical Faculty of Lund University, Malmö University Hospital, the Albert Påhlsson Research Foundation, the Crafoord Foundation, the Ernhold Lundströms Research Foundation, the Region Skåne, the Hulda and Conrad Mossfelt Foundation, the King Gustaf V and Queen Victoria Foundation, the Lennart Hanssons Memorial Fund, the Wallenberg Foundation and the Care North Consortium.

Patient consent Obtained.

Competing interests None.

Ethics approval Ethical committee at Lund University, Lund, Sweden.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Atlanta: U.S. Department of Health and Human Services CfDCaP, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2004 [accessed 2012 Jan 10]. The Health Consequences of Smoking: A Report of the Surgeon General 2012.
2 Johnson EO, Novak SP. Onset and persistence of daily smoking: the interplay of socioeconomic status, gender, and psychiatric disorders. *Drug Alcohol Depend* 2009;104(Suppl 1):S50–7.
3 Lang UE, Sander T, Lohoff FW, *et al*. Association of the met66 allele of brain-derived neurotrophic factor (BDNF) with smoking. *Psychopharmacology (Berl)* 2007;190:433–9.
4 Consortium TaG. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010;42:441–7.
5 Gunstad J, Schofield P, Paul RH, *et al*. BDNF Val66Met polymorphism is associated with body mass index in healthy adults. *Neuropsychobiology* 2006;53:153–6.
6 Thorleifsson G, Walters GB, Gudbjartsson DF, *et al*. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat Genet* 2009;41:18–24.
7 Hashimoto K, Koizumi H, Nakazato M, *et al*. Role of brain-derived neurotrophic factor in eating disorders: recent findings and its pathophysiological implications. *Prog Neuropsychopharmacol Biol Psychiatry* 2005;29:499–504.
8 Gratacos M, Gonzalez JR, Mercader JM, *et al*. Brain-derived neurotrophic factor Val66Met and psychiatric disorders: meta-analysis of case-control studies confirm

association to substance-related disorders, eating disorders, and schizophrenia. *Biol Psychiatry* 2007;61:911–22.
9 Grimm JW, Lu L, Hayashi T, *et al*. Time-dependent increases in brain-derived neurotrophic factor protein levels within the mesolimbic dopamine system after withdrawal from cocaine: implications for incubation of cocaine craving. *J Neurosci* 2003;23:742–7.
10 Kim TS, Kim DJ, Lee H, *et al*. Increased plasma brain-derived neurotrophic factor levels in chronic smokers following unaided smoking cessation. *Neurosci Lett* 2007;423:53–7.
11 Janzon E, Hedblad B. Swedish snuff and incidence of cardiovascular disease. A population-based cohort study. *BMC Cardiovasc Disord* 2009;9:21.
12 Rosvall M, Ostergren PO, Hedblad B, *et al*. Occupational status, educational level, and the prevalence of carotid atherosclerosis in a general population sample of middle-aged Swedish men and women: results from the Malmo Diet and Cancer Study. *Am J Epidemiol* 2000;152:334–46.
13 Li S, Zhao JH, Luan J, *et al*. Cumulative effects and predictive value of common obesity-susceptibility variants identified by genome-wide association studies. *Am J Clin Nutr* 2010;91:184–90.
14 Centers for Disease Control and Prevention. Annual smoking-attributable mortality, years of potential life lost, and economic costs—United States, 1995–1999. *MMWR Morb Mortal Wkly Rep*; 2002;51:300–3.
15 Zbikowski SM, Magnusson B, Pockey JR, *et al*. A review of smoking cessation interventions for smokers aged 50 and older. *Maturitas* 2012; 71:131–41

Paper II

Original Article

Journal of INTERNAL MEDICINE

doi: 10.1111/joim.12454

Gene variance in the nicotinic receptor cluster (*CHRNA5-CHRNA3-CHRNB4*) predicts death from cardiopulmonary disease and cancer in smokers

S. Halldén[1,2], M. Sjögren[1], B. Hedblad[1], G. Engström[1], V. Hamrefors[1,3], J. Manjer[1,4] & O. Melander[1,2]

From the [1]Department of Clinical Sciences, Lund University, Malmö, Sweden; [2]Department of Internal Medicine, Skåne University Hospital Malmö, Malmö, Sweden; [3]Department of Medical Imaging and Physiology, Skåne University Hospital, Malmö, Sweden; and [4]Department of Reconstructive Surgery, Skåne University Hospital Malmö, Malmö, Sweden

Abstract. Halldén S, Sjögren M, Hedblad B, Engström G, Hamrefors V, Manjer J, Melander O (Lund University; Skåne University Hospital Malmö, Malmö, Sweden; Skåne University Hospital; and Skåne University Hospital Malmö, Malmö, Sweden). Gene variance in the nicotinic receptor cluster (*CHRNA5-CHRNA3-CHRNB4*) predicts death from cardiopulmonary disease and cancer in smokers. *J Intern Med* 2016; **279**: 388–398.

Background. Genetic variation in the cluster on chromosome 15, encoding the nicotinic acetylcholine receptor subunits (*CHRNA5-CHRNA3-CHRNB4*), has shown strong associations with tobacco consumption and an additional risk increase in smoking-related diseases such as chronic obstructive pulmonary disease (COPD), peripheral artery disease and lung cancer.

Objectives. To test whether rs1051730 (C/T), a tag for multiple variants in the CHRNA5-CHRNA3-CHRNB3 cluster, is associated with a change in risk of smoking-related mortality and morbidity in the Malmö Diet and Cancer study, a population-based prospective cohort study.

Methods. At baseline participants were classified as current ($n = 6951$), previous ($n = 8426$) or never ($n = 9417$) smokers. Cox-proportional hazards models were used to determine the correlation between rs1051730 and incidence of first COPD, tobacco-related cancer, other cancer and cardiovascular disease (CVD), and total mortality due to these causes, during approximately 14 years of follow-up.

Results. Amongst current smokers there were 480 first incident COPD events, 852 tobacco-related cancers, 810 other cancers and 1022 CVD events. A total of 1508 deaths occurred, including 500 due to CVD, 102 due to respiratory diseases and 677 due to cancer. In adjusted additive models, an increasing number of T alleles were associated with a gradual increase in total mortality, incident COPD and tobacco-related cancer, even after adjustment for smoking quantity. No significant associations were observed amongst never smokers.

Conclusion. Our data suggest that gene variance in the *CHRNA5-CHRNA3-CHRNB4* cluster is associated with an increased risk of death, incidence of COPD and tobacco-related cancer in smokers. These findings indicate an individual susceptibility to tobacco use and its complications; this may be important when targeting and designing smoking cessation therapies.

Keywords: *CHRNA*, COPD, epidemiology, smoking genetics, tobacco-related cancer.

Introduction

More than 1 billion people around the world are smokers [1]. Negative health consequences such as cancer, heart disease, stroke and respiratory diseases are well-known complications, and cigarette smoking is responsible for about 5 million deaths annually (Data from the World Health Organization). Recent genomewide association studies have shown convincing associations between a number of genetic variations and both nicotine dependence (ND) and smoking behaviour [2–6]. The synonymous single nucleotide polymorphism (SNP) rs1051730 on chromosome 15q25, in the gene for the nicotinic acetylcholine receptor (nAChr) subunit *CHRNA3,* showed the strongest association. For this SNP,

which shows a risk allele frequency of approximately 38% in European populations, each copy of the risk allele corresponded to an increase in smoking quantity of 1 cigarette per day (CPD) [3]. It is interesting that the cluster of genes on chromosome 15q25, encoding the nACHr subunits *CHRNA5-CHRNA3-CHRNB4*, has been shown to be associated not only with smoking quantity and ND but also with smoking-related diseases such as chronic obstructive pulmonary disease (COPD), lung cancer, peripheral artery disease and bladder cancer [7]. Not all smokers develop these smoking-related diseases, a fact that might indicate that genetic differences also contribute to individual susceptibility.

This gene cluster of nACHrs is also known to be an area of high correlation and, according to the international HapMap project (http://hapmap.ncbi.nlm.nih.gov), the rs1051730 is in almost perfect correlation with rs16969968 (*CHRNA5*) in European populations, and therefore, these variants are considered to be essentially interchangeable. The rs16969968 is a coding variant, and rs1051730 should be considered as a surrogate marker [8].

The purpose of this study was to determine whether genetic variations in the 15q25 locus affect the risk of smoking-related complications amongst smokers in the Malmö Diet and Cancer study.

Methods

Study population

The prospective population-based Malmö Diet and Cancer study included a total of 18 326 women born between 1923 and 1950 and 12 121 men born between 1923 and 1945 in Malmö, Sweden [9]. Participants were recruited from 1991 to 1996. At the baseline examination, anthropometric variables and blood pressure were measured and blood samples were collected and stored for later analysis. Additionally, subjects were asked to complete a self-administered questionnaire of health- and lifestyle-related factors including current and previous disease, medication, smoking and socioeconomic factors. All participants provided written, informed consent, and study protocols were approved by the ethics committee at Lund University, Lund, Sweden.

DNA extraction and genotyping

DNA was available for 28 564 subjects, and genotyping of rs1051730 was successfully performed for 26 471 subjects (success rate 92.7%). Genotyping was performed using TaqMan (Applied Biosystems, Foster City, CA, USA) with primers and conditions according to the manufacturer's recommendations.

Baseline variables

Blood pressure was measured once in the supine position, after a 5-min rest, using a mercury sphygmomanometer. Hypertension was defined as systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg, or ongoing therapy with antihypertensive medication. Use of antihypertensive or lipid-lowering treatment (LLT), history of diabetes and smoking status were assessed from the Malmö Diet and Cancer study baseline questionnaire. Information about age, sex, rs1051730, body mass index (BMI), hypertension, previous diabetes diagnosis and LLT was available for all subjects included in the analyses (see Fig. 1).

Study participants who reported smoking daily or sometimes within the past year were classified as current smokers. Previous smokers reported smoking cessation at least 1 year before baseline and never smokers reported never having smoked. Current and previous smokers were included in the subgroup of ever smokers.

Data regarding CPD were available for current smokers, and additionally adjusted for as a continuous variable.

Assessment of end-points

The following end-points were examined: (i) incident COPD, (ii) incident tobacco-related cancer, (iii) incident non-tobacco-related cancer, (iv) incident cardiovascular disease (CVD), (v) total mortality, (vi) CVD mortality, (vii) cancer mortality and (viii) respiratory disease mortality.

Information about mortality end-points during follow-up was retrieved through linkage of the 10-digit civil registration number with the Swedish National Cause of Death Register (SNCDR). The SNCDR has previously been validated [10, 11]. Mortality was classified as attributable to cardiovascular causes for main International Classification of Diseases (ICD) ninth and 10th revision (ICD9 and ICD10, respectively) codes 390–459 (ICD9) or I00–I99 (ICD10) and was attributable to

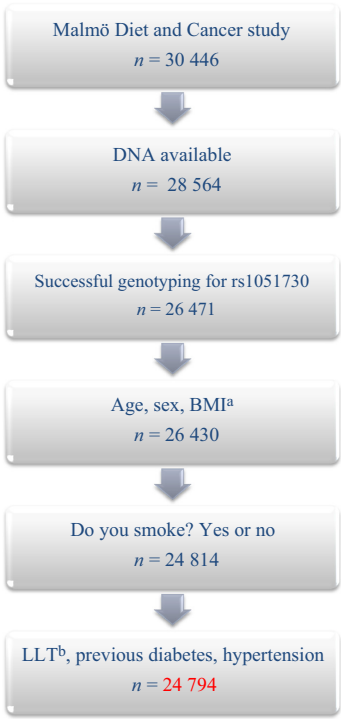


Fig. 1 *Selection of study subjects.* [a]*BMI, Body Mass Index;* [b]*LLT, Lipid-lowering therapy.*

cancer when the code was given as 140–239 (ICD9) or C00–C97 (ICD10) on the cause of death certificate.

CVD was defined as fatal or nonfatal myocardial infarction (MI), stroke or death due to ischaemic heart disease from the Swedish Hospital Discharge Register or SNCDR. MI was defined as codes 410 (ICD9) or I21 (ICD10), death due to ischaemic heart disease was defined as codes 412 and 414 (ICD9) or I22–I23 and I25 (ICD10) and stroke as codes 430, 431, 434 and 436 (ICD9) or I60–I61, I63 and I64 (ICD10). Respiratory disease mortality was defined as codes 460–519 (ICD9) or J00-99 (ICD10), and incident COPD as codes 490–496 (ICD9) or J40–44 (ICD10).

Tobacco-related cancers were defined by the International Agency for Research on Cancer (IARC) [12]: cancer of the oral cavity [ICD seventh revision (ICD7) code 140–144], oropharynx (ICD7 145, 147 and 148), nasopharynx (ICD7 146), oesophagus (ICD7 150), stomach (ICD7 151), colon (ICD7 153), rectum (ICD7 154), liver (ICD7 155 and 156), pancreas (ICD7 157), nose and sinuses (ICD7 160), larynx (ICD7 161), lung (ICD7 162 and 163), uterine cervix (ICD7 171), ovary (ICD7 175), kidney (ICD7 180) and lower urinary tract (ICD7 181) and myeloid leukaemia (ICD7 205).

Non-tobacco-related ('other') cancers included cancers of the breast (ICD7 170), prostate (ICD7 177), skin including malignant melanoma (ICD7 190–191) and nervous system (ICD7 193) and malignant lymphoma (ICD7 200–201).

Follow-up for the end-points incident COPD and CVD, and total, CVD, respiratory disease and cancer mortality, extended until 31 December 2009. Follow-up for the end-points incident tobacco-related and non-tobacco-related cancers extended until 31 December 2010.

Statistical analysis

Continuous variables are reported as means (SD) and dichotomous variables as numbers (%). The SNP rs1051730 was coded additively (CC=0, CT=1, TT=2) in all analyses. Cross-sectional relationships between genotype and smoking status were evaluated using logistic regression. Cox-proportional hazard models were used to calculate hazard ratios (HR) and 95% confidence intervals (CIs) for rs1051730 in relation to first event of each end-point during follow-up.

All *P*-values reported are two-sided. *P*-values were not adjusted for multiple tests. SPSS version 22.0 (IBM Corporation, New York, NY, USA) was used for all calculations.

Results

Baseline characteristics

Complete data were retrieved for 24 794 study participants (Fig. 1). Baseline data of the study participants are presented in Table 1. Smoking data and genotype distribution are presented in Table 2. There was no significant deviation from Hardy–Weinberg equilibrium in any of the groups studied ($P > 0.05$).

Table 1 *Baseline characteristics of the study population*

	Female	Male	All
Study participants, *n* (%)	15 094 (60.9)	9700 (39.1)	24 794 (100)
Age, years	57 (±8)	59 (±7)	58 (±8)
BMI, kg/m^2	25 (±4)	26 (±4)	26 (±4)
Hypertension[a], *n* (%)	8569 (56.8)	6655 (68.6)	15 224 (61.4)
Lipid-lowering therapy, *n* (%)	320 (2)	475 (5)	795 (3)
Previous diabetes, *n* (%)	480 (3.2)	558 (5.8)	1038 (4.2)
CPD[b]	13 (±7)	16 (±9)	14 (±8)

Data are presented as mean (±SE) unless otherwise stated.
[a]Antihypertensive treatment and/or systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg. [b]Data available for 6260 study participants. CPD, cigarettes per day.

Table 2 *Distribution of rs1051730 in the overall study population and stratified by smoking status*

Genotype	CC	CT	TT
All genotyped subjects *n* = 26 471 (%)	12 092 (45.7)	11 442 (43.2)	2937 (11.1)
Analysed subjects[a] *n* = 24 794 (%)	11 311 (45.6)	10 732 (43.3)	2751 (11.1)
Current smokers *n* = 6951 (%)	3109 (44.7)	3041 (43.8)	801 (11.5)
Previous smokers *n* = 8426 (%)	3916 (46.5)	3611 (42.9)	899 (10.7)
Never smokers *n* = 9417 (%)	4286 (45.5)	4080 (43.3)	1051 (11.2)

[a]With complete data for age, sex, rs1051730, body mass index, smoking status, hypertension and lipid-lowering treatment.

rs1051730 polymorphism and relation to smoking behaviour and risk factors

In additive models adjusted for age and sex, the risk allele (T) showed an association with current smoking (compared to nonsmokers, i.e. never and former smokers combined) [odds ratio (OR) 1.042, 95% CI 1.000–1.087; $P = 0.050$]. Conversely, the polymorphism was associated with a lower probability of being a former smoker (OR 0.961, 95% CI 0.923–1.000; $P = 0.048$). No associations were observed between the polymorphism and ever or never smoking status (OR, 95% CI: 0.994, 0.965–1.033; $P = 0.752$ and 1.002, 0.964–1.042; $P = 0.911$, respectively). Furthermore, within the group of ever smokers (i.e. never smokers excluded), the OR for current smoking compared with previous smoking per T allele was 1.057 (95% CI 1.007–1.109; $P = 0.024$).

The risk allele showed a strong linear association with smoking quantity (β = 1.137 CPD per allele; $P = 9 \times 10^{-15}$). There were no significant correlations between rs1051730 and BMI, diabetes, hypertension or LLT (data not shown).

CHRNA polymorphism and smoking-related complications in current, previous and never smokers

Multivariable-adjusted HRs and 95% CIs per genotype from categorical models are shown in Tables 3–5, where carriers of CT and TT are compared with CC carriers (defined as the reference group: HR 1.0). *P*-values for trend (from additive models), number of events, total cases and event rates per 1000 person-years for all end-points are also shown in Tables 3–5.

The HRs and 95% CIs presented in the following section are average values per T allele calculated in additive models. The mean follow-up times for all end-points are shown in Table S1.

Current smokers

Incident COPD

In model 1, adjusted for age and sex, each copy of the risk allele was associated with a significant increase in COPD incidence (HR 1.290, 95% CI 1.130–1.463). The increased risk remained significant in model 2, with further adjustments for

Table 3 *Multivariable-adjusted hazard ratios per genotype in current smokers*

Genotype			CC	CT	TT			
		Events/		HR				
End-points	Events (total cases[a,b])	1000 p-ys		95% CI[a]		*P*trend[a]	*P*trend[b]	*P*trend[c]
Incident COPD	480 (6931/6241[c])	4.97	1.0 (ref)	1.281 1.053–1.559*	1.657 1.266–2.169*	<0.001*	<0.001*	0.001*
Incident tobacco-related cancer	852 (6304/5645[c])	9.33	1.0 (ref)	1.139 0.985–1.317	1.385 1.126–1.705*	0.002*	0.002*	0.043*
Incident other cancer	810 (6676/6015[c])	8.48	1.0 (ref)	1.098 0.981–1.271	1.058 0.842–1.331	0.350	0.358	0.769
Incident CVD	1022 (6760/6084[c])	11.21	1.0 (ref)	0.956 0.840–1.089	0.958 0.780–1.177	0.534	0.683	0.618
Total mortality	1508 (6951/6260[c])	15.71	1.0 (ref)	1.036 0.930–1.155	1.258 1.076–1.472*	0.014*	0.009*	0.065
CVD mortality	500 (6951/6260[c])	5.37	1.0 (ref)	0.893 0.738–1.079	1.181 0.904–1.544	0.674	0.574	0.666
Respiratory disease mortality	102 (6946/6256[c])	1.06	1.0 (ref)	1.534 0.995–2.364	1.827 1.009–3.311*	0.022*	0.017*	0.061
Cancer mortality	677 (6946/6255[c])	7.05	1.0 (ref)	1.113 0.947–1.307	1.145 0.897–1.463	0.161	0.160	0.496

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; p-ys, patient-years; HR, hazard ratio; CI, confidence interval.
[a]Adjusted for age and sex. [b]Adjusted for age, sex, body mass index (BMI), hypertension, previous diabetes and lipid-lowering treatment. [c]Adjusted for age, sex, BMI, hypertension, previous DM, LLT and cigarettes per day (available cases 6260).
*$P < 0.05$.

BMI, diabetes, LLT and hypertension (HR 1.286, 95% CI 1.131–1.463), as well as in model 3 after adjustments for CPD (HR 1.249, 95% CI 1.091–1.429).

Incident tobacco-related cancer
Significant associations between the risk allele and tobacco-related cancer were seen in models 1, 2 and 3 (HR, 95% CI: 1.167, 1.058–1.287; 1.116, 1.057–1.286; and 1.114, 1.004–1.237, respectively).

Incident other cancers
No significant associations between the risk allele and other cancers were seen in model 1 (HR 1.050, 95% CI 0. 948–1.162), model 2 (HR 1.050, 95% CI 0. 948–1.162) or model 3 (HR 1.017, 95% CI 0.911–1.314).

Incident CVD
No significant relation between the risk allele and incident CVD was observed in model 1 (HR 0.971, 95% CI 0.886–1.065), model 2 (HR 0.981, 95% CI 0.895–1.076) or model 3 (HR 0.975, 95% CI 0.884–1.076).

Total mortality
In models 1 and 2, the risk allele was significantly associated with total mortality (HR, 95% CI: 1.098, 1.019–1.182 and 1.103, 1.025–1.188, respectively). In model 3, the association was nearly significant (HR 1.077, 95% CI 0.995–1.166).

Table 4 *Multivariable-adjusted hazard ratios per genotype in previous smokers*

Genotype			CC	CT	TT		
End-point	Events (total cases[a,b])	Event/ 1000 py-s		HR 95% CI[a]		*P*trend[a]	*P*trend[b]
Incident COPD	211 (8394)	1.73	1.0 (ref)	1.086 0.814–1.449	1.265 0.819–1.955	0.293	0.292
Incident tobacco-related cancer	693 (7919)	5.80	1.0 (ref)	1.101 0.939–1.290	1.241 0.975–1.581	0.063	0.062
Incident other cancer	1218 (8071)	10.36	1.0 (ref)	0.917 0.814–1.033	0.934 0.766–1.140	0.206	0.212
Incident CVD	1047 (8025)	9.33	1.0 (ref)	1.007 0.885–1.144	0.991 0.805–1.219	0.987	0.865
Total mortality	1425 (8426)	11.90	1.0 (ref)	1.165 1.044–1.301*	1.214 1.020–1.444*	0.004*	0.004*
CVD mortality	488 (8425)	4.18	1.0 (ref)	1.373 1.137–1.659*	1.281 0.945–1.736	0.006*	0.005*
Respiratory disease mortality	70 (8390)	0.59	1.0 (ref)	1.379 0.832–2.285	1.543 0.725–3.283	0.160	0.166
Cancer mortality	614 (8426)	5.13	1.0 (ref)	1.044 0.882–1.236	1.195 0.923–1.547	0.216	0.218

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; p-ys, patient-years; HR, hazard ratio; CI, confidence interval.
[a]Adjusted for age and sex. [b]Adjusted for age, sex, body mass index (BMI), hypertension, previous diabetes and lipid-lowering treatment.
*$P < 0.05$.

CVD mortality
No significant association between the risk allele and CVD mortality was seen in any model (HR, 95% CI: 1.028, 0.903–1.171; 1.038, 0.912–1.181 and 1.031, 0.898–1.182, in models 1–3, respectively).

Respiratory disease mortality
In models 1 and 2, the risk allele was significant correlated with respiratory disease mortality (HR, 95% CI: 1.384, 1.048–1.827 and 1.376, 1.043–1.815, respectively). After adjusting for CPD in model 3, the association was no longer significant (HR 1.318, 95% CI 0.987–1.359).

Cancer mortality
The associations between the risk allele and cancer mortality were not significant, with identical results in models 1 and 2 (HR 1.083, 95% CI 0.969–1.210) or in model 3 (HR 1.033, 95% CI 0.917–1.163).

Previous smokers

Incident disease
In contrast to current smokers, no significant relation was observed between the risk allele and incident COPD in model 1 (HR 1.113, 95% CI 0.911–1.359) or model 2 (HR 1.114, 95% CI 0.912–1.360). A nearly significant association was seen when analysing the association with incident tobacco-related cancer in models 1 and 2 (HR, 95% CI: 1.110, 0.994–1.239 and 1.111, 0.995–1.240) but not with incident other cancers (0.946, 0.869–1.031 and 0.947, 0.869–1.032) or CVD (0.999, 0.912–1.095 and 0.992, 0.905–1.087, respectively).

Table 5 *Multivariable-adjusted hazard ratios per genotype in never smokers*

Genotype			CC	CT	TT		
	Events	Events/		HR			
End-point	(total cases[a,b])	1000 p-ys		95% CI[a]		*P*trend[a]	*P*trend[b]
Incident COPD	79 (9314)	0.56	1.0 (ref)	1.269 0.801–2.013	0.731 0.306–1.744	0.990	0.987
Incident tobacco-related cancer	559 (8975)	4.27	1.0 (ref)	1.079 0.906–1.286	0.974 0.733–1.293	0.789	0.790
Incident other cancer	1233 (8906)	9.20	1.0 (ref)	0.983 0.874–1.106	0.860 0.708–1.044	0.200	0.213
Incident CVD	955 (9268)	7.18	1.0 (ref)	0.935 0.818–1.070	0.878 0.708–1.089	0.176	0.180
Total mortality	1143 (9416)	8.34	1.0 (ref)	0.965 0.854–1.091	0.909 0.746–1.107	0.327	0.309
CVD mortality	350 (9415)	2.60	1.0 (ref)	0.814 0.650–1.018	0.894 0.634–1.261	0.184	0.152
Respiratory disease mortality	33 (9310)	0.24	1.0 (ref)	0.792 0.385–1.632	0.572 0.206–2.390	0.460	0.452
Cancer mortality	520 (9406)	3.80	1.0 (ref)	1.033 0.863–1.237	0.811 0.595–1.106	0.416	0.414

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; p-ys, patient-years; HR, hazard ratio; CI, confidence interval.
[a]Adjusted for age and sex. [b]Adjusted for age, sex, body mass index (BMI), hypertension, previous diabetes and lipid-lowering treatment.

Mortality end-points

All-cause mortality events were analysed in previous smokers and, in line with the results from current smokers, significant associations between the risk allele and mortality end-points were found in both models 1 and 2 (HR, 95% CI: 1.122, 1.038–1.212 and 1.119, 1.036–1.208). Moreover, a significant association was seen for mortality caused by CVD in these two models, respectively (HR, 95% CI: 1.202, 1.055–1.369 and 1.205, 1.057–1.374.). Respiratory disease mortality was not significantly associated with the allele (HR, 95% CI: 1.277, 0.908–1.797 and 1.273, 0.904–1.792); similarly, there was no significant association with death from cancer (HR, 95% CI: 1.077, 0.957–1.212 and 1.077, 0.957–1.211, respectively).

Never smokers

Incident disease

No associations were observed in models 1 and 2 between the risk allele and incident COPD (HR, 95% CI: 0.998, 0.718–1.386 and 1.003, 0.722–1.393), tobacco-related cancer (HRs 1.017, 95% CIs 0.899–1.151 in both models), incident other cancer (HR, 95% CI: 0.947 0.871–1.029 and 0.948, 0.872–1.031) or incident CVD (HR, 95% CI: 0.936, 0.851–1.030 and 0.937, 0.852–1.031, respectively).

Mortality end-points

No significant correlations were observed in models 1 and 2 between the risk allele and all-cause mortality (HR, 95% CI: 0.957, 0.878–1.044 and 0.956, 95% CI 0.876–1.043), CVD mortality (0.956,

0.876–1.043 and 0.891, 0.760–1.045) or respiratory disease mortality (0.819, 0.483–1.390 and 0.816, 0.481–1.385). There was also no correlation between the SNP and cancer mortality (HR, 95% CI: 0.948, 0.833–1.079 and 0.947, 0.832–1.078).

Additional analyses

We could not confirm the association between the risk allele and bladder cancer [7]. Incident bladder cancer was analysed separately with 117 (1.28 event/1000 person-years) cases in the current smoking group ($n = 6304$) and 98 (1.20/1000 person-years) cases with CPD data. No risk increase was seen in models 1, 2 or 3 (HR per allele, 95% CI: 0.808, 0.625–1.044; 0.763, 0.584–0.997; and 0.768, 0.566–1.043, respectively). Lung cancer incidence was also analysed separately, and a significant risk increase was seen in current smokers with 277 events (3.03/1000 person-years) reported in 6304 participants in models 1 and 2 (HR, 95% CI: 1.289, 1.088–1.527 and 1.283, 1.083–1.519, respectively). After adjusting for CPD in model 3 ($n = 5641$) with 251 events (3.08/1000 person-years), the association was no longer significant (HR 1.176, 95% CI 0.982–1.408). Furthermore, as we excluded lung cancer diagnosis from the end-point tobacco-related cancers, the associations were no longer significant (data not shown).

The median age of death stratified by genotype was also analysed, focusing on potential differences between subjects with low- (CC) and high-risk genotypes (TT). In current smokers, the median age of death was 1.4 years lower in TT carriers compared to CC carriers, and in previous smokers the median age of death was 0.2 years lower amongst TT carriers compared to CC carriers. By contrast, in the never smoking group, the median age of death was 2.1 years higher in subjects with the TT genotype compared to those with the CC genotype.

Discussion

The novel finding of the present study is that genetic variance in the 15q25 locus predicts an increased risk of death amongst smokers. We also confirmed the associations between this variance and incident COPD [4, 13, 14], tobacco-related cancers [7, 15, 16], lung cancer [4, 7, 15, 17–21] and smoking quantity [2, 3, 7], indicating an exciting overlap of genetic influence on ND and smoking-related diseases. As mentioned above, this region of the nAChRs is characterized by high correlation and the results should be interpreted as an association with the cluster instead of the rs1051730.

The additional risk increase in all-cause mortality was observed in both current and previous smokers. To illustrate this from another perspective, the median age at death amongst current smokers was 1.4 years lower in subjects with the risk genotype (TT) compared to subjects with the CC genotype.

There was no association between the SNP and mortality amongst never smokers, despite only a slightly lower number of events in this subgroup. Furthermore, with regard to the specific causes of mortality, the SNP was significantly associated with increased respiratory disease mortality amongst current smokers. There was no such significant association amongst former smokers; however as the number of events was more than 30% lower in this subgroup, the lack of association between the SNP and respiratory disease mortality amongst former smokers may be due to a lack of power. Furthermore, the SNP was associated significantly with CVD mortality amongst former smokers, whereas there was no such significant association amongst current smokers, despite a similar number of events. It may be speculated that because current smoking is itself a strong risk factor for CVD mortality, genetic influences on the nicotine receptor may affect the risk of CVD to a lesser extent in this subgroup (i.e. the risk increase in CVD mortality caused by smoking is large regardless of genotype). However, whether the differences in cause-specific mortality in relation to the SNP amongst the subgroups of current and former smokers could be attributed to different pathophysiological implications of the SNP in different diseases could not be determined with certainty from the current analyses. In general, we acknowledge that the power for interpretation of cause-specific mortality in the different subgroups is likely to be limited.

It is likely that the increased risk of total mortality as a result of the SNP amongst current and former smokers could be attributed to multiple potential mechanisms. There may be an interaction between the inhaled substances and the receptor that in a later process results in pathophysiological changes causing disease. It may be hypothesized that genetic changes in inflammatory responses could

increase the risk of many smoking-related diseases, thus also suggesting a possible common cause of the modification of the consequences of smoking by genetic influences. Yet, whether this is a direct association, or only a proxy for the increased exposure to tobacco carcinogens, remains controversial as the risk allele is also related to smoking quantity.

The association with lung cancer has been investigated in many studies and, in line with our results, it has been argued that the increased CPD is not the sole explanation. Investigation of nicotinic receptor function and distribution may theoretically increase understanding of this clinically interesting relationship. The nAchRs are present in the central nervous system and in peripheral organs such as the lung, and nicotine addiction is mediated through nAChRs in the mesolimbic dopaminergic system. In bronchial cells, the receptors are involved in remodelling airway epithelium and in the regulation of inflammation and immunity. In theory, nicotine consequently acts as a suppressor of the immune response, and could affect the clearance of transformed cells and participate in the emergence of neoplastic lesions [22].

In the present study, although the mortality risk is higher in both current and former smokers carrying the risk allele, the data highlight the benefits of smoking cessation. In ex-smokers, the risk allele no longer confers an additional risk of COPD incidence or tobacco-related cancers.

As briefly mentioned above, whether or not the increased risk due to the polymorphism in 15q25 in smokers could be the sole consequence of increase in CPD, is worth considering. The strong correlation between genetic variants in *CHRNA5-CHRNA3-CHRNB4*, here represented by rs1051730, and CPD suggests that the minor allele (T) could be associated with reduced sensitivity to plasma nicotine levels, leading to increased tobacco consumption [23]. Smokers homozygous for the minor allele inhale more often than noncarriers and heterozygous smokers [7]. Keskitalo *et al.* [24] measured CPD and the serum levels of cotinine, a metabolite of nicotine, in 560 daily smokers from a Finnish population. Both cotinine levels and CPD were strongly association with rs1051730, with effect sizes of 0.30 and 0.13, respectively. Hence, the authors concluded that the nicotinic receptor polymorphism influences cotinine/nicotine levels, and appears to be involved in nicotine metabolism and/or regulation. The effect size for cotinine levels is greater than that for CPD, and the authors suggested that the nAChR polymorphism influences nicotine levels more than smoking quantity, or at least that cotinine is a better measure of nicotine intake. Moreover, Timofeeva *et al.* [20] investigated the effect of the polymorphism on smoking behaviour and lung cancer by measuring circulating cotinine levels in lung cancer patients within the The European Prospective Investigation into Cancer and Nutrition cohort. No association was seen with smoking behaviour but the association between increased cotinine levels and the minor allele of rs16969968 (*CHRNA5*) was confirmed. The authors concluded that the use of crude measures such as CPD could underestimate the established effects of the SNP on chromosome 15q25 on lung cancer risk mediated by smoking.

In addition to smoking, obesity is a well-known risk factor for many diseases [25]. Nicotine acts on the reward system of the brain and, because eating and smoking are behavioural attributes that at least in part are controlled by the same mechanisms [26], Thorgeirsson *et al.* [27] investigated whether SNPs associated with BMI also have an impact on smoking behaviour. In several cases, the studied variants that were correlated with elevated BMI, increased the propensity to smoke and/or increased smoking quantity. Although no association between BMI and rs1051730 was seen in our cohort, Thorgeirsson *et al.* demonstrated a significant correlation with lower BMI in smokers but not in never smokers. The authors suggested that the influence on BMI is probably through the effect of the polymorphism on smoking behaviour and consequently the increase in metabolic rate and appetite suppression attributable to nicotine.

A limitation of our study is that cotinine levels were not measured. In addition, the burden of nicotine after baseline examinations is not known. It could have been of interest to analyse individuals exposed to passive smoking. Moreover, because this genetic variance in the α5-α3-β4 nAChr gene cluster seems to increase the need for nicotine, there might be a residual confounding effect of under-reporting smoking quantity in carriers of the risk allele, compared to noncarriers. A further limitation is that our definition of cause-specific mortality is based on the main underlying cause of death as listed on the Swedish death certificate.

Autopsy rates are relatively low (27%) leading to some uncertainty about the specific cause of death. We did note a slightly higher proportion of cancer deaths than we intuitively expected. Finally, we acknowledge that the power for interpreting cause-specific mortality in the subgroups is likely to be limited.

Conclusion

In this large, prospective study we have shown that smoking and gene variance in the *CHRNA5-CHRNA3-CHRNB4* cluster correlate with increased cigarette intake, mortality and incident tobacco-related diseases. Tobacco use is a leading cause of morbidity and mortality worldwide and our results confirm the notion that genetic factors partly contribute to the development of nicotine addiction and its complications. Understanding the genetics of dependence could lead to optimization of targeted treatments and prevention of disability and death. In the future, further genetic variants and molecular pathways need to be identified, in order to develop a more individualized intervention therapy, and possibly also to help motivate smoking cessation in high-risk individuals.

Conflicts of interest statement

None declared.

Acknowledgements

Funding was received from the European Research Council (StG-282255), the Swedish Heart and Lung Foundation, the Swedish Research Council, the Novo Nordisk Foundation, Skåne University Hospital donation funds, the Medical Faculty of Lund University, governmental funding for clinical research within the National Health Services, the Albert Påhlsson Research Foundation, Region Skåne, the King Gustaf V and Queen Victoria Foundation, and the Marianne and Marcus Wallenberg Foundation.

References

1 Secretan B, Straif K, Baan R, *et al.* A review of human carcinogens–Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish. *Lancet Oncol* 2009; **10:** 1033–4.
2 Thorgeirsson TE, Gudbjartsson DF, Surakka I, *et al.* Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet* 2010; **42:** 448–53.
3 Consortium TaG. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010; **42:** 441–7.
4 Saccone NL, Culverhouse RC, Schwantes-An TH, *et al.* Multiple independent loci at chromosome 15q25.1 affect smoking quantity: a meta-analysis and comparison with lung cancer and COPD. *PLoS Genet* 2010; **6:** e1001053.
5 Bierut LJ, Madden PA, Breslau N, *et al.* Novel genes identified in a high-density genome wide association study for nicotine dependence. *Hum Mol Genet* 2007; **16:** 24–35.
6 Saccone NL, Wang JC, Breslau N, *et al.* The CHRNA5-CHRNA3-CHRNB4 nicotinic receptor subunit gene cluster affects risk for nicotine dependence in African-Americans and in European-Americans. *Cancer Res* 2009; **69:** 6848–56.
7 Kaur-Knudsen D, Bojesen SE, Tybjaerg-Hansen A, Nordestgaard BG. Nicotinic acetylcholine receptor polymorphism, smoking behavior, and tobacco-related cancer and lung and cardiovascular diseases: a cohort study. *J Clin Oncol* 2011; **29:** 2875–82.
8 Bierut LJ, Stitzel JA, Wang JC, *et al.* Variants in nicotinic receptors and risk for nicotine dependence. *Am J Psychiatry* 2008; **165:** 1163–71.
9 Berglund G, Elmstahl S, Janzon L, Larsson SA. The Malmo diet and cancer study. Design and feasibility. *J Intern Med* 1993; **233:** 45–51.
10 Inghammar M, Engstrom G, Lofdahl CG, Egesten A. Validation of a COPD diagnosis from the Swedish Inpatient Registry. *Scand J Public Health* 2012; **40:** 773–6.
11 Hammar N, Alfredsson L, Rosen M, *et al.* A national record linkage to study acute myocardial infarction incidence and case fatality in Sweden. *Int J Epidemiol* 2001; **30(Suppl 1):** S30–4.
12 Agudo A, Bonet C, Travier N, *et al.* Impact of cigarette smoking on cancer risk in the European prospective investigation into cancer and nutrition study. *J Clin Oncol* 2012; **30:** 4550–7.
13 Kaur-Knudsen D, Bojesen SE, Nordestgaard BG. CHRNA3 and CYP3A5*3 genotype, lung function and chronic obstructive pulmonary disease in the general population. *Pharmacogenet Genomics* 2014; **24:** 220–9.
14 Pillai SG, Ge D, Zhu G, *et al.* A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet* 2009; **5:** e1000421.
15 Lips EH, Gaborieau V, McKay JD, *et al.* Association between a 15q25 gene variant, smoking quantity and tobacco-related cancers among 17 000 individuals. *Int J Epidemiol* 2010; **39:** 563–77.
16 Russo P, Cardinale A, Margaritora S, Cesario A. Nicotinic receptor and tobacco-related cancer. *Life Sci* 2012; **91:** 1087–92.
17 Amos CI, Wu X, Broderick P, *et al.* Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* 2008; **40:** 616–22.
18 Hung RJ, McKay JD, Gaborieau V, *et al.* A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 2008; **452:** 633–7.
19 Rafnar T, Sulem P, Besenbacher S, *et al.* Genome-wide significant association between a sequence variant at 15q15.2 and lung cancer risk. *Cancer Res* 2011; **71:** 1356–61.

20 Timofeeva MN, McKay JD, Smith GD, *et al.* Genetic polymorphisms in 15q25 and 19q13 loci, cotinine levels, and risk of lung cancer in EPIC. *Cancer Epidemiol Biomarkers Prev* 2011; **20:** 2250–61.
21 Spitz MR, Amos CI, Dong Q, Lin J, Wu X. The CHRNA5-A3 region on chromosome 15q24-25.1 is a risk factor both for nicotine dependence and for lung cancer. *J Natl Cancer Inst* 2008; **100:** 1552–6.
22 Tournier JM, Birembaut P. Nicotinic acetylcholine receptors and predisposition to lung cancer. *Curr Opin Oncol* 2011; **23:** 83–7.
23 Wang JC, Cruchaga C, Saccone NL, *et al.* Risk for nicotine dependence and lung cancer is conferred by mRNA expression levels and amino acid change in CHRNA5. *Hum Mol Genet* 2009; **18:** 3125–35.
24 Keskitalo K, Broms U, Heliovaara M, *et al.* Association of serum cotinine level with a cluster of three nicotinic acetylcholine receptor genes (CHRNA3/CHRNA5/CHRNB4) on chromosome 15. *Hum Mol Genet* 2009; **18:** 4007–12.
25 Haslam DW, James WP. Obesity. *Lancet* 2005; **366:** 1197–209.
26 Volkow ND, Wang GJ, Fowler JS, Telang F. Overlapping neuronal circuits in addiction and obesity: evidence of systems pathology. *Philos Trans R Soc Lond B Biol Sci* 2008; **363:** 3191–200.
27 Thorgeirsson TE, Gudbjartsson DF, Sulem P, *et al.* A common biological basis of obesity and nicotine addiction. *Transl Psychiatry* 2013; **3:** e308.

Correspondence: Sara Halldén, Clinical Research Center, Entrance 72, Building 91, Floor 12, Malmö University Hospital, SE 205 02 Malmö, Sweden.
(fax:+46 40 39 12 22; e-mail: sara.hallden@med.lu.se).

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Mean follow-up time (years).

Sara Halldén studied medicine at Lund University and graduated in 2008.
She started with research in 2011 during the end of her internship at Skåne University Hospital in Malmö.
Currently she is working as a resident at the department of internal medicine in Malmö.

Printed by Media-Tryck, Lund 2017 NORDIC ECOLABEL. 3041 0903

LUND UNIVERSITY
Faculty of Medicine

Department of Clinical Sciences in Malmö

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2017:103
ISBN 978-91-7619-483-6
ISSN 1652-8220