university of groningen

University Medical Center Groningen

University of Groningen

Chronic mucus hypersecretion and airway wall structure

Dijkstra, Akke

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

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

Publication date:
2014

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Dijkstra, A. (2014). *Chronic mucus hypersecretion and airway wall structure: Genes and environment*. [Thesis fully internal (DIV), University of Groningen]. s.n.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 13-02-2023

Chronic mucus hypersecretion and airway wall structure

Genes and environment

Akkelies Dijkstra

On the cover: a bronze statue by teja van hoften entitled 'lungs'.
Human structures show a strong similarity with plant-like structures. This statue is made according to the 'cire perdue' method, a process by which a duplicate metal sculpture is cast from an original wax model. But this time the original sculpture is not only made up of wax, but also of plants which can be seen in the grain structures in the surface of the bronze. The use of plants is well chosen as we depend on plants for our oxygen demand. For more information about her work, see www.tejavanhoften.nl

© A.E. Dijkstra, 2014
No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form by any means, without permission of the author, or, when appropriate, of the Publisher of the publication or illustration material.

The research in this thesis was financially supported by the European Union (EU FP7 grant 201379), Stichting Astma Bestrijding and the Dutch Lung foundation

The publication of this thesis was financially supported by the University of Groningen, GUIDE, University Medical Center Groningen (UMCG), Stichting Astma Bestrijding, Boehringer Ingelheim BV, GlaxoSmithKline BV, Novartis Pharma BV and Teva Nederland.

ISBN 978-90-6464-782-6

Design and lay-out: Ubel Smid Vormgeving, Roden, the Netherlands.
Printing: GVO drukkers & vormgevers BV, Ponsen & Looijen, Ede, the Netherlands.

Chronic mucus hypersecretion and airway wall structure

Genes and environment

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. E. Sterken
en volgens besluit van het College van Promoties.

De openbare verdediging zal plaatsvinden op

dinsdag 17 juni 2014 om 11.00 uur

door

Akke Elisabeth Dijkstra

geboren op 1 februari 1956
te Assen

Promotores
Prof. dr. D.S. Postma
Prof. dr. H.J.M. Groen
Prof. dr. H.M. Boezen

Co-promotor
Dr. J.M. Vonk

Beoordelingscommissie
Prof. dr. G.G. Brussselle
Prof. dr. J. Vestbo
Prof. dr. H.A.M. Kerstjens

Paranimfen
Marga Lunshof
Maartje Nieuwenhuis

Contents

Chapter 1	General Introduction	9
Chapter 2	Risk factors for chronic mucus hypersecretion in individuals with and without COPD	23
Chapter 3	Genetic susceptibility to chronic mucus hypersecretion; a genome wide association study	51
Chapter 4	Dissecting genetic risk factors for chronic mucus hypersecretion in heavy smokers with and without COPD	83
Chapter 5	Low-dose CT measurements of airway dimensions and emphysema associated with airflow limitation in heavy smokers	107
Chapter 6	Chronic respiratory symptoms associate with airway wall thickening measured by thin-slice low-dose CT	129
Chapter 7	Novel genes for airway wall thickness identified with combined genome wide association and expression analyses	147
Chapter 8	Summary, discussion and future perspectives	177
Chapter 9	Nederlandse samenvatting (Dutch Summary)	193
	Dankwoord	205

Chapter 1

General introduction

AE Dijkstra

General introduction

Etiology and function of mucus in the airways

Airways are remarkably resistant to damage by environmental factors, despite being constantly exposed to toxic substances, small particles and pathogenic organisms that are present in the 10,000-20,000 liter air that is inhaled during a day. This protection is for an important part provided by a layer of mucus that covers epithelial cells of the airway wall.

In the airways of a healthy person, mucus forms a tight layer and is a barrier for the airway epithelium and the underlying bronchial and lung tissue against environmental exposures. Mucus is a complex aqueous solution secreted by secretory cells (goblet-, clara- and serous cells) and in the larger airways (Ø > 2 mm) also by submucosal glands. Mucins, high-molecular-weight glycoproteins providing the viscoelasticity of mucus, are the main component of mucus. Furthermore, mucus contains antimicrobial molecules (e.g. IgA antibodies), immunomodulatory molecules (e.g. cytokines) and protective molecules (e.g. mucins, collectins).

When the airway epithelium is irritated, immediate secretion of mucus occurs to cover inhaled particles and dilute inhaled agents. Inhaled foreign particles and pathogenic organisms stick to the thin fluid film of mucus that is moved towards the throat by coordinated beating cilia of airway epithelial cells. This process is called mucociliary clearance and may be enhanced from time to time by cough, a reflex of the airways to release mucus from the upper airways. This whole system is designed to remove foreign particles and pathogenic organisms present in the inhaled air as quickly as possible and to minimize an inflammatory reaction.

Chronic Mucus Hypersecretion

When irritation of the airways persists during an extended period of time, a chronic inflammatory response of the airways arises, the protecting homeostatic function of airway mucus secretion is lost and goblet cell hyperplasia is induced. As a result, excessive production of airway mucus occurs next to other airway wall changes and leads to airway obstruction. The clinical presentation of a particular syndrome called chronic mucus hypersecretion (CMH) develops. CMH is often associated with changes in the composition and viscosity of mucus resulting in impairment of the mucociliary clearance and accumulation of mucus in the airways as a consequence. This mucus forms a culture

medium for bacteria and contributes to obstruction of the airways. In case of CMH, the role of mucus changes from a protective to a disruptive medium.

For epidemiologic studies, a definition of CMH has been formulated: the presence of sputum production during at least 3 months in two consecutive years without another explaining origin[1]. Airway obstruction is not included in this definition. Next to this definition, a great variation of definitions of CMH has been used in the literature (Table 1). Depending on the study populations and differences in the definition of CMH used, the prevalence of CMH in the general population as reported by questionnaire surveys, varies from 3.5% - 12.7% [2-14].

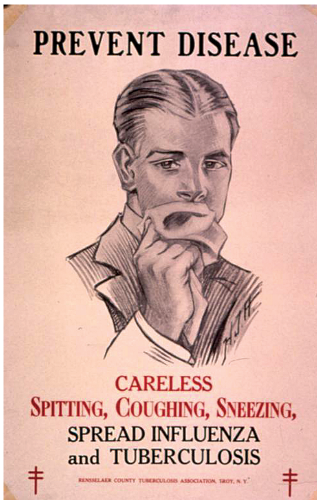
Table 1. Different definitions of chronic mucus hypersecretion

Author	year	Question or Definition
Fletcher	1959	Do you usually **bring up phlegm from your chest** (not from back of your nose), on most days for at least three months every year?
Lende	1965	Do you **expectorate sputum almost every day** more than 3 months a year?
Strachan	1988	Do you usually **bring up any phlegm from your chest** during the day or at night in the winter?
Zock	2001	**Regular expectoration of phlegm** for at least 3 months in each year.
Harmsen	2009	Have you **ever had cough with production of phlegm** for at least 3 months in 2 successive years
Johannessen	2011	Do you usually **have phlegm when coughing?** Do you cough each day for three months or more during a year?
Skorge	2009	Do you have **phlegm when coughing?**

Historic perspective of chronic mucus hypersecretion

In the first half of the 20th century, chronic production of phlegm and coughing were strongly associated with tuberculosis (TB), and when some blood was mixed with cough expectorations a TB-diagnosis was inescapable. TB was especially a disease of poor people and the cause of death for many; in the year 1900, 10,000 people died from TB in the Netherlands on a total population of 5 million inhabitants.

Assessment of the cause of TB (mycobacterium tuberculosis, Koch 1882) and development of antibiotics together with better social circumstances reduced the presence of this disease. Cough and phlegm were thought to be no longer a serious problem; subsequently, a period with less interest for these respiratory symptoms evolved.



After World War II, there was a significant increase in the number of Europeans who began to smoke. As a consequence the incidence of respiratory symptoms again increased. In 1964, British investigators hypothesized that recurrent bronchial infections were the reason that some smokers developed progressive airway obstruction and others did not. They proposed that the presence of mucus was a prerequisite to the development of airway obstruction. In 1977 Fletcher designed a study to test the hypothesis that extensive mucus production and airway obstruction are related. This relation was, however, not existent: chronic cough, sputum production, and respiratory infections appeared not to be related to airway obstruction in the general population. Both, CMH and lung function level indeed were associated with smoking, but they were not causally related to each other 15. This result was confirmed by subsequent studies that have revealed either no, or little, relation between airflow limitation and CMH [16, 17].

Finally, in 1996 a longitudinal well-powered study showed a clear relationship between CMH and lung function decline in the general population [18]. Moreover, a longitudinal study by Annesi et al. in 1986 showed that CMH was indeed not merely an innocent symptom as they observed a correlation between the presence of CMH and increased mortality risk [19].

Nowadays, it is generally accepted that CMH is an important symptom, associated with excess decline in forced expiratory volume in 1 second (FEV_1) (Figure 1), longer duration and more frequent respiratory infections, and increased hospitalization and mortality rates [5, 18, 20, 21].

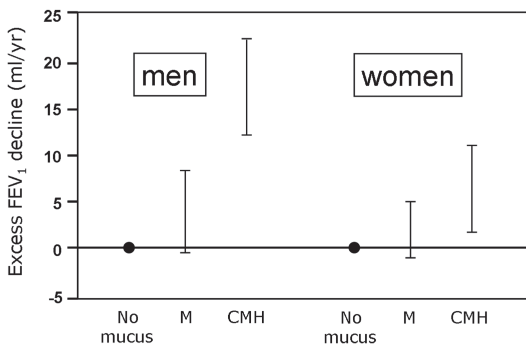


Figure 1. The association between any mucus secretion (M) or chronic mucus hypersecretion (CMH) and FEV1 decline, adjusted for age, height, weight change, and smoking for men and for women (Vestbo, 1996).

Risk factors for CMH, association with disease

Risk factors for CMH are lower socioeconomic status, higher age, male gender, and environ-mental and occupational exposures [22, 23]. The best studied and most important environmental risk factor for CMH is active smoking [2, 24]. Moreover, there is evidence that exposure to maternal smoking during pregnancy and environmental tobacco smoke exposure (ETS) in childhood are additional risk factors associated with the presence of CMH in adulthood [25-29].

In population-based studies, investigating effects of occupational exposures on CMH, an increased risk for CMH was found in agricultural-, textile-, paper-, wood-, chemical-, and food processing workers, the risks being most prominent in smokers [7, 10, 30].

Chronic airway inflammation, airway wall remodeling and thickening

Mucus is secreted by subtypes of secretory cells, goblet-, clara- and serous cells in epithelium, and in the larger airways (Ø > 2 mm) also by submucosal glands. In case of mucus secretory response, goblet cell hyperplasia and submucosal gland hyperplasia can extend to the peripheral airways.

The pathologic process underlying CMH is continuous inflammation of the epithelial layer of the airways induced by environmental factors (e.g. smoking and occupational hazards) accompanied by increased numbers of inflammatory cells like macrophages, mast cells and T lymphocytes. Chemokines generated by cigarette smoke together with inflammatory cells injure the airway and activate programs for remodeling, including goblet cell hyperplasia and/or *MUC* gene regulation, contributing to abnormal altered tissue structure and subsequently airway wall thickening (Figure 2). In addition to the ongoing inflammatory process, there is constant tissue breakdown and tissue repair in the form of fibrosis and angiogenesis, also contributing to remodeling and airway wall thickening.

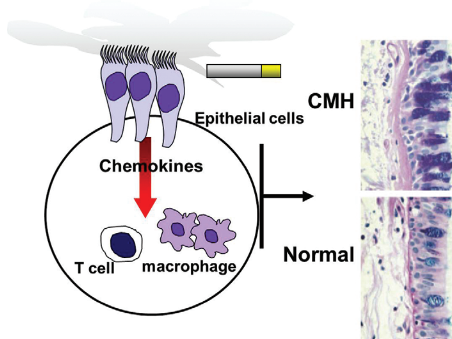


Figure 2. Pathology of chronic mucus hypersecretion

Airway wall remodelling and thickening is found over the total length of the respiratory tract and is associated with chronic CMH in larger airways and with airway obstruction in smaller airways [31].

In the past, knowledge on the process of airway wall thickening was mainly obtained by autopsy and bronchoscopic studies. Nowadays, quantification of airway dimensions has become feasible with the introduction of multi detector computed tomography (CT) and software tools for image analysis. This software automatically extracts airway centerlines, re-samples images perpendicular to the airway direction at equally spaced positions along the centerline and detects inner and outer airway wall borders in these images, based on density differences on CT (Figure 3).

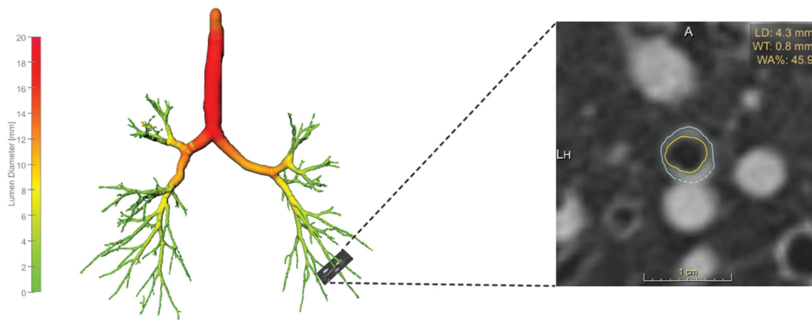


Figure 3. Projected cross-sectional image of an airway oriented perpendicular to the local centerline direction.

This non-invasive technique offers opportunities for longitudinal research, in particular when low-dose CTs are used, as the radiation-dose for low-dose CT is much lower compared to the dose of the usually used high-resolution CTs, and thus has acceptable health risks for the profits it renders.

CMH and COPD

COPD is a complex disease defined by the presence of not fully reversible and progressive airflow limitation that seriously reduces quality of life and is mainly associated with cigarette smoking [32]. In about 30% of all heavy smokers COPD is accompanied by CMH. For more than 10 years CMH without airway obstruction has even been marked as an early stage of COPD (GOLD stage 0) [33]. However, research by Vestbo et al. determined that GOLD Stage 0 was not a predictor of airway obstruction whereas smoking is; smokers with and without respiratory symptoms had comparable risk for developing airway obstruction [34]. The question rises whether risk factors (environmental and genetic) for CMH are the same in individuals with and without COPD, as not every individual with CMH has COPD and vice versa.

Worldwide, COPD affected 65 million people in 2004 and more than 3 million people died of COPD in 2005, representing 5% of all deaths. It is predicted that COPD will be the third leading cause of death worldwide in 2030 [35]. COPD seriously reduces quality of life and is responsible for high costs, i.e. yearly costs of COPD in the United States of America were $14,500 million in 2000 [36]. In the Netherlands, the total number of people diagnosed with COPD was 323,600 in 2007 and 5,984 persons died from COPD as the primary cause of death in 2010 (Figure 4). Costs of health care for Dutch people with COPD in 2000 were about €280 million and the expectation is that these costs will increase to about €440 to €495 million, at constant prices for healthcare in 2025 [37].

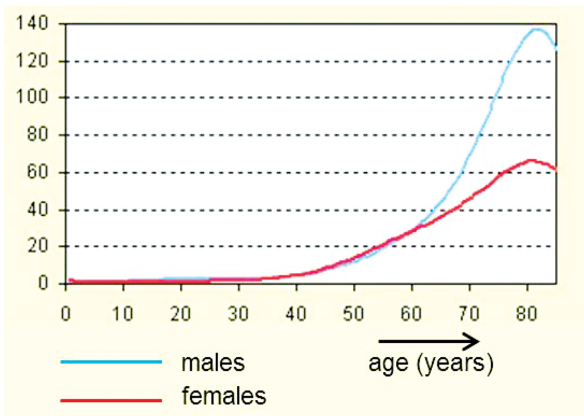


Figure 4. Prevalence of COPD (per 1,000), presented by age and sex, in January 2007 in the Netherlands [37].

Three sub-phenotypes of COPD are identified based on clinical and pathological characteristics and the localization of the inflammation in the lungs: small airways disease (SAD, airways ø < 2mm), emphysema (parenchymal destruction) and chronic bronchitis (larger airways). These sub-phenotypes are present in different combinations in COPD. SAD and emphysema are predominantly responsible for the severity of airway obstruction. Chronic bronchitis - characterized by chronic cough and/or chronic mucus hypersecretion (CMH) - constitutes a risk for more rapid lung function decline, increased duration and frequency of respiratory infections, hospitalization and mortality in individuals with established COPD [38]. Moreover, individuals with CMH have a 4-fold risk of mortality compared to those without CMH [2].

Genetic susceptibility to airway disease

As not every individual under the same environmental and lifestyle conditions suffers from CMH it is likely that, next to environmental factors, a genetic predisposition plays a role. This is supported by increased prevalence of CMH in monozygotic twins and familial

aggregation of persons with increased mucus production [39-41]. There is little known about genes that are associated with CMH. One publication suggested that the *CTLA4* gene is associated with chronic bronchitis in COPD, but not directly with COPD [42].

Also airway wall thickening is not, or not to the same extent, present in individuals under the same environmental and lifestyle conditions. Evidence for a genetic predisposition of airway wall thickening is also obtained by a familial aggregation study and shows that emphysema and airway wall thickening, both phenotypes of COPD, demonstrate independent aggregation within families of individuals with COPD, suggesting that different genetic factors influence the development of these diseases [43]. Furthermore, several candidate genes for COPD showed evidence for association with airway wall thickness in severe COPD-patients [44].

Outline of this thesis

In this thesis, studies are described that investigate environmental and genetic factors associated with chronic mucus hypersecretion (CMH) and/or with airway wall thickening. Exposure to cigarette smoke is regarded as the main causal environmental factor for these airway diseases. However, also occupational exposures are potential risk factors. For CMH and for airway wall thickening a genetic background is plausible since not all individuals under the same environmental conditions are affected.

The first four chapters in this thesis, after the introduction in **chapter 1**, focus on CMH.

In **chapter 2** we investigate whether risk factors for CMH are the same in individuals with and without COPD, as not every individual with CMH has COPD and vice versa.

Chapter 3 describes the search for a genetic origin of CMH in a heavy smoking population. Not all these heavy smokers have CMH suggesting a segregation of genetic predisposition. Since little is known about genes involved in CMH, we used a hypothesis-free genome-wide association study (GWAS) to investigate the genetic factors.

In **chapter 4** we hypothesized that apart from the well-known genes associated with COPD development, additional genes might contribute to CMH in COPD.

To investigate this idea we performed a GWAS in individuals with COPD. Hereafter, we focus on airway wall thickening.

First, in **chapter 5** we used a novel technology that makes accurate measurement of airway wall thickness possible. Low-dose CT and new software for rapid evaluations of large datasets, both never used in a large clinical study, have been used to study the relation between airway wall thickening and smoking, respiratory symptoms, emphysema, and airflow limitation.

In **chapter 6** we investigated the relation between respiratory symptoms and airway wall thickening in larger airways.

In **chapter 7** the question raises whether besides environmental factors also genetic factors explain why some individuals do develop airway wall thickening and others do not. Subsequently we describe the GWAS we performed to search for genes involved in airway wall thickening.

Finally, in **chapter 8** the main results and conclusions of the different studies are summarized and presented.

References

1 Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. American Thoracic Society. *American journal of respiratory and critical care medicine*. 1995;152(5 Pt 2):S77-121

2 Lange P, Nyboe J, Appleyard M, Jensen G and Schnohr P. Relation of ventilatory impairment and of chronic mucus hypersecretion to mortality from obstructive lung disease and from all causes. *Thorax*. 1990;45(0040-6376; 0040-6376; 8):579-585

3 Suadicani P, Hein HO, Meyer HW and Gyntelberg F. Exposure to cold and draught, alcohol consumption, and the NS-phenotype are associated with chronic bronchitis: an epidemiological investigation of 3387 men aged 53-75 years: the Copenhagen Male Study. *Occupational and environmental medicine*. 2001;58(3):160-164

4 de Meer G, Kerkhof M, Kromhout H, Schouten JP and Heederik D. Interaction of atopy and smoking on respiratory effects of occupational dust exposure: a general population-based study. *Environmental health : a global access science source*. 2004;3(1):6

5 Lange P, Parner J, Prescott E and Vestbo J. Chronic bronchitis in an elderly population. *Age and Ageing*. 2003;32(0002-0729; 6):636-642

6 Sunyer J, Zock JP, Kromhout H et al. Lung function decline, chronic bronchitis, and occupational exposures in young adults. *American journal of respiratory and critical care medicine*. 2005;172(9):1139-1145

7 Matheson MC, Benke G, Raven J et al. Biological dust exposure in the workplace is a risk factor for chronic obstructive pulmonary disease. *Thorax*. 2005;60(8):645-651

8 Halbert RJ, Natoli JL, Gano A, Badamgarav E, Buist AS and Mannino DM. Global burden of COPD: systematic review and meta-analysis. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*. 2006;28(3):523-532

9 Jaen A, Zock JP, Kogevinas M, Ferrer A and Marin A. Occupation, smoking, and chronic obstructive respiratory disorders: a cross sectional study in an industrial area of Catalonia, Spain. *Environmental health : a global access science source*. 2006;5:2

10 Zock JP, Sunyer J, Kogevinas M, Kromhout H, Burney P and Anto JM. Occupation, chronic bronchitis, and lung function in young adults. An international study. *American journal of respiratory and critical care medicine*. 2001;163(7):1572-1577

11 LeVan TD, Koh WP, Lee HP, Koh D, Yu MC and London SJ. Vapor, dust, and smoke exposure in relation to adult-onset asthma and chronic respiratory symptoms: the Singapore Chinese Health Study. *American Journal of Epidemiology*. 2006;163(12):1118-1128

12 Krstev S, Ji BT, Shu XO et al. Occupation and chronic bronchitis among Chinese women. *Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine*. 2008;50(1):64-71

13 Skorge TD, Eagan TM, Eide GE, Gulsvik A and Bakke PS. Occupational exposure and incidence of respiratory disorders in a general population. *Scandinavian journal of work, environment & health*. 2009;35(6):454-461

14 de Oca MM, Halbert RJ, Lopez MV et al. The chronic bronchitis phenotype in subjects with and without COPD: the PLATINO study. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*. 2012;40(1):28-36

15 Fletcher C and Peto R. The natural history of chronic airflow obstruction. *British medical journal*. 1977;1(0007-1447; 0007-1447; 6077):1645-1648

16 Clement J and Van de Woestijne KP. Rapidly decreasing forced expiratory volume in one second or vital capacity and development of chronic airflow obstruction. The American Review of Respiratory Disease. 1982;125(5):553-558

17 Higgins MW, Keller JB, Becker M et al. An index of risk for obstructive airways disease. The American Review of Respiratory Disease. 1982;125(2):144-151

18 Vestbo J, Prescott E and Lange P. Association of chronic mucus hypersecretion with FEV1 decline and chronic obstructive pulmonary disease morbidity. Copenhagen City Heart Study Group. American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association. 1996;153(1073-449; 5):1530-1535

19 Annesi I and Kauffmann F. Is respiratory mucus hypersecretion really an innocent disorder? A 22-year mortality survey of 1,061 working men. *The American Review of Respiratory Disease*. 1986;134(4):688-693

20 Vestbo J and Rasmussen FV. Respiratory symptoms and FEV1 as predictors of hospitalization and medication in the following 12 years due to respiratory disease. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology. 1989;2(0903-1936; 8):710-715

21 Ekberg-Aronsson M, Pehrsson K, Nilsson JA, Nilsson PM and Lofdahl CG. Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. Respiratory research. 2005;6:98

22 Mahesh PA, Jayaraj BS, Prabhakar AK, Chaya SK and Vijayasimha R. Prevalence of chronic cough, chronic phlegm & associated factors in Mysore, Karnataka, India. *The Indian journal of medical research*. 2011;134(1):91-100

23 Ferre A, Fuhrman C, Zureik M et al. Chronic bronchitis in the general population: influence of age, gender and socio-economic conditions. *Respiratory medicine*. 2012;106(3):467-471

24 Vestbo J and Hogg JC. Convergence of the epidemiology and pathology of COPD. *Thorax*. 2006;61(1):86-88

25 Cook DG, Strachan DP and Carey IM. Health effects of passive smoking. 9. Parental smoking and spirometric indices in children. *Thorax*. 1998;53(10):884-893

26 Svanes C, Omenaas E, Jarvis D, Chinn S, Gulsvik A and Burney P. Parental smoking in childhood and adult obstructive lung disease: results from the European Community Respiratory Health Survey. *Thorax*. 2004;59(4):295-302

27 Skorge TD, Eagan TM, Eide GE, Gulsvik A and Bakke PS. Indoor exposures and respiratory symptoms in a Norwegian community sample. *Thorax*. 2005;60(11):937-942

28 David GL, Koh WP, Lee HP, Yu MC and London SJ. Childhood exposure to environmental tobacco smoke and chronic respiratory symptoms in non-smoking adults: the Singapore Chinese Health Study. *Thorax*. 2005;60(12):1052-1058

29 Johannessen A, Bakke PS, Hardie JA and Eagan TM. Association of exposure to environmental tobacco smoke in childhood with chronic obstructive pulmonary disease and respiratory symptoms in adults. *Respirology (Carlton, Vic.)*. 2012;17(3):499-505

30 Blanc PD, Iribarren C, Trupin L et al. Occupational exposures and the risk of COPD: dusty trades revisited. *Thorax*. 2009;64(1):6-12

31 Hogg JC, Chu F, Utokaparch S et al. The nature of small-airway obstruction in chronic obstructive pulmonary disease. *The New England journal of medicine*. 2004;350(1533-4406; 0028-4793; 26):2645-2653

32 Rabe KF, Hurd S, Anzueto A et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association*. 2007;176(1073-449; 6):532-555

33 Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS and GOLD Scientific Committee. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. *American journal of respiratory and critical care medicine*. 2001;163(5):1256-1276

34 Vestbo J and Lange P. Can GOLD Stage 0 provide information of prognostic value in chronic obstructive pulmonary disease? *American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association*. 2002;166(1073-449; 3):329-332

35 http://www.who.int/respiratory/copd/burden/en/index 2013

36 Wilson L, Devine EB and So K. Direct medical costs of chronic obstructive pulmonary disease: chronic bronchitis and emphysema. *Respiratory medicine*. 2000;94(3):204-213

37 Rijks instituut voor gezondheid en milieu, ministerie voor Volksgezondheid, Welzijn en Sport. http://www.nationaalkompas.nl/gezondheid-en-ziekte/ziekten-en-aandoeningen/ademhalingswegen/copd/cijfers-copd-prevalentie-incidentie-en-sterfte-uit-de-vtv-2010

38 Vestbo J. Epidemiological studies in mucus hypersecretion. *Novartis Foundation symposium*. 2002;248(1528-2511; 1528-2511):3-12

39 Viegi G, Carrozzi L, Di Pede F et al. Risk factors for chronic obstructive pulmonary disease in a north Italian rural area. European journal of epidemiology. 1994;10(6):725-731

40 Silverman EK and Speizer FE. Risk factors for the development of chronic obstructive pulmonary disease. The Medical clinics of North America. 1996;80(0025-7125; 0025-7125; 3):501-522

41 Hallberg J, Dominicus A, Eriksson UK et al. Interaction between smoking and genetic factors in the development of chronic bronchitis. *American journal of respiratory and critical care medicine*. 2008;177(5):486-490

42 Zhu G, Agusti A, Gulsvik A et al. CTLA4 gene polymorphisms are associated with chronic bronchitis. *Eur.Respir.J.* 2009;(1399-3003)

43 Patel BD, Coxson HO, Pillai SG et al. Airway wall thickening and emphysema show independent familial aggregation in chronic obstructive pulmonary disease. American Journal of Respiratory & Critical Care Medicine. 2008;178(5):500-505

44 Kim WJ, Hoffman E, Reilly J et al. Association of COPD candidate genes with computed tomography emphysema and airway phenotypes in severe COPD. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology. 2011;37(1):39-43

Chapter 2

Risk factors for chronic mucus hypersecretion in individuals with and without COPD

Influence of smoking and job exposure on CMH

AE Dijkstra, K de Jong, HM Boezen, H Kromhout, R Vermeulen, HJM Groen, DS Postma, JM Vonk

Journal of Occupational and Environmental Medicine, 2014 March

Abstract

Background

Chronic mucus hypersecretion (CMH) is highly prevalent in smokers and associated with an accelerated lung function decline and COPD. Several risk factors contribute to CMH and to COPD. It is however unknown if risk factors for CMH are similar in subjects with and without COPD.

Methods

1,479 subjects with and 8,529 without COPD, participating in the general population based LifeLines cohort, completed questionnaires and underwent spirometry. Occupational exposure was assessed using the ALOHA+ job exposure matrix. Analyses were performed using multiple logistic regression models.

Results

In COPD, a significantly higher risk for CMH was associated with higher packyears smoking (per 10 packyears) (OR = 1.28; 1.12-1.46) and environmental tobacco smoke (ETS) (OR = 2.06; 1.33-3.19). In non-COPD; male gender (OR = 1.91; 1.51-2.41), higher body mass index (OR = 1.04; 1.01-1.06), higher packyears smoking (OR = 1.28; 1.14-1.44), current smoking (OR = 1.50; 1.04-2.18), low and high exposure to mineral dust (OR = 1.39; 1.04-1.87 and OR = 1.60; 1.02-2.52), high exposure to gases & fumes (OR = 2.19; 1.49-3.22). Significant interactions were found between COPD and exposure to gases & fumes ($p = 0.018$) and aromatic solvents ($p = 0.038$).

Conclusions

A higher risk for CMH was associated with higher packyears smoking regardless of COPD status. However, a higher risk for CMH was associated with high occupational exposure to gases & fumes in individuals without COPD only.

Introduction

The secretion of mucus is a normal response of epithelial cells in order to protect the airways and lung tissue against inhaled pathogens, particles and noxious chemicals. In contrast, chronic mucus hypersecretion (CMH) is abnormal. CMH is a condition of mucus overproduction defined by mucus production for at least 3 months during the last 2 years, when specific causes have been excluded [1]. The prevalence of CMH in the general population varies from 3.5% to 12.7% depending on the study population and the CMH definitions used [2, 3]. In the general population, CMH is associated with an increased duration and frequency of respiratory infections, excess decline of the forced expiratory volume in 1 second (FEV1), and increased hospitalization and mortality rates [2, 4-6].

The best studied and most important risk factor for CMH is cigarette smoking [2, 7]. Other risk factors for CMH are higher age and male gender [8, 9]. Of interest, the presence of respiratory infections in childhood is a risk factor for CMH and also for development of Chronic Obstructive Pulmonary Disease (COPD), as is smoking [7, 10]. Next to active smoking there is evidence that exposure to maternal smoking during pregnancy (passive smoking in utero) and environmental tobacco smoke exposure (ETS) in childhood are additional risk factors for the presence of CMH in adulthood [11-15]. Occupational exposures have been mentioned as risk factors for CMH in many general population based studies, and have also been reported as risk factors for COPD in different studies [3, 16, 17]. In addition, CMH is present in about 30% of COPD patients where it constitutes a risk factor for increased duration and frequency of respiratory infections, hospitalization and mortality and higher risk for exacerbations [18, 19].

Above studies show that CMH can be present, both in subjects with and without COPD and some risk factors for COPD overlap with those for CMH, like smoking and bacterial infections. However, not all patients with COPD have CMH and conversely not all individuals with CMH have COPD. We therefore investigated whether risk factors for CMH differ between subjects with and without COPD. To this aim we used data of the LifeLines cohort, a general population based study in the northern part of The Netherlands, and determined risk factors for CMH in subjects with and without COPD taking into account well-known clinical, demographic and environmental factors contributing to CMH (active smoking, exposure to environmental tobacco smoke, and occupational exposures).

Methods

Study Population and Methods

"To investigate risk factors for CMH we included subjects participating in the Dutch LifeLines cohort study. The LifeLines study is a multidisciplinary prospective general population-based study among residents of the three northern provinces of The Netherlands, investigating the origins and the development of chronic diseases and multimorbidity [20]. Subjects were recruited via general practitioners. In the current study, we included 13,301 Caucasian adults, aged between 18 and 90 years, from the second data release of the LifeLines cohort All participants gave written informed consent, completed questionnaires and underwent a medical examination and standardized spirometry, according to the ERS guidelines [21]. In this population-based study we did not administer a bronchodilator.

The exact question used to define CMH was "do you usually expectorate sputum during day or night in winter on the majority of days ≥ 3 months a year? (yes/no)".

Since it is known that the presence of asthma can cause symptoms of CMH, subjects with asthma (ever having asthma confirmed by a physician) were excluded from the analyses (n = 953).

Environmental tobacco smoke (ETS) exposure and smoking habits

Exposure to smoke during childhood was determined by the question:"did your mother/ father smoke regularly during your childhood?" (yes/no). Furthermore, current ETS exposure was determined by questions about regularly exposure to smoke from others during the last year for at least 1 hour per day (yes/no), and in case of a paid job, whether smoking was present in the workplace (yes/no). Smoking habits were defined as never smoking, ex smoking and current smoking and the lifetime number of packyears smoked. An individual was defined as being a current smoker if he/she answered 'yes' to the question: do you smoke now or have you been smoking in the last month. A never smoker when answered 'no' to the question: have you ever smoked for as long as a year, and an ex-smoker answered 'yes' to the question: have you ever smoked for as long as a year and 'no' to the question: do you smoke now or have you been smoking in the last month and 'yes' to the question: did you currently quit smoking. Packyears of smoking were calculated as the number of packs of cigarettes (1 pack = 20 cigarettes) smoked per day times the number of years of smoking.

Occupational exposure

Information on employment status, job title and description of work tasks of the current job (or last held job in case of retirement) was obtained by questionnaire and coded according to the International Standard Classification of Occupations version 1988 (ISCO-88) [22].

Employed and unemployed subjects were included in this study. The ALOHA+ Job Exposure Matrix (JEM) was used to classify the reported jobs into no, low or high exposure to various agents (coded respectively 0, 1 or 2) [16]. If someone had two or more jobs (n = 232, 2.3%), the average occupational exposure was determined by rounding the average to the nearest integer (0.5 = 1 and 1.5 = 2).

Statistical analyses

Analyses were stratified for COPD defined as $FEV_1/FVC < 70\%$. Body mass index (BMI) was defined as weight/height^2 (kg/m^2). Differences in characteristics and occupational exposures between subjects with and without CMH stratified by COPD were analyzed using chi-square tests and 2-tailed unpaired Student's t-tests. Characteristics significantly associated with CMH (except for the occupational exposures and lung function), were included in a multivariate logistic regression model. Subsequently, each occupational exposure was included in this model one-by-one without taking into account other occupational exposures. The interaction effect between COPD and the other possible risk factors was tested by using a multivariate regression model including COPD x risk factor as an extra variable in the model.

Since the prevalence of exposures to herbicides and insecticides was very low in our population (1.3% vs. 3.5%), we analyzed all pesticides as one variable (prevalence 4.0%). Differential effects of the possible risk factors between subjects with and without COPD were tested in unstratified multivariate models by including the appropriate interaction terms. In an additional analysis, retired and unemployed subjects were excluded (n = 1,996) to assess the effect of current occupational exposure only. Finally, analyses were stratified by age, gender, and smoking habits to investigate possible effect modification by these variables.

Analyses were conducted using SPSS version 19.0. A two-sided $p < 0.05$ was considered statistically significant.

Results

From the initial LifeLines sample of 13,301 subjects a total of 2,340 was excluded because of incomplete data on CMH (n = 356), lacking information on smoking habits and ETS (n = 1,568) and incomplete data on lung function (n = 416). After exclusion of asthmatics (n = 953) 10,008 subjects (75.8% of all subjects) remained, including 1,479 (14.8%) with and 8,529 without COPD.

Characteristics, ETS and smoking habits related to CMH

Table 1 presents the demographics of subjects with and without CMH, stratified by COPD status. The overall prevalence of CMH was 4.2% and was significantly higher in subjects with COPD (8.7%) than in subjects without COPD (3.4%, p < 0.001). In both, subjects with and without COPD, the prevalence of CMH was significantly higher in males, in ever- and current smokers and in subjects with ETS exposure; the number of packyears smoked was also significantly higher in subjects with CMH. COPD subjects with CMH had significantly worse lung function than those without CMH.

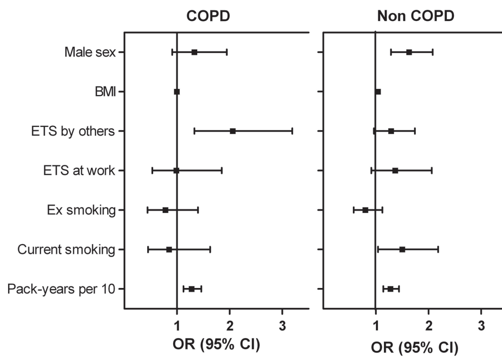
Table 1. Demographic characteristics of subjects with and without chronic mucus hypersecretion, and association of these characteristics with chronic mucus hypersecretion (expressed as p-value), stratified by COPD (n = 10,008).

Variable	Non-COPD n = 8,529 (85.2%)			COPD n = 1,479 (14.8%)		
	No CMH	**CMH**	**p**	**No CMH**	**CMH**	**p**
N (%)	**8,236 (96.6)**	**293 (3.4)**		**1,350 (91.3)**	**129 (8.7)**	
Males, n (%)	**1,907 (40.0)**	**164 (56.0)**	**< 0.001**	**619 (45.9)**	**73 (56.6)**	**0.020**
Age (years), median (range)	46.4 (18-89)	47.0 (26-79)	0.257	52.3 (26-86)	51.0 (34-86)	0.933
BMI (kg/height2), mean (SD)	**26.2 (4.2)**	**27.0 (4.7)**	**0.001**	26.0 (3.7)	26.3 (4.6)	0.549
Lung function						
FEV1 (liter), mean (SD)	3.5 (0.8)	3.6 (0.8)	0.142	**2.8 (0.8)**	**2.7 (0.8)**	**0.006**
FEV1/FVC (%), median (range)	**78.8 (70-100)**	**78.0 (70-96)**	**0.016**	**66.0 (39-70)**	**65.0 (33-70)**	**< 0.001**
# FEV1, % predicted (%), mean (SD)	105.2 (12.6)	104.0 (12.6)	**0.097**	**91.7 (14.3)**	**84.6 (17.1)**	**< 0.001**
ETS						
By mother during childhood, n (%)	2,820 (34.3)	92 (31.5)	0.316	492 (31.8)	48 (33.8)	0.400
By father during childhood, n (%)	6,288 (76.6)	229 (78.2)	0.545	1,283 (83.4)	124 (87.3)	0.284
*By others, n (%)	**1,756 (21.3)**	**108 (36.9)**	**< 0.001**	**323 (23.9)**	**56 (43.4)**	**< 0.001**
At work, n (%)	**472 (5.7)**	**36 (12.3)**	**< 0.001**	97 (7.2)	15 (11.6)	0.071
Smoking habits						
Never smoking, n (%)	**3,711 (45.1)**	**97 (33.1)**	**< 0.001**	**357 (26.4)**	**23 (17.8)**	**< 0.001**
Ex smoking, n (%)	**3,006 (36.5)**	**86 (29.4)**	**0.013**	580 (43.0)	47 (36.4)	0.153
Packyears, median (range)	7.0 (0.05-7.5)	7.9 (0.1-47.0)	0.525	**10.9 (0.05-100)**	**16.3 (0.05-67.5)**	**0.005**
Current smoking, n (%)	**1,519 (18.4)**	**110 (37.5)**	**< 0.001**	**413 (30.6)**	**59 (45.7)**	**0.001**
Packyears, median (range)	**13.5 (0.25-70.5)**	**21.0 (1.05-84.0)**	**< 0.001**	**21.4 (0.45-100)**	**26.3 (2.75-69.0)**	**0.003**

*CMH = chronic mucus hypersecretion; BMI = body mass index; ETS = environmental tobacco smoke; * at least one hour per day during the last year; bold = p-value < 0.05; # Lung function reference equations according to Quanjer et al.* [21]

Figure 1 and Table 2 present the results of the multivariate logistic regression analysis on the associations between risk factors and CMH, stratified by COPD, and the results of interaction analysis between risk factors and COPD.

Figure 1. Odds ratios and 95% CI for multivariate analysis showing association between chronic mucus hypersecretion and gender, BMI, ETS and smoking habits, stratified by COPD.



Packyears per 10: the unit in the analysis is 10 packyears so the OR is the estimate of 10 packyears.

In subjects with COPD, a higher number of packyears and current ETS exposure were significantly associated with a higher risk for CMH. In subjects without COPD, next to a higher number of packyears also male gender, higher BMI and current smoking were associated with a significant higher risk for CMH. None of the investigated interactions between the risk factors and COPD was statistically significant.

Table 2. Interaction analysis between COPD and characteristics, ETS (by others and at work), ex- and current smoking and packyears and multivariate logistic regression on association between chronic mucus hypersecretion and gender, BMI, ETS (by others and at work), ex- and current smoking and packyears, stratified by COPD.

Variables	Non-COPD		COPD		Interaction with COPD
	OR (95% CI)	p	OR (95% CI)	p	p
Gender (male)	**1.63 (1.29-2.10)**	**< 0.001**	1.33 (0.91-1.94)	0.142	0.276
BMI	**1.04 (1.01-1.06)**	**0.010**	1.00 (0.96-1.05)	0.860	0.345
ETS					
By others*	1.29 (0.96-1.74)	0.088	**2.06 (1.33-3.19)**	**0.001**	0.475
At work	1.37 (0.91-2.06)	0.128	0.99 (0.53-1.85)	0.975	0.561
Smoking habits					
Ex smoking	0.80 (0.58-1.12)	0.191	0.78 (0.44-1.40)	0.408	0.853
Current smoking	**1.50 (1.04-2.18)**	**0.032**	0.85 (0.45-1.63)	0.629	0.180
Packyears per 10	**1.28 (1.14-1.44)**	**< 0.001**	**1.28 (1.12-1.46)**	**< 0.001**	0.362

*BMI = body mass index; ETS = environmental tobacco smoke; * at least one hour per day during the last year; bold = p-value < 0.05. Packyears per 10: the unit in the analysis is 10 packyears so the OR is the estimate of 10 packyears*

Occupational exposure and risk for CMH

Table 3 presents the proportion of subjects without, with low or high exposure to occupational agents according to the ALOHA+ JEM, in subjects with and without chronic mucus hypersecretion, stratified by COPD. Almost 45% of the population had some occupational exposure, either low or high. Exposure to gases & fumes was the most frequent occupational exposure (40.1%). An overview of the most prevalent occupations within those exposed is given in Table 1 in the supplement.

In subjects with COPD, there was no significant difference in occupational exposures between subjects with and without CMH. In contrast, in subjects without COPD, the prevalence of 5 out of the 8 investigated occupational exposures was significantly different between subjects with and without CMH.

Statistically significant interactions were found between COPD and high exposure to gases & fumes and between COPD and low exposure to aromatic solvents (Supplement Table 2). In the stratified analyses, significant associations were found particularly between low and high exposure to mineral dust and CMH and between high exposure to gases & fumes, chlorinated solvents or heavy metals and CMH (adjusted for gender, BMI, ETS and smoking habits) in subjects without COPD. Figure 2 shows the odds ratios and 95% confidence intervals of occupational risk factors studied with respect to the presence of CMH, stratified by COPD. In subjects with COPD there were no significant associations between occupational exposures and CMH (Supplement, Table 2).

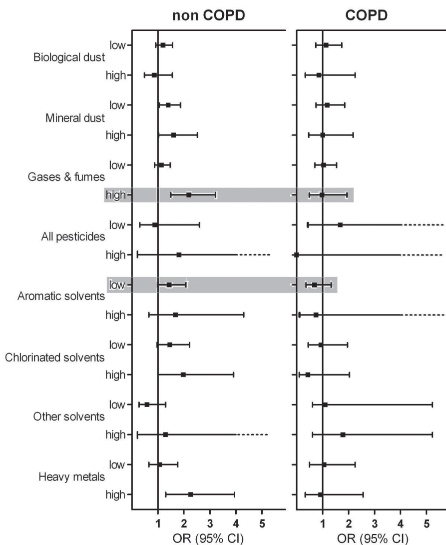


Figure 2. Odds ratios and 95% CI for multivariate analysis showing association between chronic mucus hypersecretion and occupational exposures, stratified by COPD.

Reference is not exposed; analysis corrected for gender, BMI, ETS, ex- and current smoking and packyears. Occupational exposures were added one by one; Gray frame: significant interaction ($p < 0.05$) between occupational exposure and COPD.

Table 3. Prevalence of occupational exposures, according to the ALOHA+ JEM, in subjects with and without chronic mucus hypersecretion, stratified by COPD.

Exposure		Non-COPD (n = 8,529)			COPD (n = 1,479)		
		No CMH (n = 8,236)	CMH (n = 293)		No CMH (n = 1,350)	CMH (n = 129)	
		n (%)	n (%)	p*	n (%)	n (%)	p*
Biological dust	No	5,674 (68.9)	196 (66.9)		935 (69.3)	88 (68.2)	
	Low	2,240 (27.2)	84 (28.7)	0.541	359 (26.6)	36 (27.9)	0.747
	High	322 (4.2)	13 (4.4)	0.715	56 (4.1)	5 (3.9)	0.882
Mineral dust	No	6,673 (81.0)	201 (68.6)		1,032 (76.4)	89 (69.0)	
	Low	**1,234 (15.0)**	**66 (22.5)**	**< 0.001**	240 (17.8)	31 (24.0)	0.079
	High	**329 (4.0)**	**26 (8.9)**	**< 0.001**	78 (5.8)	9 (7.0)	0.580
Gases & fumes	No	5,008 (60.8)	146 (49.8)		775 (57.4)	68 (52.7)	
	Low	2,812 (34.1)	105 (35.8)	0.423	482 (35.7)	48 (37.2)	0.733
	High	**416 (5.1)**	**42 (14.3)**	**< 0.001**	93 (6.9)	13 (10.1)	0.180
All pesticides	No	7,992 (96.2)	279 (95.2)		1,285 (95.2)	122 (94.6)	
	Low	251 (3.0)	11 (3.8)	0.607	47 (3.5)	4 (3.1)	0.821
	High	63 (0.8)	3 (1.0)	0.212	18 (1.3)	3 (2.3)	0.363
Aromatic solvents	No	7,559 (91.8)	250 (85.3)		1,212 (89.8)	116 (89.9)	
	Low	**618 (7.5)**	**38 (13.0)**	**0.003**	131 (9.7)	12 (9.3)	0.883
	High	59 (0.7)	5 (1.7)	0.081	7 (0.5)	1 (0.8)	0.704
Chlorinated solvents	No	7,665 (93.1)	250 (85.3)		1,247 (92.4)	118 (91.5)	
	Low	**464 (5.6)**	**27 (9.2)**	**0.013**	78 (5.8)	9 (7.0)	0.580
	High	**107 (1.3)**	**10 (3.4)**	**0.013**	25 (1.9)	2 (1.6)	0.807
Other solvents	No	6,301 (76.8)	221 (75.4)		1,055 (78.1)	103 (79.8)	
	Low	1,788 (21.7)	64 (21.8)	0.729	279 (20.7)	24 (18.6)	0.579
	High	147 (1.8)	8 (2.7)	0.302	16 (1.2)	2 (1.6)	0.718
Heavy metals	No	7,729 (93.8)	258 (88.1)		1,244 (92.1)	115 (89.1)	
	Low	**366 (4.4)**	**19 (6.5)**	**0.046**	70 (5.2)	9 (7.0)	0.387
	High	**141 (1.7)**	**16 (5.5)**	**< 0.001**	36 (2.7)	5 (3.9)	0.424

*CMH = chronic mucus hypersecretion; bold = p-value < 0.05; *p-value: unadjusted logistic regression, reference is not exposed (to the current investigated agent)*

Exclusion of retired and unemployed subjects to assess the effect of current occupational exposures did not change the results (results not shown). Stratification by age, gender or smoking habits (never-, ex- and current smoking) did not consistently indicate effect modification by these variables of the associations between occupational exposures and CMH (Supplement Table 3, 4 and 5).

Discussion

We report results from a large cross-sectional general population based study, relating demographic characteristics, environmental smoke exposure, smoking habits and occupational exposures to CMH in subjects with and without COPD. Subjects with COPD had a higher prevalence of CMH (defined by expectoration of sputum on most days ≥3 months during the last year) (8.7%) than to those without COPD (3.4%). The risk for CMH in subjects with COPD increased with higher packyears and ETS exposure only, without any effect of occupational exposures. In contrast, risk factors for CMH in subjects without COPD were male gender, higher BMI, current smoking, higher packyears and several occupational exposures. Interestingly, the association between CMH and high occupational exposure to gases & fumes differed significantly between subjects with and without COPD. Although the differences in the associations of the other occupational risk factors with CMH between subjects with and without COPD failed to reach statistical significance, the observed differences in effect sizes may be important.
The commonly reported prevalence of CMH in the general population ranges from 3.5% to 12.7% [2, 9]. The prevalence of CMH was 4.2% in our study, which is in the lower range of reported prevalences. When also asthmatics were included the prevalence was 4.8%. The prevalence of CMH in our study was comparable with the prevalence of CMH (defined in the same way), in another general population based cohort from the northern part of The Netherlands (Vlagtwedde), also when stratified for gender, smoking habits or COPD.

It has been well established that the presence of CMH increases with the severity of airflow limitation [18, 23]. Since our population encompassed subjects with relatively mild COPD according to GOLD the guidelines (80% stage 1, 20% stage 2), the relatively low prevalence of CMH in subjects with COPD of 8.7% is in line with the association of CMH with the lung function level [24].

We had only prebronchodilator lung function available in this population based study, which may have affected our prevalence of COPD and especially very mild COPD. For this same reason it is also possible that few undiagnosed asthmatics may incorrectly have been included in the COPD group. In a sensitivity analysis we used the lower limit of normal (LLN) to define COPD [25]. The results of this analysis showed that the prevalence of CMH and the directions and magnitudes of the associations remained similar (Supplement, Table 6).

In accordance with many other general population based studies we found that CMH is significantly more prevalent in males than in females [2, 3, 17]. A potential reason for this difference is a tendency for women to report more dyspnea and cough but less phlegm symptoms than men [26].

The association between packyears smoking and CMH is in accordance with the literature but was rarely examined separately for subjects with and without COPD in the general population [27]. We found this association to be present in both groups. This could mean that the cigarette smoke induced chronic inflammatory process and its associated remodeling of the airway walls, are the most important risk factors for CMH, thereby reducing the effects of other potential risk factors.
In addition to packyears, current smoking was significantly associated with an increased CMH-risk in subjects without but not in subjects with COPD. Since some individuals would have quitted for only a short time, this may have affected the results. Even when we excluded individuals who quitted smoking for only a short period (smoking cessation < 1 year, n = 31) or added these 31 subjects to the analysis in current smokers with COPD, current smoking was still not a significant risk factor for CMH. It is possible that the extensive and longstanding smoking history in subjects with COPD has resulted in irreversible airway damage which constitutes an overwhelming important contributor to CMH, more so than the current smoking status.

Occupational exposures

The ALOHA+ JEM assigns exposures to gases & fumes as well as exposures to mineral and biological dusts. Exposure to gases & fumes includes exposures to; aromatic, chlorinated and other solvents, to heavy metals and to all pesticides, which were also additionally separately allocated. Exposure to heavy metals contributes also to exposure to mineral dust.

In our study occupational exposures like mineral dust, gases & fumes, chlorinated solvents and heavy metals are significantly contributing to CMH in subjects without COPD, but not at all in subjects with COPD.

Supplementary Table 7 shows how this is related to findings in the literature published since 2000, reporting risk factors for CMH including occupational exposures, demographic characteristics and smoking habits in the general population. Of importance we have not found any study in general populations that performed stratified analyses for COPD status combined with detailed information on occupational exposures (job exposure matrix) and our findings are new in this respect.

Given the low numbers of subjects with COPD in the general population, results of the above mentioned studies will be driven primarily by subjects without COPD. This makes the results of these population-based studies comparable to our results in subjects without COPD. However, a considerable variation in the definitions used for CMH or chronic bronchitis (CB), and in definitions for (extent of) occupational exposures complicates comparisons.

Comparison of studies is further complicated by differences in age between populations, differences in habits (exposure in home caused by cooking) belonging to a continent, the registration of exposure (lifetime versus last job, self-reported versus a job exposure matrix).

Notwithstanding this, some studies have found an association between CMH and exposure to gases & fumes, and most studies have not found an association between CMH and biological dust, similar to our results.
The significant associations between CMH and low or high exposure to mineral dust, and between CMH and high exposure to heavy metals (separately) we found, were not found in other studies.

Since there are differential effects of occupational exposures on CMH in subjects with and without COPD, the question arises whether the pathophysiology of CMH is different as well. This clearly needs further study into differences given the composition, tenacity, viscosity and produced volume of sputum, as well as the type and level of inflammation, the involved genes and epigenetic phenomena. Furthermore, cigarette smoke causes damage from the central to the peripheral airways. This is a slow process which is accompanied by metaplasia of goblet cells and mucus hypersecretion that is located in the larger airways and also in the small airways in a later stage, accompanied by closure

of the small airways and subsequently airway obstruction. It remains to be established whether occupational exposures mainly affect the larger airways in subjects without COPD, yet with similar symptoms of CMH as occurring in smoking related COPD.

The strength of this study is that we had access to a large population, with a very wide age range and a considerable number of subjects with airflow limitation, which allowed us to study risk factors for CMH in subjects with and without COPD separately.
A limitation is the lack of information about life-time occupational exposure since we had information about occupational exposures during the current or last job only. Symptomatic subjects might have left jobs with exposures to occupational exposures before (early) retirement. However, an additional analysis in which unemployed and retired subjects were excluded contradicts the possibility of selective avoidance of hazardous occupational exposures; subjects with COPD had a similar or even higher prevalence of occupational exposures in their current job than subjects without (19% had exposure to mineral dust in non-COPD versus 24.4% in COPD, for gases and fumes being 39.9% and 44.9% respectively (results not shown)).
Comparison of provided reasons for unemployment in non-COPD and COPD revealed that the mean age in the COPD-group was considerably higher explaining the higher number of subjects who were retired or pre-retired in this group. The percentage of subjects who were incapable to work was comparable in both groups.

We believe that through legislation and awareness of the danger of these exposures, people are nowadays less exposed. We hypothesize that with using current or last held job we rather have under- than over-estimated the association between occupational exposures or ETS and risk for CMH. Clearly, studies including information on lifetime (cumulative) exposure are desirable to confirm the effects found.

We conclude that occupational exposures contribute differentially to CMH in subjects with and without COPD. In subjects with established COPD only the number of packyears smoked is associated with an increased risk for CMH and occupational exposures do not contribute. In contrast, high occupational exposure to gases & fumes (among which solvents, all pesticides and heavy metals) is an important driver of CMH in subjects without airflow limitation, next to packyears smoking.

References

1 Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. American Thoracic Society. *American journal of respiratory and critical care medicine*. 1995;152(5 Pt 2):S77-121

2 Lange P, Parner J, Prescott E and Vestbo J. Chronic bronchitis in an elderly population. *Age and Ageing*. 2003;32(0002-0729; 6):636-642

3 de Oca MM, Halbert RJ, Lopez MV et al. The chronic bronchitis phenotype in subjects with and without COPD: the PLATINO study. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*. 2012;40(1):28-36

4 Vestbo J and Rasmussen FV. Respiratory symptoms and FEV1 as predictors of hospitalization and medication in the following 12 years due to respiratory disease. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*. 1989;2(0903-1936; 8):710-715

5 Vestbo J, Prescott E and Lange P. Association of chronic mucus hypersecretion with FEV1 decline and chronic obstructive pulmonary disease morbidity. Copenhagen City Heart Study Group. *American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association*. 1996;153(1073-449; 5):1530-1535

6 Ekberg-Aronsson M, Pehrsson K, Nilsson JA, Nilsson PM and Lofdahl CG. Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. *Respiratory research*. 2005;6:98

7 Vestbo J and Hogg JC. Convergence of the epidemiology and pathology of COPD. *Thorax*. 2006;61(1):86-88

8 Mahesh PA, Jayaraj BS, Prabhakar AK, Chaya SK and Vijayasimha R. Prevalence of chronic cough, chronic phlegm & associated factors in Mysore, Karnataka, India. *The Indian journal of medical research*. 2011;134(1):91-100

9 Ferre A, Fuhrman C, Zureik M et al. Chronic bronchitis in the general population: influence of age, gender and socio-economic conditions. *Respiratory medicine*. 2012;106(3):467-471

10 Viegi G, Carrozzi L, Di Pede F et al. Risk factors for chronic obstructive pulmonary disease in a north Italian rural area. *European journal of epidemiology*. 1994;10(6):725-731

11 Cook DG, Strachan DP and Carey IM. Health effects of passive smoking. 9. Parental smoking and spirometric indices in children. *Thorax*. 1998;53(10):884-893

12 Svanes C, Omenaas E, Jarvis D, Chinn S, Gulsvik A and Burney P. Parental smoking in childhood and adult obstructive lung disease: results from the European Community Respiratory Health Survey. *Thorax*. 2004;59(4):295-302

13 Skorge TD, Eagan TM, Eide GE, Gulsvik A and Bakke PS. Indoor exposures and respiratory symptoms in a Norwegian community sample. *Thorax*. 2005;60(11):937-942

14 David GL, Koh WP, Lee HP, Yu MC and London SJ. Childhood exposure to environmental tobacco smoke and chronic respiratory symptoms in non-smoking adults: the Singapore Chinese Health Study. *Thorax*. 2005;60(12):1052-1058

15 Johannessen A, Bakke PS, Hardie JA and Eagan TM. Association of exposure to environmental tobacco smoke in childhood with chronic obstructive pulmonary disease and respiratory symptoms in adults. *Respirology (Carlton, Vic.)*. 2012;17(3):499-505

16 Matheson MC, Benke G, Raven J et al. Biological dust exposure in the workplace is a risk factor for chronic obstructive pulmonary disease. *Thorax*. 2005;60(8):645-651

17 Skorge TD, Eagan TM, Eide GE, Gulsvik A and Bakke PS. Occupational exposure and incidence of respiratory disorders in a general population. *Scandinavian journal of work, environment & health*. 2009;35(6):454-461

18 Vestbo J. Epidemiological studies in mucus hypersecretion. *Novartis Foundation symposium*. 2002;248(1528-2511; 1528-2511):3-12

19 Kim V and Criner GJ. Chronic bronchitis and chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine*. 2013;187(3):228-237

20 Stolk RP, Rosmalen JG, Postma DS et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *European journal of epidemiology*. 2008;23(1):67-74

21 Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R and Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *The European respiratory journal.Supplement*. 1993;16:5-40

22 International Labour Organization, International Standard Classification of Occupations. http://www.ilo.org/public/english/bureau/stat/isco/isco88/index.htm. 1990

23 Agusti A, Calverley PM, Celli B et al. Characterisation of COPD heterogeneity in the ECLIPSE cohort. *Respiratory research*. 2010;11:122-9921-11-122

24 Rabe KF, Hurd S, Anzueto A et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association*. 2007;176(1073-449; 6):532-555

25 Swanney MP, Ruppel G, Enright PL et al. Using the lower limit of normal for the FEV1/FVC ratio reduces the misclassification of airway obstruction. *Thorax*. 2008;63(12):1046-1051

26 Chapman KR, Tashkin DP and Pye DJ. Gender bias in the diagnosis of COPD. *Chest*. 2001;119(6):1691-1695

27 Forey BA, Thornton AJ and Lee PN. Systematic review with meta-analysis of the epidemiological evidence relating smoking to COPD, chronic bronchitis and emphysema. *BMC pulmonary medicine*. 2011;11:36-2466-11-36

Chapter 2

Supplement

Table 1. Most prevalent occupations in subjects exposed.

Exposure		Jobs
Biological dust	Low	Institution- and home-based personal care workers, Domestic and office cleaners, Nursing and midwifery professionals, Nursing associate professionals
	High	Dairy and livestock producers, Carpenters, Freight handlers, Bakers
Mineral dust	Low	Cleaners, Dairy and livestock producers, Machine operators, Heavy truck and lorry drivers
	High	Welders and flame-cutters , Freight handlers, Agricultural and industrial mechanics, Gardeners, (Building) Construction workers, Field crop and vegetable growers
Gases & Fumes	Low	Institution- and home-based personal care workers, Cleaners, Nursing and midwifery professionals
	High	Heavy truck and lorry drivers, Motor vehicle mechanics, Welders and flame-cutters, Agricultural and industrial mechanics, Plumbers and pipe fitters, Painters
Aromatic solvents	Low	Carpenters and joiners, Motor vehicle mechanics, Agricultural and industrial mechanics, Life science technicians, Gardeners and horticultural growers, Plumbers and pipe-fitters
	High	Painters, Printing machine operators, Varneshers and related painters
Chlorinated solvents	Low	Hairdressers and beauticians, Plumbers and pipe fitters, Painters, Mechanical engineers, Decorators and commercial designers, Electronics mechanics and servicers
	High	Motor vehicle mechanics, Agricultural and industrial mechanics sheet metal workers
Heavy metals	Low	Plumbers and pipe fitters, Painters, Machine tool operators, Mechanical engineers, Electronics mechanics and servicers, Building construction laborers, Mechanical engineering technicians, Electrical mechanics, Electronic equipment assemblers
	High	Welders and flame-cutters, Motor vehicle mechanics, Agricultural and industrial mechanics, Sheet metal workers

Table 2. Multivariate logistic regression on association between chronic mucus hypersecretion and occupational exposures, stratified by COPD, and p-values for the interaction between the occupational exposures and COPD.

Exposure		Non-COPD			COPD			Interaction with COPD
		OR (95% CI)	p*	p**	OR (95% CI)	p*	p**	p_{int}
Biological dust	Low	1.19 (0.91-1.56)	0.203	0.622	1.14 (0.74-1.74)	0.560	0.827	0.902
	High	0.86 (0.48-1.55)	0.626		0.86 (0.33-2.25)	0.766		0.801
Mineral dust	Low	**1.39 (1.04-1.87)**	**0.028**	**0.007**	1.18 (0.75-1.85)	0.479	0.684	0.420
	High	**1.60 (1.02-2.52)**	**0.041**		1.01 (0.47-2.17)	0.980		0.190
Gases & fumes	Low	1.13 (0.87-1.47)	0.357	0.001	1.04 (0.70-1.54)	0.858	0.959	0.666
	High	**2.19 (1.49-3.22)**	**< 0.001**		0.98 (0.49-1.94)	0.948		**0.018**
All pesticides	Low	0.88 (0.30-2.60)	0.817	0.835	1.68 (0-44-6.44)	0.451	0.720	0.532
	High	1.81 (0.22-14.9)	0.584		0.00 (0.00-)	0.999		0.784
Aromatic solvents	Low	1.43 (0.99-2.07)	0.057	**0.036**	0.69 (0.36-1.33)	0.269	0.285	**0.038**
	High	1.68 (0.65-4.30)	0.282		0.75 (0.09-6.47)	0.794		0.496
Chlorinated solvents	Low	1.46 (0.96-2.22)	0.075	**0.010**	0.94 (0.45-1.96)	0.862	0.356	0.269
	High	**1.98 (1.00-3.91)**	**0.049**		0.44 (0.10-2.03)	0.292		0.066
Other solvents	Low	0.58 (0.27-1.27)	0.172	0.438	1.10 (0.61-5.23)	0.713	0.454	0.464
	High	1.29 (0.26-6.51)	0.757		1.78 (0.61-5.23)	0.295		0.597
Heavy metals	Low	1.07 (0.65-1.76)	0.795	**0.015**	1.06 (0.50-2.25)	0.881	0.956	0.843
	High	**2.26 (1.30-3.94)**	**0.004**		0.92 (0.33-2.56)	0.876		0.093

*Reference is not exposed; Analysis corrected for gender, BMI, ETS, ex- and current smoking and packyears; Occupational exposures were added one-by-one; bold = p-value < 0.05; * p = p-value for separated (low and high) exposure (no exposure = reference); ** p = p-value for linear trend of intensity of exposure; p_{int} = p-value for interaction analysis.*

Table 3. Age-stratified (< 50 years and > 50 years) multivariate logistic regression on association between chronic mucus hypersecretion and gender, BMI, ETS (by others and at work), ex- and current smoking, packyears, and occupational exposures (added one by one), stratified by COPD.

		< 50 years				≥ 50 years			
		non COPD		COPD		non COPD		COPD	
N		5,621		566		2,908		913	
CMH %		3.3		9.5		4.8		8.2	
		OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p
Gender (male)		**1.54 (1.14-2.08)**	**0.005**	1.16 (0.64-2.12)	0.630	**1.77 (1.18-2.64)**	**0.005**	1.55 (0.93-2.56)	0.091
BMI, kg/m^2		1.03 (1.00-1.06)	0.089	1.04 (0.97-1.11)	0.304	1.05 (1.00-1.10)	0.053	0.98 (0.91-1.04)	0.492
Smoking									
ETS by others*		1.21 (0.84-1.75)	0.302	**2.56 (1.30-5.05)**	**0.006**	1.53 (0.92-2.57)	0.103	1.58 (0.85-2.92)	0.148
ETS at work		**1.66 (1.05-2.65)**	**0.031**	0.65 (0.27-1.55)	0.328	0.84 (0.35-1.99)	0.685	1.56 (0.62-3.92)	0.348
Ex smoking		**0.61 (0.38-0.96)**	**0.033**	0.41 (0.14-1.20)	0.105	1.02 (0.60-1.72)	0.952	0.90 (0.43-1.89)	0.779
Current smoking		1.44 (0.91-2.27)	0.116	0.55 (0.19-1.55)	0.256	1.58 (0.78-3.22)	0.204	0.79 (0.32-1.96)	0.605
Packyears, per 10**		**1.44 (1.14-1.61)**	**0.001**	**1.66 (1.24-2.21)**	**0.001**	1.19 (1.00-1.42)	0.056	**1.21 (1.03-1.42)**	**0.017**
Occupational exposure									
Biological dust	Low	**1.42 (1.02-1.98)**	**0.040**	1.05 (0.54-2.06)	0.885	0.86 (0.54-1.37)	0.525	1.18 (0.67-2.07)	0.560
	High	0.91 (0.43-1.92)	0.812	0.51 (0.06-4.07)	0.527	0.83 (0.32-2.13)	0.694	1.02 (0.34-3.10)	0.969
Mineral dust	Low	**1.64 (1.14-2.36)**	**0.008**	1.30 (0.65-2.60)	0.453	1.05 (0.63-1.76)	0.839	0.93 (0.50-1.76)	0.834
	High	**2.12 (1.25-3.60)**	**0.006**	0.42 (0.09-2.03)	0.278	0.84 (0.32-2.17)	0.714	1.31 (0.54-3.17)	0.548
Gases & fumes	Low	**1.43 (1.03-1.98)**	**0.032**	1.11 (0.58-2.12)	0.746	0.75 (0.48-1.17)	0.207	0.94 (0.56-1.57)	0.809
	High	**2.54 (1.57-4.11)**	**<0.001**	1.98 (0.74-5.34)	0.175	1.83 (0.94-3.55)	0.073	0.55 (0.20-1.53)	0.252
All pesticides	Low	1.05 (0.48-2.32)	0.902	0.00 (0.00-)	0.998	0.96 (0.34-2.71)	0.942	1.36 (0.45-4.12)	0.583
	High	2.33 (0.68-7.96)	0.176	0.00 (0.00-)	0.999	0.00 (0.00-)	0.998	2.13 (0.58-7.78)	0.255
Aromatic Solvents	Low	1.25 (0.78-2.00)	0.355	0.62 (0.22-1.79)	0.378	1.85 (1.02-3.37)	0.043	0.66 (0.28-1.54)	0.341
	High	1.42 (0.42-4.80)	0.569	0.70 (0.06-8.61)	0.784	2.40 (0.54-10.79)	0.252	0.00 (0.00-)	0.999
Chlorinated solvents	Low	1.26 (0.74-2.15)	0.400	0.93 (0.29-3.01)	0.905	**2.01 (1.03-3.94)**	**0.040**	0.87 (0.33-2.33)	0.785
	High	1.85 (0.81-4.18)	0.142	0.76 (0.15-3.93)	0.740	2.55 (0.75-8.74)	0.136	0.00 (0.00-)	0.999
Other solvents	Low	1.00 (0.70-1.44)	0.981	0.89 (0.44-1.82)	0.755	1.15 (0.71-1.86)	0.557	0.77 (0.40-1.48)	0.437
	High	1.25 (0.49-3.16)	0.635	0.56 (0.06-5.70)	0.626	2.13 (0.63-7.17)	0.221	1.13 (0.13-9.54)	0.909
Heavy metals	Low	0.94 (0.50-1.76)	0.848	0.90 (0.27-3.02)	0.871	1.33 (0.59-3.03)	0.491	1.01 (0.37-2.73)	0.981
	High	**2.19 (1.11-4.32)**	**0.023**	1.33 (0.37-4.83)	0.661	2.61 (0.97-7.01)	0.057	0.37 (0.05-2.94)	0.348

*BMI = body mass index; ETS = environmental tobacco smoke; bold = p-value < 0.05; * at least one hour per day during the last year; ** Packyears, per 10: the unit in the analysis is 10 packyears so the OR is the estimate of 10 packyears.*

Table 4. Gender stratified multivariate logistic regression on association between chronic mucus hypersecretion and BMI, ETS (by others and at work), ex- and current smoking, packyears, and occupational exposures (added one by one), stratified by COPD.

						females			
		non COPD		COPD		non COPD		COPD	
N		3,294		692		5,071		787	
CMH %		4.7		10.5		2.5		7.1	
		OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p
BMI, kg/m^2		1.04 (0.99-1.08)	0.101	1.00 (0.93-1.07)	0.897	**1.04 (1.00-1.07)**	**0.038**	1.02 (0.95-1.09)	0.592
Smoking									
ETS by others*		1.35 (0.91-2.01)	0.134	1.43 (0.78-2.60)	0.246	1.20 (0.76-1.88	0.432	3.09 (1.60-6.00)	0.001
ETS at work		1.51 (0.93-2.47)	0.099	1.08 (0.48-2.45)	0.854	1.12 (0.53-2.38)	0.762	1.08 (0.40-2.89)	0.879
Ex smoking		0.82 (0.52-1.29)	0.396	1.00 (0.46-2.17)	0.998	0.76 (0.46-1.26)	0.287	0.46 (0.18-1.16)	0.099
Current smoking		1.25 (0.76-2.06)	0.388	0.95 (0.39-2.29)	0.906	1.84 (1.05-3.24)	0.033	0.58 (0.21-1.60)	0.290
Packyears, per 10**		**1.25 (1.08-1.45)**	**0.003**	**1.18 (1.00-1.38)**	**0.049**	**1.35 (1.10-1.65)**	**0.004**	**1.56 (1.22-2.00)**	**<0.001**
Occupational exposure									
Biological dust	Low	**1.49 (1.01-2.19)**	**0.045**	0.55 (0.26-1.20)	0.135	0.99 (0.69-1.44)	0.970	1.74 (0.98-3.08)	0.057
	High	0.88 (0.46-1.66)	0.686	0.85 (0.32-2.23)	0.736	1.02 (0.24-4.28)	0.976	0.00 (0.00-)	0.999
Mineral dust	Low	1.32 (0.90-1.92)	0.155	1.08 (0.60-1.94)	0.794	1.50 (0.94-2.38)	0.087	1.35 (0.66-2.77)	0.417
	High	1.57 (0.97-2.54)	0.069	0.92 (0.41-2.08)	0.850	1.54 (0.36-6.62)	0.563	2.84 (0.32-25.01)	0.347
Gases & fumes	Low	1.09 (0.75-1.59)	0.639	0.58 (0.32-1.04)	0.068	1.17 (0.81-1.68)	0.405	1.84 (1.03-3.28)	0.041
	High	**2.25 (1.48-3.42)**	**<0.001**	0.85 (0.41-1.75)	0.657	1.38 (0.32-5.95)	0.667	1.88 (0.21-16.43)	0.570
All pesticides	Low	0.89 (0.43-1.86)	0.757	0.95 (0.32-2.78)	0.924	1.69 (0.52-5.52)	0.383	0.00 (0.00-)	0.999
	High	1.31 (0.40-4.31)	0.652	1.75 (0.48-6.31)	0.394	0.00 (0.00-)	0.394	0.00 (0.00-)	0.999
Aromatic Solvents	Low	1.50 (1.02-2.23)	0.042	0.74 (0.38-1.47)	0.396	0.91 (0.28-2.96)	0.880	0.37 (0.04-3.401)	0.383
	High	1.85 (0.72-4.76)	0.204	1.04 (0.12-9.04)	0.970	0.00 (0.00-)	0.999	0.00 (0.00-)	1.000
Chlorinated solvents	Low	1.55 (0.96-2.50)	0.076	1.03 (0.44-2.38)	0.952	1.25 (0.54-2.93)	0.602	0.62 (0.13-3.01)	0.551
	High	**1.99 (1.00-3.94)**	**0.048**	0.60 (0.13-2.72)	0.511	0.00 (0.00-)	0.999	0.00 (0.00-)	0.999
Other solvents	Low	1.38 (0.95-2.00)	0.094	0.65 (0.33-1.28)	0.210	0.75 (0.47-1.17)	0.205	1.07 (0.54-2.11)	0.854
	High	1.86 (0.72-4.81)	0.199	0.90 (0.11-0.759)	0.920	1.12 (0.34-3.61)	0.855	1.10 (0.13-9.45)	0.928
Heavy metals	Low	1.03 (0.61-1.74)	0.919	1.09 (0.50-2.41)	0.825	1.35 (0.31-5.94)	0.691	0.57 (0.05-5.99)	0.638
	High	**2.24 (1.28-3.93)**	**0.005**	1.11 (0.40-3.03)	0.843	0.00 (0.00-)	0.999	0.00 (0.00-)	0.999

*BMI = body mass index; ETS = environmental tobacco smoke; bold = p-value < 0.05; * at least one hour per day during the last year; ** Packyears, per 10: the unit in the analysis is 10 packyears so the OR is the estimate of 10 packyears.*

Table 5. Smoking habits stratified (never, ex and current smokers) multivariate logistic regression on association between chronic mucus hypersecretion and gender, BMI, ETS (by others and at work), packyears, and occupational exposures (added one by one), stratified by COPD.

		Never smokers				Ex smokers				Current smokers			
		non COPD		COPD		non COPD		COPD		non COPD		COPD	
N		3,808		380		3,092		N = 627		N = 1,629		N = 472	
CMH %		2.5		6.1		2.8		7.5		6.8		12.5	
		OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p
Gender(male)		**1.90 (1.26-2.87)**	**0.002**	1.68 (0.72-3.95)	0.231	**2.03 (1.05-3.89)**	**0.034**	1.18 (0.79-1.77)	0.414	0.85 (0.57-1.27)	0.424	0.85 (0.47-1.52)	0.577
BMI, kg/m2		1.04 (1.00-1.09)	0.079	0.93 (0.81-1.06)	0.276	1.03 (0.98-1.09)	0.211	1.00 (0.92-1.08)	0.992	1.04 (0.99-1.08)	0.117	1.04 (0.97-1.12)	0.268
Smoking													
ETS by others*		1.43 (0.81-2.53)	0.223	1.37 (0.36-5.26)	0.648	**1.71 (0.94-3.12)**	**0.07**	1.37 (0.57-3.33)	0.483	0.99 (0.65-1.50)	0.99 (0.65-1.50)	0.962	0.001
ETS at work		1.41 (0.60-3.27)	0.430	1.68 (0.27-10.45)	0.579	1.37 (0.56-3.33)	0.488	1.37 (0.32-5.85)	0.668	1.26 (0.72-2.19)	1.26 (0.72-2.19)	0.422	0.956
Packyears, per 10**						0.94 (0.75-1.20)	0.633	**1.23 (1.02-1.49)**	**0.027**	**1.54 (1.33-1.80)**	**1.54 (1.33-1.80)**	**<0.001**	**0.007**
Occupational exposure													
Biological dust	Low	1.08 (0.67-1.74)	0.752	2.11 (0.83-5.35)	0.116	1.01 (0.60-1.68)	0.981	1.24 (0.62-2.49)	0.547	1.39 (0.90-2.15)	0.139	0.78 (0.40-1.52)	0.468
	High	0.55 (0.17-1.76)	0.311	0.00 (0.00----)	0.999	1.04 (0.37-2.96)	0.937	1.47 (0.41-5.32)	0.557	1.02 (0.40-2.56)	0.972	0.78 (0.17-3.61)	0.747
Mineral dust	Low	1.65 (1.00-2.74)	0.051	1.16 (0.37-3.65)	0.799	1.04 (0.58-1.86)	0.895	0.90 (0.41-1.98)	0.796	1.39 (0.87-2.23)	0.174	1.45 (0.76-2.76)	0.262
	High	2.00 (0.94-4.28)	0.074	0.00 (0.00-)	0.999	0.99 (0.37-2.59)	0.977	0.82 (0.26-2.58)	0.736	1.85 (0.91-3.76)	0.088	1.82 (0.60-5.56)	0.290
Gases & fumes	Low	1.02 (0.64-1.62)	0.942	1.31 (0.53-3.25)	0.557	0.86 (0.53-1.40)	0.547	0.84 (0.44-1.60)	0.598	1.48 (0.95-2.28)	0.080	1.07 (0.58-1.97)	0.825
	High	**3.19 (1.69-6.02)**	<0.001	0.95 (0.10-8.67)	0.965	1.31 (0.60-2.83)	0.497	0.54 (0.15-1.97)	0.350	**2.41 (1.26-4.59)**	**0.008**	1.51 (0.60-3.81)	0.380
All pesticides	Low	1.20 (0.47-3.03)	0.703	0.00 (0.00-)	0.999	0.82 (0.20-3.43)	0.785	0.50 (0.06-3.86)	0.503	0.88 (0.30-2.60)	0.817	1.68 (0.44-6.43)	0.451
	High	0.86 (0.12-6.44)	0.887	0.00 (0.00-)	0.999	1.27 (0.17-9.65)	0.819	2.53 (0.65-9.82)	0.179	1.81 (0.22-14.98)	0.584	0.00 (0.00-)	0.999
Aromatic Solvents	Low	1.47 (0.79-2.75)	0.224	1.23 (0.26-5.85)	0.796	1.23 (0.61-2.50)	0.560	0.96 (0.40-2.34)	0.936	1.54 (0.83-2.84)	0.171	0.31 (0.09-1.11)	0.073
	High	3.38 (0.75-15.21)	0.113	0.00 (0.00-)	1.000	2.09 (0.46-9.44)	0.338	0.00 (0.00-)	0.999	0.67 (0.09-5.15)	0.697	1.08 (0.10-11.68)	0.947
Chlorinated solvents	Low	1.23 (0.56-2.72)	0.608	1.30 (0.16-10.79)	0.807	1.00 (0.43-2.37)	0.993	1.11 (0.36-3.43)	0.850	2.03 (1.09-3.77)	0.025	0.75 (0.25-2.28)	0.612
	High	**4.40 (1.74-11.14)**	**0.002**	4.00 (0.36-44.77)	0.260	0.89 (0.21-3.83)	0.876	0.00 (0.00-)	0.999	1.29 (0.29-5.71)	0.733	0.25 (0.03-2.55)	0.244
Other solvents	Low	1.09 (0.67-1.77)	0.731	1.95 (0.73-5.23)	0.186	0.95 (0.55-1.62)	0.845	0.79 (0.37-1.69)	0.543	1.09 (0.68-1.77)	0.713	0.58 (0.27-1.27)	0.172
	High	1.38 (0.33-5.82)	0.658	0.00 (0.00-)	0.999	1.09 (0.25-4.65)	0.911	0.00 (0.00-)	0.999	1.78 (0.60-5.23)	0.295	1.29 (0.26-6.51)	0.757
Heavy metals	Low	0.93 (0.36-2.38)	0.874	1.17 (0.13-10.41)	0.887	0.76 (0.27-2.15)	0.599	1.13 (0.36-3.51)	0.831	1.40 (0.68-2.89)	0.361	1.04 (0.33-3.26)	0.950
	High	3.41 (1.44-8.06)	0.005	2.03 (0.21-19.62)	0.540	1.48 (0.50-4.36)	0.473	0.61 (0.08-5.00)	0.648	2.02 (0.74-5.52)	0.170	0.88 (0.21-3.63)	0.855

*BMI = body mass index; ETS = environmental tobacco smoke; bold = p-value < 0.05; * at least one hour per day during the lastyear;*
*** Packyears, per 10: the unit in the analysis is 10 packyears so the OR is the estimate of 10 packyears.*

Table 6. Multivariate logistic regression on the association between chronic mucus hypersecretion and gender, BMI, ETS (by others and at work), ex- and current smoking, packyears, and occupational exposures (added one by one), stratified by COPD based on the lower limit of normal (LLN)

		Non COPD based on LLN		**COPD based on LLN**	
N (CMH%)		9,060 (3.7)		948 (9.3)	
		OR (95% CI)	**p**	**OR (95% CI)**	**p**
Gender (male)		**1.63 (1.29-2.08)**	**< 0.001**	1.33 (0.91-1.94)	0.142
BMI, kg/m^2		**1.04 (1.01-1.06)**	**0.010**	1.00 (0.96-1.05)	0.860
ETS					
By others		1.29 (0.96-1.74)	0.088	**2.06 (1.33-3.19)**	**0.001**
At work		1.37 (0.91-2.06)	0.128	0.99 (0.53-1.85)	0.975
Smoking habits					
Ex smoking		0.80 (0.57-1.12)	0.191	0.78 (0.44-1.40)	0.408
Current smoking		1.50 (1.04-2.18)	0.032	0.85 (0.44-1.63)	0.629
Packyears per 10**		**1.28 (1.14-1.44)**	**< 0.001**	**1.28 (1.12-1.46)**	**<0.001**
Occupational exposures					
Biological dust	Low	1.18 (0.90-1.55)	0.219	1.14 (0.74-1.74)	0.555
	High	0.87 (0.48-1.56)	0.629	0.87 (0.33-2.26)	0.770
Mineral dust	Low	**1.39 (1.04-1.87)**	**0.028**	1.18 (0.75-1.85)	0.479
	High	**1.60 (1.02-2.52)**	**0.041**	1.01 (0.47-2.17)	0.980
Gases & fumes	Low	1.13 (0.87-1.47)	0.357	1.04 (0.70-1.54)	0.858
	High	**2.19 (1.49-3.22)**	**< 0.001**	0.98 (0.49-1.94)	0.948
All pesticides	Low	1.02 (0.54-1.91)	0.953	0.80 (0.28-2.30)	0.677
	High	1.15 (0.36-3.74)	0.812	1.64 (0.46-5.83)	0.446
Aromatic solvents	Low	1.43 (0.99-2.07)	0.057	0.69 (0.36-1.33)	0.269
	High	1.68 (0.65-4.30)	0.282	0.75 (0.09-6.47)	0.794
Chlorinated solvents	Low	1.46 (0.96-2.21)	0.075	0.94 (0.45-1.96)	0.862
	High	**1.98 (1.00-3.91)**	**0.049**	0.44 (0.09-2.03)	0.292
Other solvents	Low	1.05 (0.78-1.39)	0.757	0.82 (0.51-1.32)	0.414
	High	1.46 (0.70-3.04)	0.313	0.91 (0.20-4.15)	0.904
Heavy Metals	Low	1.07 (0.65-1.76)	0.795	1.06 (0.50-2.25)	0.881
	High	**2.26 (1.29-3.94)**	**0.004**	0.92 (0.33-2.56)	0.876

*LLN = lower limit of normal; BMI = body mass index; ETS = environmental tobacco smoke; bold = p-value < 0.05; * at least one hour per day during the last year; ** Packyears, per 10: the unit in the analysis is 10 packyears so the OR is the estimate of 10 packyears*

Table 7. Studies on occupational risk for CMH or CB in the general population, published since 2000.

Investigator year country	N	Age (yr)	Male (%)	Exposure according to	Used definition	Exposed to	Exposure effect compared to	CMH OR (95% CI)	CB OR (95% CI)	Adjusted for
Suadicani, 2001, Denmark [1]	3,331	53-75	100	Self-reported	**CB:** cough & phlegm ≥3 months/yr ≥2 successive yrs	Dust	≥ 5 years exp/ ≤ 5 years exp		**1.5 (1.1-2.1)**	smoking, age, alcohol, blood pressure, MNS phenotype
						Organic solvents	≥ 5 years exp/ ≤ 5 years exp		**1.5 (1.1-2.1)**	
Lange, 2003, Denmark [2]	3.736	≥ 65	40	Self-reported to be ever exp	**CMH:** phlegm ≥3 months /yr ≥2 successive yrs	Industrial dust + fumes	Exp/non exp	**2.2 (1.7-2.7)**		smoking, gender, childhood resp. infections, alcohol.
De Meer, 2004, the Netherlands [3]	1,906	45-70	52	JEM	**CB:** morning phlegm or cough ≥3 months ≥1 yr	Organic dust	Exp/non exp		0.89 (0.56-1.42)	living area, age, gender, smoking, working years
						Mineral dust	Exp/non exp		**2.22 (1.16-4.23)**	
						Gases & fumes	Exp/non exp		0.67 (0.36-1.26)	
Sunyer, 2005, Spain [4]	1,735	20-44	48	JEM (Current job or job quitted because of respiratory symptoms)	**CMH:** phlegm ≥3 months, last yr	Biological dust	Low -/non exp High -/non exp	1.1 (0.7-1.5) **2.0 (1.1-3.8)**		age, gender, centre, other exposure, smoking, (asthma excluded)
						Mineral dust	Low -/non exp High -/non exp	1.1 (0.6-1.8) 0.9 (0.6-1.8)		
						Gases & fumes	Low -/non exp High -/non exp	1.0(0.7-1.4) 1.5 (0.9-2.5)		

Investigator year country	N	Age (yr)	Male (%)	Exposure according to	Used definition	Exposed to	Exposure effect compared to		CMH OR (95% CI)	CB OR (95% CI)	Adjusted for
Matheson 2005, Australia [5]	1,213	45-70	52	ALOHA JEM (Lifetime)	**CMH:** phlegm ≥3 months/yr ≥2 successive yrs	Biological dust	Exp/non exp		1.74 (0.97-2.06)		gender, age, smoking
						Mineral dust	Exp/non exp		1.32 (0.71-2.44)		
						Gases & fumes	Exp/non exp		1.31 (0.72-2.40)		
					***CMH + COPD:** phlegm ≥3 months/yr ≥2 successive yrs + FEV1/FVC<0.70	Biological dust	Exp/non exp		**3.19 (1.27-7.97)***		
							Exp/non exp	Men	2.02 (0.17-2.08)*		
							Exp/non exp	Women	**5.83 (1.24-27.4)***		
						Mineral dust	Exp/non exp		1.40 (0.56-3.51)*		
							Exp/non exp	Men	0.59 (0.17-2.08)*		
							Exp/non exp	Women	**3.60 (1.06-12.3)***		
						Gases & fumes	Exp/non exp		**2.81 (1.01-7.79)***		
							Exp/non exp	Men	1.62 (0.42-6.26)*		
							Exp/non exp	Women	**4.85 (1.03-22.9)***		
Jaen, 2006, Spain [6]	576	20-70	49	Jobs self-reported exposure >1 yr. (Lifetime)	**CMH:** phlegm ≥3 months/yr ≥2 successive yrs	Dust, gases or fumes	Exp/non exp		**2.0 (1.1-3.7)**		gender, age, smoking
							Never smoking	Exp/non exp	3.7 (0.9-15)		
							Ex smoking	Exp/non exp	1.9 (0.5-7.2)		
							Current smoking	Exp/non exp	1.6 (0.7-3.6)		
Zock, 2006, The Nether-lands [7]	13,253	20-44	48	JEM (Current job or quitted because of respiratory symptoms)	**CMH:** phlegm ≥3 months/ last yr	VGDF	Never smoking	Low-/non exp	1.0 (0.7-1.4)	0.8 (0.4-1.5)	age, gender, centre, height, current nr. of cigarettes (asthma excluded)
								High-/non exp	1.4 (0.9-2.1)	1.0 (0.4-2.1)	
					CB: cough + phlegm ≥3 months/last yr		Ex smoking	Low-/non exp	1.1 (0.6-1.8)	0.7 (0.3-1.8)	
								High-/non exp	1.1 (0.6-2.2)	1.0 (0.4-2.9)	
							Current smoking	Low-/non exp	1.2 (1.0-1.5)	1.3 (0.9-1.8)	
								High-/non exp	**1.6 (1.2-2.1)**	1.7 (1.2-2.4)	

Investigator year country	N	Age (yr)	Male (%)	Exposure according to	Used definition	Exposed to	Exposure effect compared to		CMH OR (95% CI)	CB OR (95% CI)	Adjusted for
LeVan, 2006, Singapore [8]	52,325	45-74	43	Industry or occupation > 1 yr been working	**CMH:** phlegm ≥3 months/yr ≥2 successive yrs	Dust from mineral, cotton, wood, metal or asbestos	Exp/non exp		1.08 (0.95-1.22)	**1.26 (1.01-1.57)**	age, gender, region, smoking, age start smoking, cigarettes/day
							Exp/non exp		1.03 (0.89-1.18)	0.93 (0.71-1.22)	
					CB: cough & phlegm ≥3 months/yr ≥ 2 successive yrs	Vapors from chemical solvents	Exp/non exp			**2.22 (1.16-4.23)**	
						Mineral dust	Current smoking	Exp/non exp		**3.97 (1.08-14.5)**	
						Gases & fumes	Exp/non exp			0.67 (0.36-1.26)	
Krstev, 2008, China [9]	14,500	40-70	0	Every job > 1yr, categorized	**CB:** ever diagnosed with bronchitis by a physician	#Textile/sewing,	Ever employed/ never employed.			**1.11 (1.03-1.19)**	smoking, education, income, asthma
						Postal/ communication,				**1.57 (1.14-2.17)**	
						Education/culture				**1.14 (1.05-1.25)**	
Skorge, 2009, Norway [10]	2,312	15-65	49	ALOHA JEM (last 11 years)	**CMH:** phlegm when coughing	Biological dust	Men	Low-/non exp	1.37 (0.9-2.1)		age, education level, smoking, occupational exposure >11 yrs ago.
								High-/non exp	0.89 (0.4-2.0)		
							Women	Low-/non exp	0.87 (0.6-1.3)		
								High-/non exp	1.57 (0.6-4.2)		
						Mineral dust	Men	Low-/non exp	0.64 (0.4-1.0)		
								High-/non exp	0.55 (0.3-1.0)		
							Women	Low-/non exp	0.53 (0.3-0.9)		
								High-/non exp	1.68 (0.7-4.1)		
						Gases & fumes	Men	Low-/non exp	1.01 (0.6-1.6)		
								High-/non exp	0.67 (0.4-1.1)		
							Women	Low-/non exp	1.42 (1.0-2.1)		
								High-/non exp	2.03 (0.9-4.8)		

CMH = chronic mucus hypersecretion; CB = chronic bronchitis; JEM = job exposure matrix; VGDF = vapors, gases, dusts and fume

2

References

1. Suadicani P, Hein HO, Meyer HW and Gyntelberg F. Exposure to cold and draught, alcohol consumption, and the NS-phenotype are associated with chronic bronchitis: an epidemiological investigation of 3387 men aged 53-75 years: the Copenhagen Male Study. *Occupational and environmental medicine*. 2001;58(3):160-164

2 Lange P, Parner J, Prescott E and Vestbo J. Chronic bronchitis in an elderly population. *Age and Ageing*. 2003;32(0002-0729; 6):636-642

3 de Meer G, Kerkhof M, Kromhout H, Schouten JP and Heederik D. Interaction of atopy and smoking on respiratory effects of occupational dust exposure: a general population-based study. *Environmental health : a global access science source*. 2004;3(1):6

4 Sunyer J, Zock JP, Kromhout H et al. Lung function decline, chronic bronchitis, and occupational exposures in young adults. *American journal of respiratory and critical care medicine*. 2005;172(9):1139-1145

5 Matheson MC, Benke G, Raven J et al. Biological dust exposure in the workplace is a risk factor for chronic obstructive pulmonary disease. *Thorax*. 2005;60(8):645-651

6 Jaen A, Zock JP, Kogevinas M, Ferrer A and Marin A. Occupation, smoking, and chronic obstructive respiratory disorders: a cross sectional study in an industrial area of Catalonia, Spain. *Environmental health : a global access science source*. 2006;5:2

7 Zock JP, Sunyer J, Kogevinas M, Kromhout H, Burney P and Anto JM. Occupation, chronic bronchitis, and lung function in young adults. An international study. *American journal of respiratory and critical care medicine*. 2001;163(7):1572-1577

8 LeVan TD, Koh WP, Lee HP, Koh D, Yu MC and London SJ. Vapor, dust, and smoke exposure in relation to adult-onset asthma and chronic respiratory symptoms: the Singapore Chinese Health Study. *American Journal of Epidemiology*. 2006;163(12):1118-1128

9 Krstev S, Ji BT, Shu XO et al. Occupation and chronic bronchitis among Chinese women. *Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine*. 2008;50(1):64-71

10 Skorge TD, Eagan TM, Eide GE, Gulsvik A and Bakke PS. Occupational exposure and incidence of respiratory disorders in a general population. *Scandinavian journal of work, environment & health*. 2009;35(6):454-461

Chapter 3

Susceptibility to chronic mucus hypersecretion, a genome wide association study

Genetic influence of *SATB1* on airway disease

AE Dijkstra, J Smolonska, M van den Berge, C Wijmenga, P Zanen, MA Luinge, M Platteel, JWJ Lammers, M Dahlback, K Tosh, PS Hiemstra, PJ Sterk, A Spira, J Vestbo, BG Nordestgaard, M Benn, SF Nielsen, M Dahl, WM Verschuren, HSJ Picavet, HA Smit, M Owsijewitsch, HU Kauczor, HJ de Koning, E Nizankowska-Mogilnicka, F Mejza, P Nastalek, CC van Diemen, MH Cho, EK Silverman, JD Crapo, TH Beaty, DA Lomas, P Bakke, A Gulsvik, Y Bossé, M Obeidat, DW Loth, L Lahousse, F Rivadeneira, AG Uitterlinden, A Hofman, BH Stricker, GG Brusselle, CM van Duijn,U Brouwer, GH Koppelman, JM Vonk, MC Nawijn, HJM Groen, W Timens, HM Boezen, DS Postma

PLOS ONE, 2014 April

Abstract

Background

Chronic mucus hypersecretion (CMH) is associated with an increased frequency of respiratory infections, excess lung function decline, and increased hospitalisation and mortality rates in the general population. It is associated with smoking, but it is unknown why only a minority of smokers develops CMH. A plausible explanation for this phenomenon is a predisposing genetic constitution. Therefore, we performed a genome wide association (GWA) study of CMH in Caucasian populations.

Methods

GWA analysis was performed in the NELSON-study using the Illumina 610 array, followed by replication and meta-analysis in 11 additional cohorts. In total 2,704 subjects with, and 7,624 subjects without CMH were included, all current or former heavy smokers (≥ 20 packyears). Additional studies were performed to test the functional relevance of the most significant single nucleotide polymorphism (SNP).

Results

A strong association with CMH, consistent across all cohorts, was observed with rs6577641 ($p = 4.25 \times 10^{-6}$, OR = 1.17), located in intron 9 of the *special AT-rich sequence-binding protein 1 locus (SATB1)* on chromosome 3. The risk allele (G) was associated with higher mRNA expression of *SATB1* (4.3×10^{-9}) in lung tissue. Presence of CMH was associated with increased *SATB1* mRNA expression in bronchial biopsies from COPD patients. *SATB1* expression was induced during differentiation of primary human bronchial epithelial cells in culture.

Conclusions

Our findings, that SNP rs6577641 is associated with CMH in multiple cohorts and is a *cis*-eQTL for *SATB1,* together with our additional observation that *SATB1* expression increases during epithelial differentiation provide suggestive evidence that *SATB1* is a gene that affects CMH.

Introduction

The secretion of mucus is a natural part of the airway defense against inhaled noxious particles and substances. Chronic mucus hypersecretion (CMH) is a condition of overproduction of mucus and defined as the presence of sputum production during at least three months in two consecutive years without any explaining origin whereas airway obstruction is not a prerequisite [1]. Smoking is a risk factor for CMH, i.e. the prevalence of CMH in the general population is reported to be 7.4% in current smokers, 3.7% in ex-smokers and 2.4% in never smokers [2]. CMH is the key presenting symptom in chronic bronchitis, one of the three main sub-groups of chronic obstructive pulmonary disease (COPD), a complex disease characterized by the presence of incompletely reversible and generally progressive airflow limitation [3]. Moreover, CMH is a risk factor for the development of COPD [4, 5].

Worldwide, COPD affected 65 million people in 2004 and more than 3 million people died of COPD in 2005, representing 5% of all deaths. It is predicted that COPD will be the third leading cause of death worldwide in 2030 [6]. COPD markedly reduces quality of life and is responsible for high healthcare costs. For instance, the combined (direct and indirect) yearly costs of COPD and asthma in the United States of America were projected at $68 billion in 2008 [7]. CMH is not only associated with COPD but also with an increased duration and frequency of respiratory infections, excess decline in forced expiratory volume in 1 second (FEV_1) and increased hospitalization and mortality rates in the general population [4, 5, 8, 9].

It is not known why only a minority of all smokers develops CMH, yet a plausible explanation is the presence of a genetic predisposition for CMH, as evidenced by familial aggregation of mucus overproduction and higher prevalence of CMH in monozygotic than in dizygotic twins [10-12]. Little is known about the identity of the genes that predispose to CMH. One publication suggested that *CTLA4* is associated with chronic bronchitis in COPD [13].

The aim of our study was to identify genetic factors for CMH, thereby obtaining a better insight into the origins of this disorder.

Materials & Methods

Ethics Statement

The Dutch ministry of health and the Medical Ethics Committee of the hospital approved the study protocol for all Dutch centers. Ethics approval and written informed consent was obtained from all participants in all studies participating. For detailed information, see Supplement.

Subjects and genotyping

We performed GWA studies in participants of the NELSON-study (n=3,729), a male population-based lung cancer screening study investigating heavy smokers (≥ 20 packyears) [14]. Replication of SNPs with $p \leq 10^{-4}$ was attempted in six cohorts participating in 'COPD Pathology: Addressing Critical gaps, Early Treatment & diagnosis and Innovative Concepts' (COPACETIC) and in five non-COPACETIC cohorts. Caucasian subjects with ≥ 20 packyears smoking with genotype-, spirometric- and demographic data were included. An overview of the CMH definitions used in this study is presented in Table 1.

Table 1. Questions used to define chronic mucus hypersecretion in the corresponding cohorts.

Cohort	Question
NELSON[14]	Do you expectorate sputum on the majority of days more than 3 months a year, even when you do not have a cold?
Rotterdam[18, 27]	Do you expectorate sputum on the majority of days during ≥ 3 months during the last 2 years?
LifeLines[28]	Do you usually expectorate sputum during day or night in winter? If yes: Do you expectorate sputum on the majority of days > 3 months a year?
Vlagtwedde- Vlaardingen[29, 30]	Do you expectorate sputum on the majority of days > 3 months a year?
Doetinchem[31]	Do you expectorate sputum during winter, day and night, each day for 3 months?
Poland[32, 33]	Do you usually bring up phlegm from your chest, or do you usually have phlegm in your chest that is difficult to bring up when you don't have a cold? If yes: Are there months in which you have this phlegm on most days? If yes: Do you bring up this phlegm on most days for as much as three months each year? A positive answer to all (3) questions identifies CMH.
Heidelberg[34]	Do you expectorate sputum on the majority of days > 3 months a year?
GLUCOLD[17]	Do you expectorate sputum immediately after getting up on the majority of days in winter > 3 months a year?
Rucphen[30]	Do you expectorate sputum during day or night in winter? If yes: Do you have expectoration on the majority of days > 3 months a year?
ECLIPSE[35]	Do you usually bring up phlegm from your chest on getting up, first thing in the morning, during the rest of the day or at night, on most days for 3 consecutive months or more during the year?
COPDGene[36]	Do you usually bring up phlegm from your chest on getting up, first thing in the morning, during the rest of the day or at night, on most days for 3 consecutive months or more during the year?
Norway[37, 38]	Do you usually bring up phlegm from your chest on getting up, first thing in the morning, during the rest of the day or at night, on most days for 3 consecutive months or more during the year?

A brief description of the included cohorts and details according to the period of data collection, type of population, genotyping platforms and genetic imputation software are presented in Table 2.

Table 2. Overview of populations. Populations and corresponding period of data collection, type of population, genotyping platform and soft-ware used for imputation.

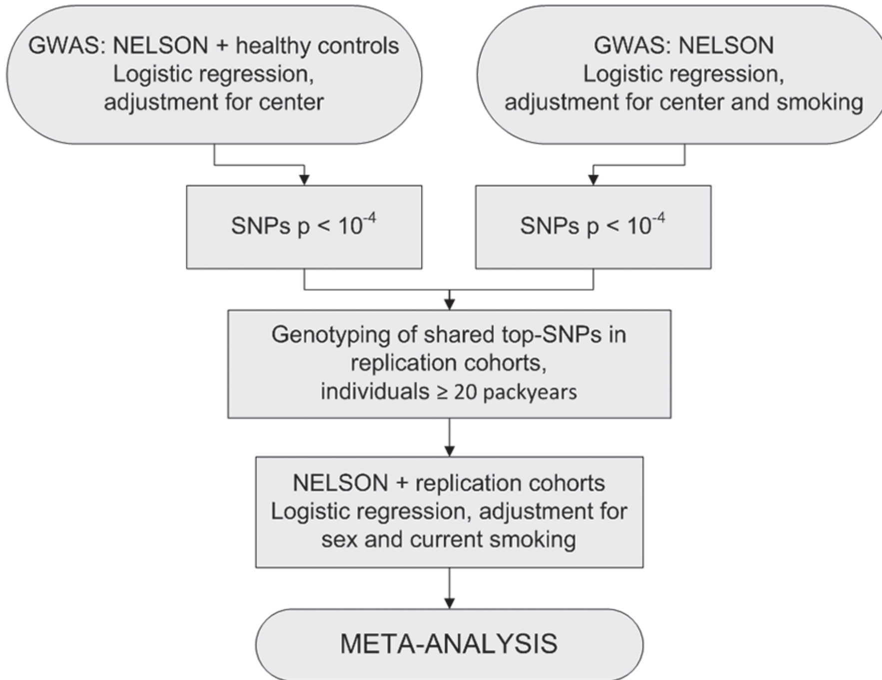
Study	Data Collection	Type of population	Genotyping platform	Imputation software
NELSON	2004/2005	general population	Illumina Quad 610	NA
GLUCOLD	2005/2006	COPD case	Illumina Veracode	NA
Vlagtwedde Vlaardingen	1989/1990	general population	Illumina Veracode	NA
Doetinchem	1998/2002	general population	Illumina Veracode	NA
Poland	2005/2006	general population	Illumina Veracode	NA
Heidelberg	2004/2005	general population	Illumina Veracode	NA
Rucphen	2002	Family based COPD on a doctor diagnosis	Illumina Veracode	NA
Rotterdam	2002/2008	general population	Illumina 550K	MaCH
LifeLines	2008/2010	general population	Illumina Human CytoSNP-12	BEAGLE v3.1.0
COPDGene	2008/2009	COPD case/control (stage I-IV)	Illumina Human Omni1-Quad	MaCH
ECLIPSE	2005/2007	COPD case/control (stage II–IV)	Illumina Human HAP 550 V3	MaCH
Norway	2003/2005	COPD case/control (stage II–IV)	Illumina Human HAP 550 V1, V3, and DUO	MaCH

NA = not applicable

Strategy

We searched for SNPs associated with CMH by using a two-stage strategy followed by a replication stage and meta-analysis (Figure 1).

Figure 1. Study design



We performed GWA studies in the NELSON cohort and in additional healthy controls. CMH was analyzed using logistic regression with adjustment for center (Groningen and Utrecht). Since current smoking can affect the presence of CMH, we additionally performed the GWAS in the NELSON cohort correcting for center and smoking. SNPs with a p-value < 10^{-4} present in both GWA studies were selected for replication. To test for generalizability of associations with CMH in other populations, we compared our results with data in CMH-cases and controls with a smoking history of ≥ 20 packyears with eleven replication populations using logistic regression with adjustment for sex and current smoking. Finally, we performed a meta-analysis on shared SNPs across the NELSON identification population and the 11 replication populations.

Statistical analysis

General characteristics of CMH-cases and controls were compared using Student's t- and Mann-Whitney-U tests for continuous variables and χ^2 tests for dichotomous variables with SPSS 20.0. Sample and SNP quality control (QC), regression- and meta-analysis were performed with PLINK 1.07 [15]. QC criteria are described in the supplement. Logistic regression analysis under an additive model was used to identify SNPs associated with CMH. SNPs with a p-value < 10^{-4} were included for replication. When two SNPs were in strong linkage disequilibrium ($r^2 \geq 0.8$), the SNP with the lowest p-value was further analyzed.

SNPs in COPACETIC cohorts and in LifeLines were analyzed using logistic regression with adjustment for sex and smoking (ex-/current smoking). In LifeLines, imputed SNPs with an info-score < 0.3 (imputation quality score) were removed. SNPs in non-COPACETIC cohorts were analyzed by the cohort investigators using the same model.

Meta-analysis was performed on SNPs across NELSON and the 11 replication cohorts. The Cochran's Q test was used to test for heterogeneity in the meta-analysis.

We performed multivariate logistic regression analysis, adjusted for packyears and lung function, to associate CMH with the risk allele of rs6577641 in the identification cohort.

Functional relevance of *SATB1* and rs6577641, our highest ranked-SNP

We performed 4 functional studies with the identified top-SNP. Details on their methods are given in the supplement.

We assessed:

1) whether rs6577641 is an eQTL, by analyzing the association of SATB1 expression levels with rs6577641 genotypes in lung tissue from three independent cohorts recruited from Laval University, University of British Columbia, and University of Groningen as described previously [16];
2) CMH-associated mRNA expression in airway wall biopsies from 77 COPD participants in the GLUCOLD-study [17];
3) the association of homozygous genotypes for rs6577641 with a) immune-histochemical staining (IHC) for SATB1 and b) the fraction of mucus positivity on bronchial tissue explanted from COPD or lung cancer subjects that underwent lung surgery;
4) SATB1 expression levels during mucociliary differentiation of primary bronchial epithelial cells cultured at air-liquid interface [18].

Results

Populations

Characteristics of the identification and replication populations are presented in Table 3. Subjects with CMH were more often current smokers and had worse lung function, except for populations including subjects with COPD only.

Table 3. Demographic and clinical characteristics of CMH-cases and -controls with ≥ 20 packyears, present in the meta-analysis.

Population	CMH	N	Population %	Female, %	Age, yrs (SD)	Packyears (range)	Current smoking, %	FEV1 %, pred. (SD)	FEV_1/FVC, % (SD)
NELSON	Control	1,795	71.5	0	60.2 (5.3)	34 (21-156)	47.5	100.3 (17.2)	72.9 (8.7)
	Case	717	28.5	0	60.4 (5.6)	39 (21-140)	74.2	93.5 (20.0)	69.2 (11.0)
Rotterdam	Control	1,043	84.1	46.1	68.0 (9.3)	45 (20-149)	40.1	92.4 (23.5)#	72.8 (8.7)#
	Case	197	15.9	43.7	72.0 (8.4)	40 (20-168)	45.2	85.0 (26.9)#	68.0 (11.1)#
LifeLines	Control	1,431	88.1	80.1	52.9 (9.2)	27 (20-100)	56.4	98.2 (15.6)	72.4 (8.2)
	Case	193	11.5	46.9	53.2 (9.9)	29 (20-97)	75.4	90.5 (18.0)	68.3 (11.3)
Vlagtwedde-Vlaardingen*	Control	234	82.4	27.4	52.9 (10.1)	29 (20-128)	51.7	94.5 (12.1)	76.6 (4.5)
	Case	50	17.6	18	53.4 (10.5)	33 (22-83)	68	86.7 (18.6)	71.0 (8.9)
Doetinchem	Control	250	80.6	37.2	54.7 (8.8)	30 (20-90)	55.6	94.8 (17.6)	71.5 (9.9)
	Case	60	19.4	36.7	56.4 (7.7)	33 (20-72)	68.3	89.1 (19.6)	69.3 (11.4)
Poland	Control	97	85.1	22.7	56.7 (10.5)	30 (20-116)	52.6	96.4 (21.4)	72.5 (0.5)
	Case	17	14.9	11.8	55.8 (9.4)	35 (22-86)	82.4	93.5 (24.0)	69.2 (13.1)
Heidelberg	Control	608	84.2	35.7	58.1 (5.2)	37 (23-138)	54.3	96.4 (17.6)	78.9 (9.7)
	Case	114	15.8	29.8	58.0 (5.2)	37 (23-91)	91.2	86.2 (21.5)	75.3 (10.6)
GLUCOLD**	Control	48	55.2	8.3	62.6 (7.6)	46 (21-182)	62.5	63.4 (9.8)	50.4 (9.1)
	Case	39	44.8	20.5	59.6 (7.4)	40 (22-83)	61.5	63.9 (8.8)	53.1 (7.8)
Rucphen**	Control	28	53.8	46.4	66.5 (7.9)	42 (21-120)	57.1	74.5 (15.7)	57.2 (7.8)
	Case	24	46.2	41.7	62.2 (10.5)	43 (21-100)	70.8	70.2 (21.6)	53.1 (9.7)
ECLIPSE**	Control	961	62	37.5	64.1 (6.7)	53 (21-205)	28.1	48.0 (15.7)	44.5 (11.3)
	Case	590	38	24.1	62.9 (7.4)	54 (22-220)	47.5	46.2 (15.5)	44.3 (11.7)
COPDGene	Control	628	71.8	53.5	63.1 (8.6)	50 (21-173)	28.2	75.0 (28.3)	63.7 (17.6)
	Case	247	28.2	40.5	61.9 (8.4)	54 (21-237)	50.2	60.4 (27.4)	54.6 (17.9)
Norway	Control	501	52.4	44.9	61.5 (10.3)	34 (20-130)	46.9	71.7 (24.2)	64.6 (15.7)
	Case	456	47.6	20.4	64.1 (10.1)	39 (20-119)	59	56.5 (24.4)	55.0 (17.3)

*CMH = Chronic mucus hypersecretion; * lung function is based on FEV1/IVC; ** all individuals in this cohort have COPD; # based on lung function of 700 subjects who returned for follow-up study 4 years later.*

Identification analysis

After QC, 492,700 SNPs and 2,512 individuals (717 CMH cases, 1,795 controls) from the NELSON study remained. Logistic regression analysis was performed including these individuals supplemented with 590 additional healthy controls, adjusting for center. The QQ-plot provided no evidence of population stratification ($\lambda = 1.019$). 77 SNPs were associated with CMH with a p-value $< 10^{-4}$. CMH was associated with current smoking in our identification cohort ($p < 0.001$). Therefore, we performed a second GWA adjusting for center and current/ex-smoking (717 CMH-cases, 1,795 controls). The QQ-plot showed no evidence of population stratification ($\lambda = 1.0056$).

We observed 64 SNPs with a p-value $< 10^{-4}$. Genome wide association for CMH ordered by chromosome is shown in the Manhattan plot. Figure 2 shows QQ-plots (A, C) and genome wide association signals for CMH ordered by chromosome (Manhattan-plots, B and D) of these sequential analyses. We identified 36 SNPs associated with CMH with a p-value $< 10^{-4}$ in both analyses (Table 4). Of these, 32 SNPs were included for replication and 4 SNPs were removed because they were in strong linkage disequilibrium ($r^2 > 0.8$) with another associated SNP.

Replication of associated SNPs

Genotyping of SNP rs4775569 failed in the COPACETIC populations, and was removed for further analysis. CMH-associated top-SNPs for all cohorts are presented in Table 5, with a complete overview in Table 1 in the supplement.

When applying Bonferroni correction in the meta-analysis ($p = 1.61 \times 10^{-3}$ for 31 SNPs), we found a strong association with one SNP:

- rs6577641, a SNP located on chromosome 3 in intron 9 of the *special AT-rich sequence-binding protein 1 locus* (*SATB1*) (combined p-value = 4.25×10^{-6}, OR = 1.17; 1.10-1.26).

The *SATB1* SNP rs6577641 had the lowest p-value for association with CMH in the meta-analysis. Figure 3 shows the forest plot of rs6577641 in the identification and replication cohorts and meta-analysis.

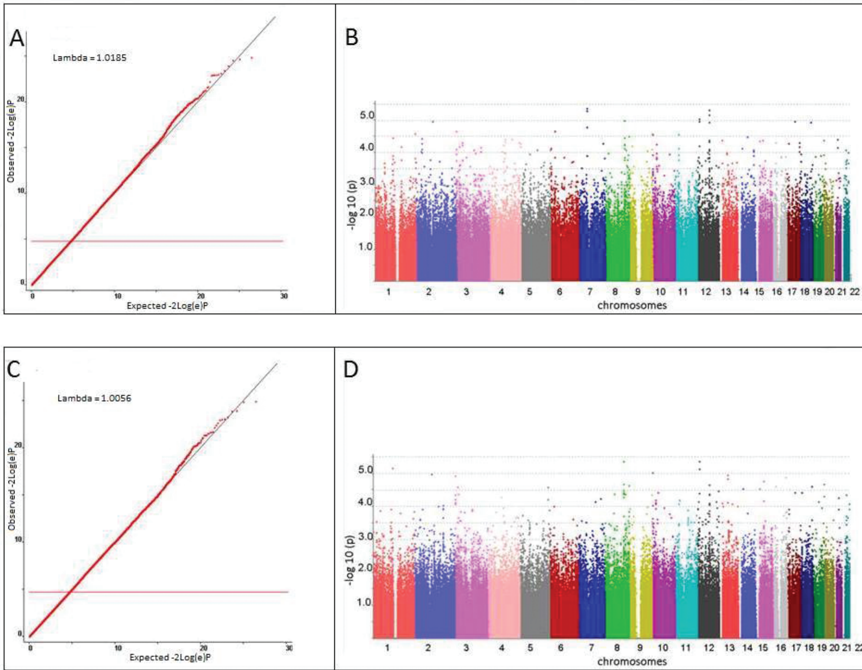


Figure 2. Quantile-quantile plot and Manhattan plot of GWA results for association of SNPs with CMH in NELSON.

A and B, amplified with bloodbank controls and corrected for center. C and D, corrected for center and smoking habits

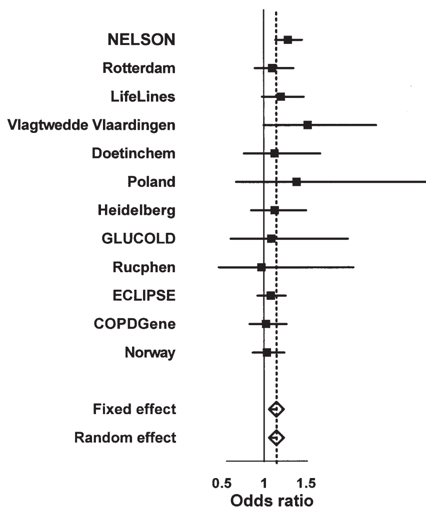


Figure 3. Forest plot showing evidence of association for rs6577641 with chronic mucus hypersecretion in the identification and replication cohorts.

Vertically left, the identification cohort and the replication cohorts included in the meta-analysis. The boxes represent the precision and the horizontal lines represent the confidence intervals. The squares represent the pooled effect estimate from the meta-analysis of all cohorts. The horizontal axis shows the scale of the effects.

Table 4. SNPs associated with CMH with a p-value < 10^{-4}, present in GWAS-I and in GWAS-II, in the NELSON identification cohort.

Chromosome	SNP	Base pair position	p-value GWAS I	p-value GWAS II
2	rs6735868	103582093	1.11×10^{-05}	1.08×10^{-05}
3	rs1387089	1940922	7.94×10^{-05}	4.56×10^{-05}
3	rs1488757	1981567	2.17×10^{-05}	1.16×10^{-05}
3	rs6577641	18397849	6.83×10^{-05}	2.57×10^{-05}
4	rs4306981	79924121	9.74×10^{-05}	5.18×10^{-05}
8	rs4242562	115475287	7.66×10^{-05}	5.13×10^{-05}
8	rs7836298	115504434	1.03×10^{-05}	4.37×10^{-06}
8	rs7823554*	115553109	6.05×10^{-05}	5.22×10^{-05}
8	rs7836963*	115568426	5.52×10^{-05}	4.24×10^{-05}
8	rs16886291	115711436	3.54×10^{-05}	2.09×10^{-05}
8	rs10098746	125838127	8.47×10^{-05}	4.34×10^{-05}
8	rs7831595	144974963	3.08×10^{-05}	2.32×10^{-05}
9	rs4842047	138816796	2.63×10^{-05}	4.51×10^{-05}
10	rs943189	22842590	5.57×10^{-05}	6.33×10^{-05}
11	rs11026531	22379184	2.76×10^{-05}	8.55×10^{-05}
12	rs1894307*	12005720	9.04×10^{-06}	7.18×10^{-06}
12	rs2255953	12010736	1.13×10^{-05}	4.33×10^{-06}
12	rs2855708	12013572	6.47×10^{-05}	3.97×10^{-05}
12	rs10879509*	73242131	6.98×10^{-06}	4.44×10^{-05}
12	rs4760851	73284781	4.85×10^{-06}	2.29×10^{-05}
12	rs952394	73441110	4.18×10^{-05}	4.22×10^{-05}
12	rs12822199	75458164	4.82×10^{-05}	8.58×10^{-05}
12	rs1379963	75493882	1.18×10^{-05}	2.20×10^{-05}
12	rs1795669	76273692	8.01×10^{-05}	7.86×10^{-05}
13	rs9578362	21882381	4.28×10^{-05}	7.99×10^{-05}
13	rs1211304	50381016	9.96×10^{-05}	1.12×10^{-05}
14	rs992745	27810095	7.67×10^{-05}	2.99×10^{-05}
15	rs754661	26934277	4.54×10^{-05}	2.88×10^{-05}
15	rs4775569	46850317	4.20×10^{-05}	1.72×10^{-05}
16	rs13333521	19904082	5.08×10^{-05}	2.50×10^{-05}
17	rs11652469	49565797	1.13×10^{-05}	3.80×10^{-05}
18	rs8086262	69227590	1.15×10^{-05}	2.53×10^{-05}
20	rs4815628	3891896	4.17×10^{-05}	2.15×10^{-05}
21	rs2032257	27774870	3.97×10^{-05}	5.39×10^{-05}
22	rs1009147	30088257	8.41×10^{-05}	4.51×10^{-05}
22	rs1005239	47687170	9.86×10^{-05}	8.67×10^{-05}

**SNP not selected for replication because of strong linkage disequilibrium with another SNP*

We assessed the percentage of subjects with CMH in each genotyping group for rs6577641 in NELSON-total and stratified for current and ex smokers (Figure 4). Multivariate logistic regression analysis, corrected for packyears and FEV_1%predicted, showed that CMH was significantly associated with the number of G-alleles in the 1,385 current smokers (reference = AA: heterozygous mutant (AG) p = 0.001; OR = 1.50, homozygous mutant (GG) p = 0.001; OR = 1.80) but not in 1,127 ex-smokers (reference = AA: heterozygous mutant (AG) p = 0.380; OR = 1.18, homozygous mutant (GG) p = 0.143; OR = 1.42).

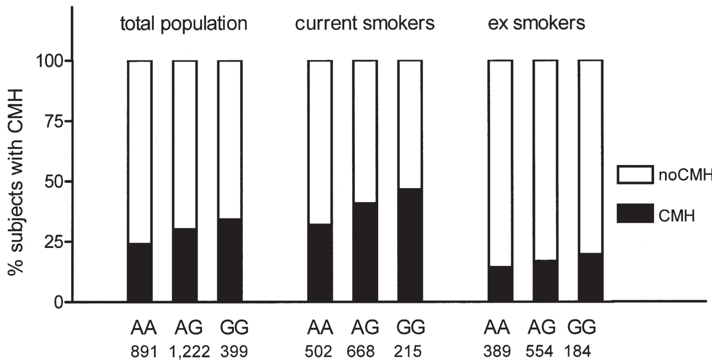


Figure 4. Percentage of subjects with chronic mucus hypersecretion (CMH) within genotypes (AA, AG and GG) of rs6577641 in the identification cohort (NELSON), and distributed among ex- and current smokers.

Table 5. Meta-analysis of top SNPs associated with CMH in replication cohorts, in identification and replication cohorts and corresponding direction of effect in all cohorts and associated feature and gene(s).

Chr	SNP	Minor allele	MAF	Meta-analysis replication cohorts		Meta-analysis identification and replication cohorts				Closest gene(s)
				p-value	OR	Dir. of effect	p-value	OR	Q	
3	rs6577641	G	0.40	5.01E-03	1.12	++++++++0+0+	4.25 E-06	1.17	6.20E-01	*SATB1*
3	rs1488757	G	0.11	2.34E-01	0.92	-00+--+++--0	1.10E-03	0.83	1.55E-01	*LOC727810 & CNTN4*
12	rs2855708	G	0.27	2.18E-01	1.06	+0+0--+-+++0	1.20E-03	1.13	1.76E-01	*ETV6*
14	rs992745	G	0.23	2.94E-01	0.95	--+----+++--	2.74E-03	0.89	4.59E-02	*LOC7288755*
4	rs4306981	G	0.31	3.37E-01	1.04	++-0-+++-0-+	1.38E-03	1.12	5.19E-02	*PAQR3 & ARD1B*
12	rs1795669	A	0.06	2.83E-01	1.09	+++++----+++	2.90E-03	1.22	1.77E-01	*LOC100130336 & LOC100131830*
9	rs4842047	A	0.30	3.88E-01	0.96	-0-XX+-X0-00	3.44E-03	0.89	3.03E-01	*CAMPSAP1 & UBAC1*
13	rs95788362	A	0.40	8.05E-01	1.01	-+---+0+--00	3.61E-03	0.91	2.88E-02	*LOC6500794 & GRK6PS*
12	rs2255953	G	0.21	5.31E-01	0.97	+-X---+-0++0	5.12E-03	1.13	4.54E-02	*ETV6*
15	rs754661	G	0.41	5.45E-01	0.96	-00X-++---0+	6.29E-03	0.91	1.08E-01	*GABRB3*
8	rs16886291	A	0.13	5.01E-03	1.12	---+--+-+00+	5.41E-03	0.86	1.55E-01	*hCG_1644355 & TRPS1*

*MAF = minor allele frequency in NELSON; *Direction of effect per cohort: each sign reflects one cohort, direction of effect is presented by: + = (OR > 1.05), – = (OR < 0.95), 0 = (0.95 < OR < 1.05) and x = (missing result); cohorts are presented in the same order as in Table 2; OR is odds ratio; Q = p-value for heterogeneity; NA is not annotated.*

Functional relevance of SATB1 and rs6577641

1) Transcriptional regulation of *SATB1* mRNA expression

We analyzed the association of *SATB1* expression levels in lung tissue with rs6577641 genotype in 3 independent data sets of the Universities of Groningen, Laval and UBC [16]. A *cis*-acting effect of rs6577641 on *SATB1* expression was identified and present in all three datasets (n = 1,095), with the same direction of effect across all three *SATB1* probes on the array. The (susceptibility) G allele increased expression, the (protective) A allele reduced expression ($p = 4.3 \times 10^{-9}$) in the meta-analysis across the three datasets and across all three SATB1 probes measured (Table 6).

Table 6. Meta-analysis of the effect of rs6577641 on mRNA expression levels of SATB1 in the lung*

Gene	Affymetrix Probe ID	Z-score Groningen N=351	Z-score Laval N=335	Z-score UBC N=409	Z-Score Meta-Analysis	p-value Meta-Analysis
SATB1	100148784_TGI_at	-2.28	-0.08	-1.62	-2.29	0.022
SATB1	100150253_TGI_at	-0.84	-0.49	-1.62	-1.70	0.088
SATB1	100305926_TGI_at	-2.81	-1.38	-1.46	-3.26	0.001

** To assess the effect of the SNP rs6577641 on gene expression, a Kruskal-Wallis test was performed. This test generates a p-value, but does not give a direction of the effect. To assess the direction of the effect, a Spearman's correlation test was performed. Next, a Z-score was calculated for each center and a meta-analyses performed for each of the three SATB1 probes across all centers. Finally, a meta-analyses for all three SATB1 probes was performed across all centers. This generated a Z-score of -5.87 and a corresponding p-value of 4.3 * 10-9, indicating that the susceptibility G allele of the SNP rs6577641 increases SATB1 expression.*

2) *SATB1* mRNA expression and CMH

We compared *SATB1* expression in baseline airway wall biopsies of COPD patients with (n = 38) and without (n = 39) CMH in GLUCOLD [17]. CMH was significantly associated with *SATB1* expression levels (corrected for ex-/current smoking; p = 0.0045; Figure 5). After stratification, the same direction of effect was present in ex- and current smokers. However, this association reached statistical significance in current smokers (p = 0.021) and not in ex- smokers (p = 0.132), probably due to a difference in power as 46 subjects were current smokers versus 33 ex-smokers.

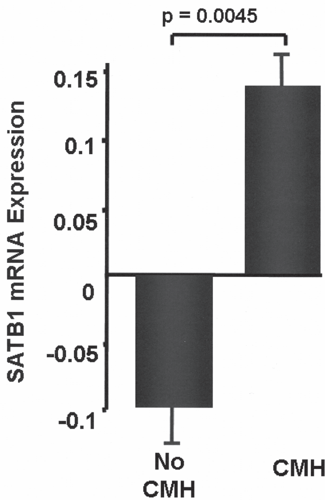


Figure 5. Bronchial biopsy mRNA-expression levels of SATB1 in COPD patients with chronic mucus hypersecretion (n = 38) compared to patients without chronic mucus hypersecretion (n = 39).

3) Genotype related protein expression and mucus positivity in bronchial epithelium

SATB1 protein expression has previously been observed in IHS analysis of bronchial epithelial cells [19]. Therefore, we stained SATB1 on paraffin embedded lung tissue biopsies of individuals from the Groningen population contributing to the eQTL analysis. We observed clear nuclear staining for SATB1 in bronchial epithelial cells. No significant difference for % of strong positive, positive and weak positive cells was observed between the protective (AA, n=9) and risk (GG, n=14) rs6577641 genotypes (11.8% ± 5.8 versus 12.7% ± 6.9, p = 0.74).

We determined whether the fraction of mucus positive bronchial epithelium was different in subjects with different homozygous rs6577641 genotypes and performed PAS-staining on tissue biopsies from the same cohort. We observed no significant difference between individuals with the homozygous protective (AA, n=10) and risk (GG, n=7) alleles (19.7% ± 11.9 versus 14.3% ± 9.6, p = 0.34).

4) *SATB1* expression levels during bronchial epithelial cell mucociliary differentiation

We investigated whether SATB1 expression was induced during mucociliary differentiation of primary human bronchial epithelial (HBE) cells *in vitro* and compared SATB1 mRNA expression levels at different time points of an air-liquid interface (ALI) culture for up to 45 days. ALI culture of HBE cells induced mucociliary differentiation,

as confirmed by induction of expression of FOXJ1, a marker for ciliated cells (19) and MUC5AC, a marker of goblet cells. SATB1 expression was induced over time (Figure 6), with an approximately 8-fold increased expression from the start to the end of the 45-day ALI culture period.

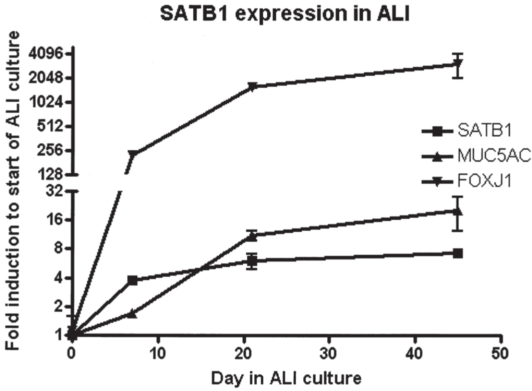


Figure 6. SATB1, MUC5AC and FOXJ1 mRNA expression levels during mucociliary human airway epithelial cell differentiation (n=2 donors). Expression of SATB1, the identified gene in our study, MUC5AC a marker of mucus, and FOXJ1, representing ciliated cells in epithelial cell culture on air liquid interface.

3

Discussion

Since not every ex- or current heavy smoker suffers from chronic mucus hypersecretion (CMH), we aimed to identify genetic variants conferring susceptibility to CMH. Therefore, we performed the first GWA study on CMH, the key presenting symptom in chronic bronchitis. CMH was associated with 36 SNPs at the $p < 10^{-4}$ significance level in the identification cohort. In the meta-analysis combining our identification and replication cohorts, strong association was observed with rs6577641, a SNP located on chromosome 3 in intron 9 of *SATB1*. Although the association of rs6577641 with CMH did not reach conventional genome-wide significance, its effect was in the same direction and was significant (4.25×10^{-6}) at nominal levels (1.61×10^{-3}) across eleven study populations, showing the robustness of this finding. The detected odds ratio for this SNP suggests an additional risk of 17% per G allele to develop CMH in a population of ex- and current heavy smokers.

Multivariate regression analysis, stratified for current an ex-smoking, showed essentially the same effect sizes and direction of the association of CMH and the risk allele of rs6577641. It is likely that lack of power is the reason for not reaching the level of significance in ex-smokers.

These data strongly suggest that *SATB1* plays a role in the susceptibility to CMH in subjects with a history of heavy smoking (≥ 20 packyears) within the general population. Moreover, rs6577641 has a *cis*-eQTL effect on *SATB1* lung tissue expression, the risk allele at rs6577641 (G) increasing and the A-allele reducing expression of *SATB1* significantly. Additionally, we found a higher *SATB1* expression in bronchial biopsies of COPD-patients with CMH. We found no differences between the GG and AA genotypes for protein expression of SATB1 in airway epithelium by IHC in a small sample from our lung tissue registry. Finally, we demonstrate that *SATB1* mRNA expression is induced during mucociliary differentiation in ALI cultures of human bronchial epithelial cells of 2 donors supporting our eQTL findings. Interestingly, expression of the mucin gene *MUC5AC* was also induced during this culture period, with a slightly delayed kinetics compared to *SATB1*. Together these data strongly suggest that *SATB1* is induced during differentiation of bronchial epithelial cells and affects chronic mucus hypersecretion.

The forest plot clearly shows that the effect of SNP rs6577641 is lower in cohorts including COPD patients only (GLUCOLD, Rucphen, COPDGene, ECLIPSE and Norway) than in the other cohorts. Additional meta-analysis of COPD-cohorts and general population based cohorts separately confirmed this (COPD cohorts, combined p-value = 0.236, OR

= 1.07 and general population based cohorts, combined p-value = 5.18×10^{-7}, OR = 1.26). This suggests genetic heterogeneity of CMH in subjects with and without COPD.

The SNP most significantly associated with CMH, rs6577641, is located in an intron of *SATB1*. *SATB1* is a transcription factor and chromatin (re)organizer important for controlling the expression of many genes in a tissue or cell-type specific fashion, for instance in differentiating thymus T-cells [20] or differentiating skin keratinocytes [21]. Expression of SATB1 has been observed in normal human bronchial epithelial cells by immunohistochemistry and lower levels were observed in non-small lung cancer cells [19]. In our study, we also showed the presence of SATB1 in bronchial epithelial cells by IHC staining of lung tissue. However, no significant differences were found between patients homozygous for the protective and risk alleles, for either specific SATB1 staining or for PAS staining, the latter specifically detecting mucus. This inability to detect a genotype effect on protein staining may be due to lack of power, as we found a large variation in SATB1 and PAS protein expression in the relatively small number of lung tissue samples. Other explanations include possible expression regulation of SATB1 by smoke exposure which could be a dynamic process not readily detected at the protein level by any single-time point analysis such as IHC staining on lung biopsies. Alternatively SATB1 expression levels may vary throughout the lungs or the technique used here is not sensitive enough to detect relatively small differences in protein levels. To further explore the association of SATB1 protein and its underlying regulation, it would be of interest to perform longitudinal investigations on lung tissue samples of subjects with and without CMH, or time series of *in vitro* cultured epithelial cells from donors with a specific genotype and cigarette smoke exposure. This would also allow further studies on epigenetic regulation with methylation, microRNA or histone modifications.

The lack of association between the SATB1 protein and rs6577641 might additionally be due to the location of mucus positive cells in lung tissue. Mucus is produced both by goblet cells and submucosal glands, which we did not investigate further. Normal mucus consists of 97% water and 3% solids including 30% mucins. In case of dysregulation of mucus production, the concentration of solids in mucus may increase up to 15%. A further step therefore could involve investigating mucins/proteins present in mucus, e.g. MUC5AC is predominantly produced by goblet cells in proximal airways and MUC5B by secretory cells throughout the airways and by submucosal glands.

How does *SATB1* expression contribute to CMH? *SATB1* is known to be a genome organizer, a tissue specific chromatin remodeling protein with a property to modifying

chromatin architecture by formation of loops, allowing contact of condensed genomic DNA to regulatory transcription proteins [22].
Thus *SATB1* can control gene expression of a series of target genes located within a single locus at a specific chromosomal location [23]. This has for instance been elegantly shown in case of differentiating keratinocytes [21], where Satb1 expression regulates genes located in the keratinocyte-specific loci, leading to adaptation of a specific cell fate of the differentiating keratinocytes. Similarly, a mechanism by which *SATB1* could contribute to CMH is the induction of a gene expression program during differentiation of bronchial epithelial cells, leading to adaptation of a cell fate specific for mucus producing cells in the submucosal glands or a goblet cell phenotype in the bronchial epithelium. Involvement of Satb1 in pneumocyte differentiation was previously observed by Baguma et al. in mice [24].
We observed induction of *SATB1* expression in bronchial epithelial cells differentiating under ALI culture conditions. Further research will need to test whether a specific gene expression profile is induced by *SATB1* expression in differentiating bronchial epithelial cells. SATB1 is also highly expressed in thymocytes, but absent in mature non-activated T cells [25]. Moreover, Satb1 has been shown in mice to be essential for expression of $T_{helper}2$ (Th2) cells important in the regulation of genes encoding interleukin 4, 5 and 13 [22]. In Satb1-deficient mice, development of thymocytes stopped after the $CD4^+/CD8^+$ stage with deregulation of many genes [26]. Conversely, in case of excessive SATB1-production an excess of Th2 cells may be formed which all produce IL-13, which may contribute to increased mucus production. Therefore, a putative role of SATB1 in T-cells for the CMH phenotype should not be disregarded.

Strength of our study is the fact that we were able to replicate our findings in different populations, ranging from cohorts consisting of individuals with severe airflow limitation to cohorts mainly consisting of healthy smokers. There are some limitations, e.g. the presence of CMH was not based on actual measurements of the amount of sputum produced but based on questionnaires that were not completely similar in all study cohorts. Underreporting of CMH occurs since those experiencing CMH become accustomed to these symptoms, believing they are smoking related or because they are embarrassed to admit to cough and sputum. We demonstrated that *SATB1* mRNA expression is induced during mucociliary differentiation in ALI cultures of HBE cells in a small dataset (n=2). However, these data seem reliable as they are supported by eQTL data from lung tissue. Despite this drawback, we consistently found evidence for association of *SATB1* with CMH in the populations studied, showing the robustness of our finding. Moreover, we corroborated this finding by functional studies in lung tissue, airway wall biopsies of COPD patients and epithelial cultures. More extensive research is needed to investigate which factors induce *SATB1* expression in airway epithelium.

In summary, we performed identification analyses and meta-analyses using data from almost 7,000 participants to identify genes involved in susceptibility for CMH. It is remarkable that we found a genetic association for CMH given this phenotype is partly subjectively determined and not well delineated. Moreover, despite cohort differences to define CMH and severity of airflow limitation, we found consistent effects of SNP rs6577641 on CMH. This confirms that the CMH phenotype, despite the fact that it is self-reported, is a robust phenotype irrespective of the presence or absence of airflow limitation. The association of rs6577641 on chromosome 3 at the *SATB1* locus with CMH was supported by functional studies including gene expression findings, demonstrating SATB1 to be associated with CMH.

Chronic mucus hypersecretion is a bothersome symptom for many people, it increases in prevalence with aging and affects quality of life, exacerbations of symptoms due to respiratory infections and ultimately increases mortality. The involvement of *SATB1* in CMH offers opportunities to better understand the process leading to CMH, and future development of tailored medicines.

References

1 Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. American Thoracic Society. *American journal of respiratory and critical care medicine*. 1995;152(5 Pt 2):S77-121

2 Ferre A, Fuhrman C, Zureik M et al. Chronic bronchitis in the general population: influence of age, gender and socio-economic conditions. *Respiratory medicine*. 2012;106(3):467-471

3 Rabe KF, Hurd S, Anzueto A et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association*. 2007;176(1073-449; 6):532-555

4 Vestbo J and Rasmussen FV. Respiratory symptoms and FEV1 as predictors of hospitalization and medication in the following 12 years due to respiratory disease. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*. 1989;2(0903-1936; 8):710-715

5 Vestbo J, Prescott E and Lange P. Association of chronic mucus hypersecretion with FEV1 decline and chronic obstructive pulmonary disease morbidity. Copenhagen City Heart Study Group. *American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association*. 1996;153(1073-449; 5):1530-1535

6 Website World Health Organization. Available at: http://www.who.int/respiratory/copd/burden/en/index.html. 2013

7 Shurin S. Morbidity & Mortality: 2012 Chart Book on Cardiovascular, Lung, and Blood Diseases. . 2012;2013:107

8 Lange P, Parner J, Prescott E and Vestbo J. Chronic bronchitis in an elderly population. *Age and Ageing*. 2003;32(0002-0729; 6):636-642

9 Ekberg-Aronsson M, Pehrsson K, Nilsson JA, Nilsson PM and Lofdahl CG. Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. *Respiratory research*. 2005;6:98

10 Viegi G, Carrozzi L, Di Pede F et al. Risk factors for chronic obstructive pulmonary disease in a north Italian rural area. *European journal of epidemiology*. 1994;10(6):725-731

11 Silverman EK, Chapman HA, Drazen JM et al. Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. Risk to relatives for airflow obstruction and chronic bronchitis. *American journal of respiratory and critical care medicine*. 1998;157(6 Pt 1):1770-1778

12 Hallberg J, Dominicus A, Eriksson UK et al. Interaction between smoking and genetic factors in the development of chronic bronchitis. *American journal of respiratory and critical care medicine*. 2008;177(5):486-490

13 Zhu G, Agusti A, Gulsvik A et al. CTLA4 gene polymorphisms are associated with chronic bronchitis. *Eur. Respir.J.* 2009;(1399-3003)

14 van Klaveren RJ, Oudkerk M, Prokop M et al. Management of lung nodules detected by volume CT scanning. *The New England journal of medicine*. 2009;361(23):2221-2229

15 Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*. 2007;81(0002-9297; 0002-9297; 3):559-575

16 Hao K, Bosse Y, Nickle DC et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS genetics*. 2012;8(11):e1003029

17 Lapperre TS, Snoeck-Stroband JB, Gosman MM et al. Effect of fluticasone with and without salmeterol on pulmonary outcomes in chronic obstructive pulmonary disease: a randomized trial. *Annals of Internal Medicine*. 2009;151(8):517-527

18 Hofman A, van Duijn CM, Franco OH et al. The Rotterdam Study: 2012 objectives and design update. *European journal of epidemiology*. 2011;26(8):657-686

19 Selinger CI, Cooper WA, Al-Sohaily S et al. Loss of special AT-rich binding protein 1 expression is a marker of poor survival in lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer*. 2011;6(7):1179-1189

20 Cai S, Lee CC and Kohwi-Shigematsu T. SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nature genetics*. 2006;38(11):1278-1288

21 Fessing MY, Mardaryev AN, Gdula MR et al. p63 regulates Satb1 to control tissue-specific chromatin remodeling during development of the epidermis. *The Journal of cell biology*. 2011;194(6):825-839

22 Cai C, Zhang HY, Le JJ et al. Inflammatory airway features and hypothalamic-pituitary-adrenal axis function in asthmatic rats combined with chronic obstructive pulmonary disease. *Chinese medical journal*. 2010;123(13):1720-1726

23 Yasui D, Miyano M, Cai S, Varga-Weisz P and Kohwi-Shigematsu T. SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature*. 2002;419(6907):641-645

24 Baguma-Nibasheka M, Angka HE, Inanlou MR and Kablar B. Microarray analysis of Myf5-/-:MyoD-/- hypoplastic mouse lungs reveals a profile of genes involved in pneumocyte differentiation. *Histology and histopathology*. 2007;22(5):483-495

25 Dickinson LA, Joh T, Kohwi Y and Kohwi-Shigematsu T. A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell*. 1992;70(4):631-645

26 Krangel MS. T cell development: better living through chromatin. *Nature immunology*. 2007;8(7):687-694

27 Soler Artigas M, Loth DW, Wain LV et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nature genetics*. 2011;43(11):1082-1090

28 Stolk RP, Rosmalen JG, Postma DS et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *European journal of epidemiology*. 2008;23(1):67-74

29 Rijcken B, Schouten JP, Mensinga TT, Weiss ST, De Vries K and Van der Lende R. Factors associated with bronchial responsiveness to histamine in a population sample of adults. *The American Review of Respiratory Disease*. 1993;147(6 Pt 1):1447-1453

30 van Diemen CC, Postma DS, Vonk JM, Bruinenberg M, Nolte IM and Boezen HM. Decorin and TGF-beta1 polymorphisms and development of COPD in a general population. *Respiratory research*. 2006;7:89

31 Verschuren WM, Blokstra A, Picavet HS and Smit HA. Cohort profile: the Doetinchem Cohort Study. *International journal of epidemiology*. 2008;37(6):1236-1241

32 Nizankowska-Mogilnicka E, Mejza F, Buist AS et al. Prevalence of COPD and tobacco smoking in Malopolska region--results from the BOLD study in Poland. *Polskie archiwum medycyny wewnetrznej*. 2007;117(9):402-410

33 Lamprecht B, McBurnie MA, Vollmer WM et al. COPD in never smokers: results from the population-based burden of obstructive lung disease study. *Chest*. 2011;139(4):752-763

34 Becker N, Delorme S and Kauczor H,U. LUSI: the German component of the European trial on the efficacy of multislice-CT for the early detection of lung cancer [abstract]. *onkologie 31*. 2008

35 Vestbo J, Anderson W, Coxson HO et al. Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*. 2008;31(4):869-873

36 Regan EA, Hokanson JE, Murphy JR et al. Genetic epidemiology of COPD (COPDGene) study design. *Copd*. 2010;7(1):32-43

37 Cho MH, Castaldi PJ, Wan ES et al. A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19q13. *Human molecular genetics*. 2011

38 Grydeland TB, Dirksen A, Coxson HO et al. Quantitative computed tomography: emphysema and airway wall thickness by sex, age and smoking. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*. 2009;34(1399-3003; 4):858-865

Chapter 3

Supplement

Populations, genotyping and quality control

The *NELSON Study* is a Dutch multi-center lung cancer screening study and includes only male current or former smokers (≥ 20 packyears). Detailed inclusion criteria and characteristics have been described elsewhere [1].

Genome-wide single nucleotide polymorphism (SNP) genotyping for NELSON individuals was performed on Illumina Human 610-Quad BeadChip containing over 620,000 markers. Genotypes were called with the standard algorithm provided by Illumina and implemented in Genome Studio software. Quality control (QC) in NELSON implied exclusion of those individuals from whom > 5% of genotyping was missing, when detected as an ethnic outlier or when a sample was derived from a relative of another participant (based on genetic distance derived from principal components c1 and c2, and on IBS estimation, Phat > 0.5) and when lung cancer was present. SNPs were excluded if minor allele frequency (MAF) was < 5%, if a deviation from Hardy-Weinberg was observed ($p < 0.0001$) and if > 5% of samples were missing.

In the first GWA on CMH (with adjustment for center only), we included blood bank controls to increase the power. However, except for gender, there was no information available for these controls. We could not include these controls in the second GWA analysis (with adjustment for center and ex-/current smoking) as we had no information on smoking habits.

COPACETIC-populations

The *Doetinchem Study* is a general population cohort of the inhabitants of Doetinchem, an industrial town in the Netherlands. We used data collected between 1998 and 2002. Detailed inclusion criteria and characteristics have been described elsewhere [2].

The *Vlagtwedde-Vlaardingen Study* is a Dutch population based study based on a random sample of the general population from Vlagtwedde (a rural village) and Vlaardingen (an industrial village). We used data collected between 1989 and 1990. Detailed inclusion criteria and characteristics have been described elsewhere [3, 4].

GLUCOLD is an acronym for Groningen Leiden Universities and Corticosteroids in Obstructive Lung Disease. The *GLUCOLD Study* is a multicenter trial including COPD-patients. All patients had a history of ≥ 10 packyears smoking and COPD-stage II or higher. Detailed inclusion criteria and characteristics have been described previously [5].

POLAND is a general population based Polish cohort participating in the Burden of Lung Disease (BOLD) initiative, an international study designed to investigate COPD prevalence and economic burden. The Polish population consisted of a random sample aged ≥ 40 years from the Malopolska region in southern Poland. Detailed inclusion criteria and characteristics have been described elsewhere [6, 7].

Participants in the *Rucphen Study* are part of a genetically isolated population located in the southwest of The Netherlands. All participants were heavy smokers and had COPD stage II or higher based on GOLD-guidelines [4].

The *Heidelberg* cohort is a German cohort participating in a Lung Cancer Screening Intervention Trial using the same inclusion criteria as the NELSON study [8].

Genotyping of top-SNPs for replication in COPACETIC-cohorts was performed on custom made Veracode assays. QC in COPACETIC-cohorts implied exclusion of those individuals from whom > 10% of genotyping was missing. SNPs were excluded if minor allele frequency (MAF) was < 5%, if a deviation from Hardy-Weinberg was observed ($p < 0.0001$) and if > 5% of samples were missing.

Non COPACETIC populations

The *Rotterdam Study* is a prospective population-based cohort study founded in 1990 in a suburb of Rotterdam, the Netherlands. The first cohort (RS I) consists of 7,983 participants, aged 55 years and over. The second cohort (RS II) was recruited in 2000 with the same inclusion criteria. The third cohort (RS III) was recruited in 2006. Details regarding the Rotterdam study have been described elsewhere [9, 10].

The *LifeLines Cohort Study* is a prospective population based cohort study on multimorbidity, being conducted in the northern part of the Netherlands [11].

NORWAY is a COPD (GOLD stage II or worse) case-control cohort from Bergen, Norway. Details regarding the *NORWAY Study* have been published previously [12, 13]. All participants were ex- or current smokers with ≥ 2.5 packyears and age ≥ 40 years.

COPDGene is a multicenter study including subjects from 21 clinical study centers throughout USA with: a smoking history of ≥ 10 packyears. Subjects with COPD (GOLD II-IV) and normal spirometry were included. Only non-Hispanic White subjects from COPDGene were included in this analysis. Details regarding the COPDGene study have been described elsewhere [14].

ECLIPSE, the Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints, is a multicenter case-control study including individuals from 46 centers in 12 countries. Details regarding the ECLIPSE study have been published previously [15]. COPD-cases (GOLD II–IV, ≥ 10 packyears), smoking controls (≥ 10 packyears) and nonsmoking controls (< 1 packyear) without COPD are included. Only Caucasian subjects were included in this analysis.

The non-COPACETIC-studies performed their own genotyping using different commercially available platforms, quality control and imputation. Imputations of non-genotyped SNPs were carried out within each study.

The lung eQTL study

Non-tumor lung tissues were collected from patients who underwent lung resection surgery at three participating sites: Laval University (Quebec City, Canada), University of Groningen (Groningen, The Netherlands), and University of British Columbia (Vancouver, Canada). Whole-genome gene expression and genotyping data were obtained from these specimens. Gene expression profiling was performed using an Affymetrix custom array testing 51,627 non-control probe sets and normalized using RMA (Irizarry, R.A. et al. Biostatistics 2003; 4, 249-64). Genotyping was performed using the Illumina Human1M-Duo BeadChip array. At Laval, lung specimens were collected from patients undergoing lung cancer surgery and stored at the "Institut universitaire de cardiologie et de pneumologie de Québec" (IUCPQ) site of the Respiratory Health Network Tissue Bank of the "Fonds de recherche du Québec - Santé" (www.tissuebank.ca). Written informed consent was obtained from all subjects and the study was approved by the IUCPQ ethics committee. At Groningen, lung specimens were provided by the local tissue bank of the Department of Pathology and the study protocol was consistent with the Research Code of the University Medical Center Groningen and Dutch national ethical and professional guidelines ("Code of conduct; Dutch federation of biomedical scientific societies"; http://www.federa.org). At Vancouver, the lung specimens were provided by the James Hogg Research Center Biobank at St Paul's Hospital and subjects provided written informed consent. The study was approved by the ethics committees at the UBC-Providence Health Care Research Institute Ethics Board. The lung eQTL analysis was performed as described before by Fehrman and Hao [16, 17].

Gene expression analysis in GLUCOLD

RNA Isolation and Size Fractionation

Out of 114 COPD subjects in GLUCOLD, 89 individuals had endobronchial biopsies which had been immediately snap-frozen, stored at -80 ^{o}C and were available for extraction of RNA. RNA was extracted from bronchial biopsies and fractioned into low molecular weight (<200 nt) and high molecular weight (>200 nt) fractions using the miRNeasy mini kit (QIAGEN) according to manufacturer's protocol. The purity of RNA fractions was assessed using a NanoDrop 1000 UV-Vis spectrophotometer, and the integrity of the large RNA fraction was assessed by using the RNA Pico assay in the Agilent 2100 BioAnalyzer.

3

RNA processing and microarray hybridization

All procedures were performed at Boston University Microarray Resource Facility as described in GeneChip® Whole Transcript (WT) Sense Target Labeling Assay Manual (Affymetrix, Santa Clara, CA, current version available at www.affymetrix.com). The Qiagen miRNeasy Mini Kit and RNeasy MinElute Cleanup Kit were used to isolate high and low molecular weight RNA. 200 ng of high molecular weight large RNA was reverse transcribed using the Whole Transcript cDNA Synthesis kit (Affymetrix, Santa Clara, CA). The obtained cDNA was used as a template for in vitro transcription using the Whole Transcript cDNA Amplification Kit (Affymetrix, Santa Clara, CA). The obtained antisense cRNA was purified using GeneChip Sample Cleanup Module (Affymetrix, Santa Clara, CA), and used as a template for reverse transcription (Whole Transcript cDNA Synthesis kit, Affymetrix, Santa Clara, CA) to produce single-stranded DNA in the sense orientation. During this step, dUTP was incorporated. The DNA was then fragmented using uracil DNA glycosylase (UDG) and apurinic/apyrimidinic endonuclease 1 (APE 1) and labeled with DNA Labeling Reagent that was covalently linked to biotin using terminal deoxynucleotidyl transferase (TdT, Whole Transcript Terminal Labeling kit, Affymetrix, Santa Clara, CA). IVT and cDNA fragmentation quality controls were carried out by running an mRNA Nano assay in the Agilent 2100 Bioanalyzer. The labeled fragmented DNA was hybridized to the Affymetrix Human Gene 1.0 ST Arrays for 16-18 hours in GeneChip Hybridization oven 640 at 45oC with rotation (60 rpm). The hybridized samples were washed and stained using Affymetrix fluidics station 450. The first stain with streptavidin-R-phycoerythrin (SAPE) was followed by signal amplification using a biotinilated goat anti-streptavidin antibody and another SAPE staining (Hybridization, Washing and Staining Kit, Affymetrix, Santa Clara, CA). Microarrays were immediately

scanned using Affymetrix GeneArray Scanner 3000 7G Plus (Affymetrix, Santa Clara, CA).

Data acquisition, probeset summarization and normalization, and data preprocessing
Normalization was performed with Affymetrix Expression Console software using Affymetrix default Robust Multichip Analysis (RMA) sketch algorithm workflow and 1 additional sample was excluded due to low quality of the microarray data.
Microarray data quality was assessed using relative log expression (RLE) plots, normalized unscaled standard error (NUSE) plots, and principle component analysis (PCA). Based on the RLE and NUSE plots, a total of 9 microarrays were excluded, leaving 79 microarrays for subsequent analysis, 77 having data on CMH (38 CMH-cases, 39 non-CMH-controls).
Association of CMH with *SATB1* mRNA-expression levels was analyzed with logistic regression and adjustment for current smoking and RNA integrity score.

Gene expression levels during airway epithelial cell differentiation

To investigate *SATB1* gene expression levels during airway epithelial cell differentiation, a time-course series of air liquid interface cultured cells (ALIs) and submerged cultured cells was purchased (MucilAirTM, Epithelix Sàrl, Geneva, Switzerland) and analyzed as described before [18]. ALI cultured cells were analyzed harvested from two independent culture series (duplicate), whereas submerged cultured cells were analyzed from a single series of cultures. Briefly, at time-points 1, 7, 21 and 45 days after start of ALI culture, cells were harvested for RNA and analyzed for gene expression levels by qRT-PCR analysis using Taqman Assays (Applied Biosystems Europe BV, Nieuwekerk A/D IJssel, the Netherlands). qRT-PCRs were performed on the ABI7900HT cycler in 384-well format. Pre-designed assays were used for the detection of *SATB1* (Hs00161515_m1), FOXJ1 (Hs00230964_m1), MUC5AC (Hs01365616_m1) and four house-keeping genes: GAPDH (Hs99999905_m1), β-actin (Hs99999903_m1), and RPLPO (Hs99999902_m1). cDNA was produced using OmniscriptTM Reverse Transcriptase (QIAGEN Benelux BV, Venlo, the Netherlands). A total of 500 ng of RNA was reverse transcribed into cDNA using Oligo-12-18) primers in a 20 µl reaction volume, including RNase inhibitor, at 37 ^{o}C for 1 hour. Each qPCR reaction contained 17.5ng of cDNA, 250nM of probe, 900nM of forward and reverse primers, 5 µl TaqMan Universal PCR Master Mix (Applied Biosystems Europe BV, Nieuwekerk A/D IJssel, The Netherlands), in a final volume of 10 µl. All samples were measured in duplicate using recommended cycling conditions. Data was analyzed using SDS2.3 software by applying the ΔΔCt-method (Applied Biosystems

User Bulletin 2). As three house-keeping genes were used, the best combination of house-keeping genes for normalization was determined by using the Normfinder applet.

Immunohistochemical staining

Bronchial tissue for SATB1- and PAS staining was available from patients in whom lung surgery or lung transplantation was conducted because of COPD or lung cancer.

SATB1-staining

Monoclonal mouse anti-SATB1 (BD cat nr. 611182) was used for immunohistochemistry. Antigen retrieval was performed with 10mM Tris/1mMEDTA buffer pH 9.0 at 125 ^{o}C for 15 minutes in a Pascal pressure chamber (Dako), and the primary antibody was incubated overnight at 4 ^{o}C in a 1/100 dilution. The incubation of the second step, rabbit anti mouse 1/100, and the third step, goat anti rabbit 1/100, was 30 minutes. The color reaction was with Di-Amino Benzidin (DAB; Sigma, Illinois). Tonsil tissue was used as a positive control.

The number of strong positively-, positively- and weak positively stained cells per case was counted in the epithelium. Quantifiable areas were selected for evaluation when the following requirements were met 1) SATB1 and HE staining was good 2) intact basement membrane (BM) 3) presence of maximal 1-3 layer(s) of epithelial cells, including a layer of ciliated epithelium, thus avoiding hyperplasia and squamous metaplasia.

PAS-staining

Periodic acid-shift (PAS) histochemical staining was performed using the DAKO autostainer (DAKO, Glostrup, Denmark) at the Pathology department. Areas were selected for evaluation when the following requirements were met 1) PAS staining was acceptable 2) intact basement membrane (BM) 3) presence of maximal 1-3 layer(s) of epithelial cells, including a layer of ciliated epithelium.

The number of PAS-positive pixels was determined in the epithelium and expressed as the percentage of mucus-positivity in all measured epithelium per sample by using Aperio® Scanscope software.

All stainings were quantified by a blinded observer.

Table 1. Meta-analysis of top SNPs associated with CMH across replication cohorts and across identification and replication cohorts, corrected for smoking and sex.

		Meta-analysis across replication cohorts			Meta-analysis across Identification and replication cohorts			
Chr	**SNP**	**p-value**	**OR**	**Q**	**p-value**	**OR**	**Q**	**Close(st) gene(s)**
3	rs6577641	5.01E-03	1.12	9.19E-01	4.25E-06	1.17	6.20E-01	*SATB1**
18	rs8086262	2.16E-02	1.11	1.00E-02	8.91E-02	1.13	2.60E-03	*LOC100132647 & CBLN2*
8	rs4242562	5.04E-02	1.18	4.20E-01	6.15E-01	1.07	1.40E-03	*hCG_1644355 & TRPS1*
8	rs10098746	5.74E-02	0.91	1.76E-01	6.22E-01	0.95	2.00E-04	*MTSS1 & LOC100130448*
12	rs1379963	8.87E-02	1.09	6.08E-01	4.57E-01	1.06	2.10E-03	*KCNC2**
12	rs12822199	9.66E-02	1.09	5.30E-01	8.50E-01	1.02	4.70E-03	*KCNC2**
13	rs1211304	1.37E-01	1.10	6.86E-01	8.70E-01	0.98	2.50E-03	*KPNA3 & LOC220429*
12	rs2855708	2.18E-01	1.06	7.02E-01	1.20E-03	1.13	1.76E-01	*ETV6**
3	rs1488757	2.34E-01	0.92	8.44E-01	1.10E-03	0.83	1.55E-01	*LOC727810 & CNTN4*
12	rs1795669	2.83E-01	1.09	7.02E-01	2.90E-03	1.22	1.77E-01	*LOC100130336 & LOC100131830*
14	rs992745	2.94E-01	0.95	4.06E-01	2.74E-03	0.89	4.59E-02	*LOC728755**
13	rs9578362	3.06E-01	0.96	2.06E-01	3.61E-03	0.91	2.88E-02	*LOC650794 & GRK6PS*
4	rs4306981	3.37E-01	1.04	3.45E-01	2.89E-03	1.12	5.19E-02	*PAQR3 & ARD1B*
2	rs6735868	3.66E-01	1.05	8.51E-01	1.59E-01	0.94	1.38E-02	*TMEM182 & LOC728815*
9	rs4842047	3.88E-01	0.96	9.99E-01	3.44E-03	0.89	3.03E-01	*CAMSAP1 & UBAC1*
17	rs11652469	4.51E-01	0.942	2.39E-01	9.77E-01	1.00	3.50E-03	*FLJ42842 & LOC388401*
15	rs754661	5.31E-01	0.974	8.08E-01	6.29E-03	0.91	1.08E-01	*GABRB3**
8	rs16886291	5.45E-01	0.963	8.77E-01	6.47E-03	0.86	1.32E-01	*hCG_1644355 & TRPS1*
20	rs4815628	5.90E-01	1.02	2.76E-01	4.51E-01	0.95	3.90E-03	*PANK2**
10	rs943189	6.12E-01	1.02	6.86E-01	8.82E-02	0.94	3.59E-02	*SPAG6 & LOC643475*
12	rs4760851	6.15E-01	1.021	9.57E-01	7.51E-02	0.94	6.50E-02	*TRHDE & LOC100128674*
3	rs1387089	6.80E-01	0.97	4.43E-01	1.13E-02	0.86	3.64E-02	*LOC391504 & LOC727810*
22	rs1005239	7.07E-01	0.98	5.10E-01	1.44E-02	0.92	6.50E-02	*TBC1D22A & RP11-191L9.1*
21	rs2032257	7.14E-01	1.02	3.24E-01	7.68E-02	0.94	9.90E-03	*APP & CYYR1*
8	rs7836298	7.16E-01	1.03	6.19E-01	5.60E-02	0.89	7.20E-03	*hCG_1644355 & TRPS1*
12	rs952394	7.31E-01	0.99	9.15E-01	5.54E-02	1.07	8.47E-02	*TRHDE & LOC100128674*
12	rs2255953	8.05E-01	1.01	8.06E-01	5.12E-03	1.13	4.54E-02	*ETV6**
16	rs13333521	8.34E-01	1.022	6.71E-02	3.11E-01	1.18	2.70E-03	*GPRC5B & GPR139*
22	rs1009147	8.58E-01	0.99	9.91E-01	2.07E-02	0.88	2.06E-01	*NF2**
8	rs7831595	8.96E-01	1.01	1.82E-01	1.90E-02	1.08	5.90E-03	*EPPK1*
11	rs11026531	9.75E-01	1.00	3.72E-01	3.49E-02	0.92	2.31E-02	*SLC17A6**

*P-value is fixed p-value if p-value for heterogeneity (Q) > 0.005, and random p-value if p-value for heterogeneity (Q) < 0.005; OR is Odds Ratio; OR is fixed OR if p-value for heterogeneity (Q) > 0.005, and random OR if p-value for heterogeneity (Q) < 0.005; Q is p-value for heterogeneity; N = number of cohorts; * means that the corresponding SNP is an intron in this gene.*

References

1 van Klaveren RJ, Oudkerk M, Prokop M et al. Management of lung nodules detected by volume CT scanning. *The New England journal of medicine*. 2009;361(23):2221-2229

2 Verschuren WM, Blokstra A, Picavet HS and Smit HA. Cohort profile: the Doetinchem Cohort Study. *International journal of epidemiology*. 2008;37(6):1236-1241

3 Rijcken B, Schouten JP, Mensinga TT, Weiss ST, De Vries K and Van der Lende R. Factors associated with bronchial responsiveness to histamine in a population sample of adults. *The American Review of Respiratory Disease*. 1993;147(6 Pt 1):1447-1453

4 van Diemen CC, Postma DS, Vonk JM, Bruinenberg M, Nolte IM and Boezen HM. Decorin and TGF-beta1 polymorphisms and development of COPD in a general population. *Respiratory research*. 2006;7:89

5 Lapperre TS, Snoeck-Stroband JB, Gosman MM et al. Effect of fluticasone with and without salmeterol on pulmonary outcomes in chronic obstructive pulmonary disease: a randomized trial. *Annals of Internal Medicine*. 2009;151(8):517-527

6 Nizankowska-Mogilnicka E, Mejza F, Buist AS et al. Prevalence of COPD and tobacco smoking in Malopolska region--results from the BOLD study in Poland. *Polskie archiwum medycyny wewnetrznej*. 2007;117(9):402-410

7 Lamprecht B, McBurnie MA, Vollmer WM et al. COPD in never smokers: results from the population-based burden of obstructive lung disease study. *Chest*. 2011;139(4):752-763

8 Becker N, Delorme S and Kauczor H,U. LUSI: the German component of the European trial on the efficacy of multislice-CT for the early detection of lung cancer [abstract]. *onkologie 31*. 2008

9 Soler Artigas M, Loth DW, Wain LV et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nature genetics*. 2011;43(11):1082-1090

10 Hofman A, van Duijn CM, Franco OH et al. The Rotterdam Study: 2012 objectives and design update. *European journal of epidemiology*. 2011;26(8):657-686

11 Stolk RP, Rosmalen JG, Postma DS et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *European journal of epidemiology*. 2008;23(1):67-74

12 Cho MH, Castaldi PJ, Wan ES et al. A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19q13. *Human molecular genetics*. 2011

13 Grydeland TB, Dirksen A, Coxson HO et al. Quantitative computed tomography: emphysema and airway wall thickness by sex, age and smoking. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*. 2009;34(1399-3003; 4):858-865

14 Regan EA, Hokanson JE, Murphy JR et al. Genetic epidemiology of COPD (COPDGene) study design. *Copd*. 2010;7(1):32-43

15 Vestbo J, anderson W, Coxson HO et al. Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*. 2008;31(4):869-873

16 Fehrmann RS, Jansen RC, Veldink JH et al. Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS genetics*. 2011;7(8):e1002197

17 Hao K, Bosse Y, Nickle DC et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS genetics*. 2012;8(11):e1003029

18 Koning H, Sayers I, Stewart CE et al. Characterization of protocadherin-1 expression in primary bronchial epithelial cells: association with epithelial cell differentiation. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*. 2012;26(1):439-448

Chapter 4

Dissecting genetic risk factors for chronic mucus hypersecretion in heavy smokers with and without COPD

AE Dijkstra, HM Boezen, M van den Berge, JM Vonk, PS Hiemstra, RG Barr, KM Burkart, A Manichaikul, TD Pottinger, EK Silverman, MH Cho, JD Crapo, TH Beaty, P Bakke, A Gulsvik, DA Lomas, Y Bossé, DC Nickle, PD Paré, HJ de Koning, JWJ Lammers, P Zanen, J Smolonska, C Wijmenga, CA Brandsma, HJM Groen, DS Postma

Submitted

Abstract

Background

Smoking is a notorious risk factor for chronic mucus hypersecretion (CMH). CMH frequently occurs in Chronic Obstructive Pulmonary Disease (COPD), a smoking associated lung disease. Not all individuals with COPD have CMH, and conversely, many individuals with CMH do not have COPD. Therefore, the question arises whether single nucleotide polymorphisms in genes (SNPs) are related to CMH in smokers with and without COPD, and if so, whether the same SNPs are involved.

Methods

We performed two genome wide association (GWA) studies on CMH, one in male smokers (≥ 20 packyears) with COPD (n = 849, 39.9% CMH) and the other in male smokers without COPD (n = 1,348, 25.4% CMH) under an additive genetic model using logistic regression (adjusted for ex/current smoking), followed by replication and meta-analysis in comparable populations selecting individuals with ≥ 5 packyears: four cohorts with and one cohort without COPD. Additional studies assessed the functional relevance of the most significantly associated genetic variants.

Results

GWA analysis on CMH in the cohort with COPD and the cohort without COPD yielded no genome wide significance after replication. In smokers with COPD, our top SNP (rs10461985, $p = 5.43 \times 10^{-5}$) was located in the *GDNF-antisense* gene that is functionally associated with the *GDNF* gene. Of interest, expression of *GDNF* in bronchial biopsies of COPD patients was significantly associated with CMH ($p = 0.007$).

In smokers without COPD, 4 SNPs had a p-value $< 10^{-5}$ in the meta-analysis, including a SNP (rs4863687) in the *MAML3* gene, the T allele showing modest association with CMH ($p = 7.57 \times 10^{-6}$, OR = 1.48) and with significantly increased *MAML3* expression in lung tissue ($p = 2.59 \times 10^{-12}$).

Conclusions

The most significantly associated SNPs with CMH in individuals with and without COPD were different, resulting in differential gene expression in lung tissue. Our data suggest the potential for differential genetic backgrounds of CMH in individuals with and without COPD.

Introduction

Chronic mucus hypersecretion (CMH) can be present in individuals with and without COPD. The prevalence of CMH varies from 3.5% to 12.7% in the general population depending on the population studied and the CMH definition used [1, 2]. The prevalence of CMH is much higher In individuals with COPD (30%) and increases with the severity of airflow limitation [3, 4]. Some risk factors for COPD and CMH overlap, like smoking, occupational exposures and bacterial infections [5-9].
However, not all heavy smokers have CMH, which may be explained by a genetic contribution to CMH, as evidenced by familial aggregation of mucus overproduction and higher concordance of CMH in monozygotic than in dizygotic twins [10-12]. So far, only two genetic studies on CMH have been published. One study suggested that *CTLA4* is associated with chronic bronchitis in individuals with COPD without a direct association with COPD itself [13]. A second study showed that a SNP (rs6577641) in the *SATB1* gene was strongly associated with CMH in a heavy smoking population [15].

Since not all individuals with COPD have CMH and conversely not all individuals with CMH have COPD, the question arises whether similar or differential genetic factors are involved in the development of CMH in individuals with and without COPD.
Therefore, we performed a genome wide association study on CMH in a group of male individuals with COPD and a group without COPD, from the same heavy smoking general population based cohort (NELSON) [14]. Subsequently, we evaluated our findings on the association with CMH in replication cohorts including individuals with and without COPD, and searched for features of our most significant findings.

Methods

Ethics Statement

The Dutch Ministry of Health and the Medical Ethics Committee of the hospital approved the study protocol for the Dutch centers. Ethics approval and written informed consent was obtained from all participants in the studies.

Identification population

Male Caucasian participants from Groningen and Utrecht were included from the Dutch NELSON study [14], a heavy smoking population based lung cancer screening study. Information on CMH and smoking behavior was collected by questionnaires as published previously [15]. Spirometry was performed according to the European Respiratory Society guidelines, including forced expiratory volume in 1 sec (FEV_1) and forced vital capacity (FVC), without using a bronchodilator [16]. COPD was defined as $FEV_1/FVC < 0.70$.
To assess whether different genetic factors contribute to the presence of CMH in smoking individuals with and without COPD, we conducted two genome wide association (GWA) studies; one in NELSON-individuals with COPD (NELSON-COPD) and a second in NELSON participants without COPD (NELSON-non-COPD) [14].

Replication populations

Top hits associated with CMH in NELSON-COPD were in silico analyzed in individuals with ≥ 5 packyears smoking and $FEV_1/FVC < 0.70$ from four independent, Caucasian COPD-cohorts: GenKOLS, COPDGene, ECLIPSE and MESA [17-20]. Subsequently meta-analyses were performed across these replication cohorts, and across NELSON-COPD and these replication cohorts.
Top hits associated with CMH in NELSON-non-COPD, were analyzed in the general population cohort LifeLines by selecting individuals without COPD and ≥ 5 packyears smoking.
A description of the replication cohorts is given in the supplement. Details on the identification and replication cohorts concerning genotyping method, genotyping imputation software, and CMH and COPD definitions are given in Supplementary Table 1.

Functional relevance of identified top SNPs

We assessed whether the top SNPs in individuals with and without COPD were associated with gene expression levels in human lungs. Expression quantitative trait loci (eQTLs) were identified in 1,095 lung tissues from three independent cohorts recruited

from Laval University, University of British Columbia, and University of Groningen as described previously [21].

Additionally, we assessed whether CMH was associated with mRNA expression of candidate genes in bronchial biopsies from 77 COPD participants in the Groningen and Leiden Universities study of Corticosteroids in Obstructive Lung Disease study (GLUCOLD) [22, 23].

Details on the methods are given in the Supplement.

Statistical analysis

Quality control (QC) of genotyping, regression- and meta-analyses were performed with PLINK 1.07 [24]. QC was performed in cases and controls according to the following exclusion criteria: SNPs with call rate $< 95\%$, Minor Allele Frequency (MAF) < 0.05, proportion of individuals for which no genotype was called (mind) < 0.95 and Hardy Weinberg equilibrium (HWE) $p < 0.0001$. Ethnic outliers, duplicates and relatives were removed (based on the top two components from multidimensional scaling).

Logistic regression analysis under an additive genetic model with adjustment for center and smoking (ex/current) was used to identify SNPs associated with CMH in NELSON participants in two separate analyses. SNPs were included for replication if there was any nominally significant association between CMH and a SNP ($p < 2.0 \times 10^{-4}$) and analyzed using additional adjustment for gender as the replication cohorts also included females.

Results

Populations

After QC, out of 3,005 NELSON participants, 2,799 remained. Females were excluded as only 48 were present after QC. 2,194 NELSON males with complete information on CMH, spirometry and smoking history were analyzed including 849 with and 1,345 without COPD. The prevalence of CMH in individuals with COPD was 39.8% (n = 338) and in individuals without COPD 25.4% (n = 342). Demographic and clinical characteristics of NELSON participants with COPD and of the four COPD-replication cohorts are presented in Table 1 [17-20].

Table 1. Characteristics of individuals with and without CMH, in NELSON-COPD and in replication COPD cohorts.

	NELSON		GenKOLS		COPDGene		ECLIPSE		MESA	
	+ CMH	- CMH	+ CMH	- CMH	+ CMH	- CMH	+ CMH	- CMH	+ CMH	- CMH
N (%)	338 (39.9)	511 (60.1)	487 (57.1)	364 (42.7)	182 (36.6)	315 (63.4)	643 (38.1)	1,045 (61.9)	50 (21.4)	184 (78.6)
Age, yrs	61.5 (5.9)	61.2 (5.4)	65.8 (10.0)	65.2 (10.0)	63.9 (7.8)	65.2 (8.3)	62.9 (7.6)	64.1 (6.8)	64.8 (9.4)	65.6 (9.1)
Female, %	0	0	34.0	46.5	39.0	57.1	24.7	38.5	58.0	64.7
Packyears	38.7 (20-140)	38.7 (20-119)	33.2 (5-119)	31.2 (5-130)	47.8 (11-238)	47.6 (10-146)	45.0 (6-220)	45.0 (10-205)	47.0 (6 - 135)	40.6 (5- 167)
Current smoking, %	74.8	50.2	53.5	39.7	42.9	23.5	45.1	27.0	38.0	12.5
FEV_1, %predicted	81.8 (19.8)	86.3 (7.1)	48.2 (17.5)	54.0 (16.8)	46.5 (18.1)	49.9 (18.5)	46.7 (15.4)	48.2 (15.7)	67.5 (18.6)	75.4 (17.4)
FEV_1/FVC, %	60.1 (8.6)	62.5 (7.1)	49.7 (13.4)	53.5 (12.2)	45.5 (11.9)	48.6 (13.8)	44.3 (11.8)	49.7 (13.3)	59.4 (10.5)	62.6 (7.2)

CMH = chronic mucus hypersecretion; Mean (standard deviation) shown for normally distributed continuous data and median (range) for non-normally distributed continuous data.

In all cohorts, individuals with CMH had lower lung function and were more often current smokers compared to individuals without CMH.

Genome wide analyses in NELSON participants with COPD

After QC, out of 620,901 SNPs 522,636 remained for GWA analysis in 849 individuals with COPD, 338 with and 511 without CMH. The QQ-plot showed no indication of population stratification ($\lambda = 1.002$). The p-values of the GWA study are presented in the Manhattan plot (Figure 1). 78 SNPs were associated with CMH at a $p < 2 \times 10^{-4}$. SNP rs626326 located in an intron in the *StAR-related lipid transfer (START) domain containing 13 gene* (*STARD13*) on chromosome 13q13.1 showed the strongest association with CMH ($p = 3.99 \times 10^{-6}$, OR = 1.632).

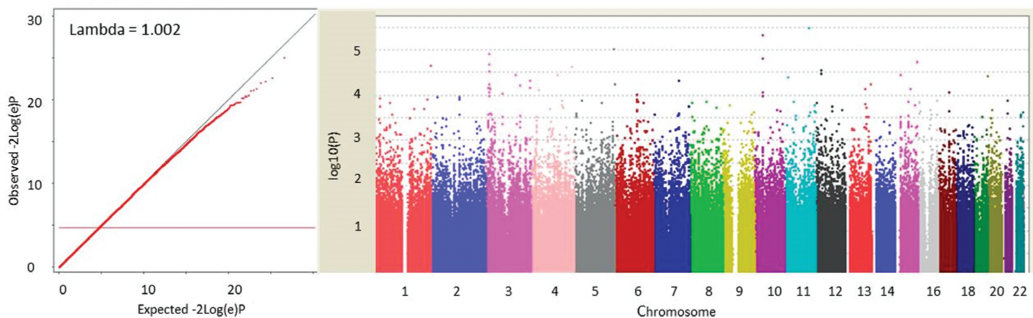


Figure 1. Quantile-quantile plot (left) and Manhattan plot (right) of GWA results for association of SNPs with CMH in NELSON participants with COPD.

Replication of top SNPs in four COPD cohorts

Table 3 shows the results of the 78 SNPs that were analyzed in 3,106 individuals with COPD, including 1,198 with and 1,908 without CMH, participating in 4 different COPD cohorts. Meta-analyses of these 78 SNPs across the replication cohorts showed borderline association to six SNPs with CMH and a similar direction of effect.

The strongest association in the meta-analysis was observed for rs10461985 on chromosome 5p13.2 showing effects in the same direction in all cohorts ($p = 5.43 \times 10^{-5}$, OR = 0.714), except for COPDGene that showed no effect. SNP rs10461985 is located in an intron in the glial cell line-derived neurotrophic factor antisense RNA 1 gene *(GDNF-AS1).*

Functional relevance of rs10461985 and GDNF

The Affymetrix chip used to investigate mRNA expression in airway wall biopsies of COPD patients did not have probe set for the *GDNF-AS1* gene. As the role of *GDNF-AS1* as an antisense RNA is to prevent translation of *GDNF*, we assessed the association of the mRNA expression of this gene and CMH. *GDNF* mRNA expression was found to be significantly lower in bronchial biopsies of COPD patients with CMH than those without CMH (b = -2.8, p = 0.007).

Table 3. Association of SNPs with CMH in identification analysis (NELSON-COPD) and in replication cohorts and subsequent meta-analysis across identification and replication cohorts.

CHR	SNP	NELSON-COPD				GenKOLS		COPDGene		ECLIPSE		MESA		Meta-analysis across identification and replication cohorts				Direction of effect	closest gene(s)
		rank	MAF	p	OR	p	OR	p	OR	p	OR	p	OR	rank	p#	OR#	Q		
1	rs2810587	33	0.23	9.90E-05	1.59	3.99E-01	1.10	3.10E-01	0.85	2.30E-01	0.90	6.49E-02	0.57	77	9.88E-01	1	<0.001	+ + - - -	*EDN2 & HIVEP3*
1	rs17518769	28	0.13	8.94E-05	2.03	1.49E-01	0.73	1.00E+00	1.00	3.00E-01	1.15	8.11E-02	0.55	70	8.59E-01	1.04	0.001	+ - 0 + -	*SLC19A2**
1	rs10753077	3	0.09	1.65E-05	1.79	4.95E-01	1.10	8.20E-01	1.05	6.70E-01	1.04	7.04E-01	1.15	14	5.44E-03	1.2	0.020	+ + + 0 +	*LOC730070 & PRDX6*
1	rs12410049	49	0.18	1.38E-04	1.79	7.96E-01	1.04	4.20E-01	0.84	2.90E-01	0.88	9.02E-01	0.96	61	6.43E-01	1.07	0.004	+ 0 - - -	*LOC100132728 & LOC100128723*
1	rs2001475	50	0.13	1.38E-04	1.79	7.96E-01	1.04	4.20E-01	0.84	2.90E-01	0.88	9.28E-01	0.97	60	6.37E-01	1.08	0.004	+ 0 - - 0	*LOC100132728 & LOC100128723*
1	rs3123695	36	0.12	1.08E-04	1.85	2.12E-01	0.78	7.40E-01	0.92	3.90E-01	0.90	6.49E-01	0.83	72	8.84E-01	1.03	0.002	+ - - - -	*ADSS**
2	rs4671197	63	0.33	1.67E-04	1.50	6.85E-01	0.96	3.90E-01	1.15	3.90E-01	1.07	5.82E-01	0.86	24	2.01E-02	1.13	0.030	+ 0 + + -	*ETAA1 & LOC402076*
2	rs216626	25	0.11	7.95E-05	1.89	2.44E-01	1.22	8.80E-01	1.03	2.50E-01	1.14	1.93E-01	0.67	13	4.94E-03	1.23	0.016	+ + 0 + -	*CTNNA2**
2	rs216640	59	0.11	1.55E-04	1.86	2.55E-01	1.21	8.40E-01	1.04	2.70E-01	1.13	1.84E-01	0.67	17	8.06E-03	1.21	0.020	+ + 0 + -	*CTNNA2**
2	rs3821072	20	0.10	6.69E-05	1.93	2.00E-01	1.25	7.90E-01	1.06	3.50E-01	1.11	1.89E-01	0.67	15	6.25E-03	1.22	0.013	+ + + + -	*CTNNA2**
2	rs6760631	68	0.21	1.78E-04	0.60	4.55E-01	0.91	5.00E-02	1.35	5.20E-01	1.06	4.37E-02	0.61	43	3.84E-01	0.88	<0.001	- - + + -	*MRPS18BP2 & LRP1B*
3	rs6442701	70	0.33	1.82E-04	0.66	7.29E-01	0.96	3.90E-01	0.88	9.50E-01	1.00	1.57E-01	1.45	32	5.92E-02	0.91	0.010	- 0 - 0 +	*LOC39150 4 &LOC727810*
3	rs6799163	73	0.33	1.90E-04	0.66	7.11E-01	0.96	4.70E-01	0.90	9.30E-01	0.99			25	2.44E-02	0.89	0.023	- 0 - 0 x	*LOC391504 & LOC727810*
3	rs492476	67	0.32	1.76E-04	0.64	1.14E-01	1.20	1.10E-01	1.28	7.90E-01	0.98	4.64E-01	1.24	73	9.28E-01	1.01	0.001	- + + - +	*GRM7 & LMCD1*
3	rs4420851	69	0.33	1.80E-04	0.65	1.20E-01	1.19	1.30E-01	1.26	6.70E-01	0.96	4.79E-01	1.23	78	9.95E-01	1	0.001	- + + 0 +	*GRM7 & LMCD1*
3	rs547906	39	0.33	1.13E-04	1.54	9.05E-01	0.99	7.00E-02	1.29	2.10E-01	0.90	9.57E-01	0.99	40	3.22E-01	1.12	0.002	+ 0 + - 0	*GRM7 & LMCD1*
3	rs12632517	29	0.32	9.02E-05	1.56	9.23E-01	1.01	1.00E-01	1.27	5.00E-02	0.85	9.28E-01	0.98	45	4.12E-01	1.11	<0.001	+ 0 + - 0	*GRM7 & LMCD1*
3	rs4515036	40	0.32	1.16E-04	1.55	9.76E-01	1.00	1.00E-01	1.27	4.00E-02	0.85	9.28E-01	0.98	46	4.31E-01	1.11	<0.001	+ 0 + - 0	*GRM7 & LMCD1*
3	rs3856798	66	0.13	1.74E-04	0.55	1.93E-01	1.21	5.50E-01	1.13	7.70E-01	1.03	2.33E-02	2.63	63	7.45E-01	1.09	<0.001	- + + 0 +	*VGLL4**

CHR	SNP	NELSON-COPD				GenKOLS		COPDGene		ECLIPSE		MESA		Meta-analysis across identification and replication cohorts				Direction of effect	closest gene(s)
		rank	MAF	p	OR	p	OR	p	OR	p	OR	p	OR	rank	p#	OR#	Q		
3	rs9831604	55	0.13	1.47E-04	0.55	1.73E-01	1.22	5.10E-01	1.14	8.40E-01	1.02	2.30E-02	2.62	67	7.94E-01	1.05	<0.001	- + + 0 +	*VGLL4**
3	rs339668	34	0.42	1.02E-04	1.51	1.61E-01	1.15	2.00E-02	0.71	8.20E-01	1.02	4.08E-01	0.81	65	7.58E-01	1.04	0.001	+ + - 0 -	*CCDC13**
3	rs12485872	27	0.13	8.24E-05	1.85	2.15E-01	0.84	6.70E-01	1.09	9.00E-01	1.01	5.27E-01	1.30	44	3.90E-01	1.21	0.003	+ - + 0 0	*LOC402152 & ATP11B*
4	rs4306981	12	0.31	4.40E-05	1.57	4.84E-02	1.25	6.70E-01	0.94	8.90E-01	0.99	1.32E-01	1.52	10	4.12E-03	1.16	0.005	+ + - 0 +	*PAQR3 & ARD1B*
5	rs7732527	43	0.44	1.25E-04	1.50	4.38E-01	1.08	8.00E-01	1.03	9.00E-01	1.01	7.12E-01	0.92	26	2.46E-02	1.12	0.033	+ + 0 0 -	*LOC646241 & LOC646273*
5	rs4867387	23	0.43	6.82E-05	1.73	4.28E-01	1.12	7.10E-01	0.92	6.50E-01	1.05	4.80E-01	1.27	16	7.70E-03	1.2	0.037	+ + - + +	*PDZD2**
5	rs11111	21	0.17	6.70E-05	0.56	7.72E-01	1.04	1.60E-01	0.76	2.40E-01	0.89	6.12E-01	0.84	8	2.74E-03	0.82	0.033	- 0 - - -	*GDNF*
5	rs10461985	71	0.18	1.82E-04	0.52	1.87E-01	0.78	9.80E-01	1.00	2.00E-02	0.74	3.70E-01	0.69	1	5.43E-05	0.71	0.228	- - 0 - -	*GDNF-AS1*
5	rs1501977	19	0.12	6.48E-05	0.62	1.94E-01	1.16	1.90E-01	0.81	6.00E-01	1.05	4.14E-01	0.78	39	3.13E-01	0.88	0.001	- + - + -	*PARP8 & LOC133569*
5	rs1229729	52	0.27	1.42E-04	0.66	4.91E-01	1.07	2.50E-01	1.17	1.90E-01	1.11	9.62E-01	1.01	71	8.80E-01	0.98	0.001	- + + + 0	*SPOCK1**
5	rs1229708	11	0.50	4.39E-05	1.54	8.06E-01	0.98	3.50E-01	0.88	7.60E-01	0.98	4.78E-01	1.19	48	4.48E-01	1.08	0.003	+ 0 - 0 +	*SPOCK1**
5	rs7736228	74	0.28	1.91E-04	0.64	5.68E-01	0.94	1.70E-01	0.81	2.80E-01	0.91	7.86E-01	1.08	5	1.94E-03	0.85	0.100	- - - - +	*LOC100132712 & ASSP10*
5	rs13178728	78	0.10	1.99E-04	1.91	8.49E-01	1.04	4.30E-01	1.22	9.70E-01	1.00	2.14E-01	1.80	21	1.59E-02	1.23	0.037	0 0 + 0 +	*LARP1**
5	rs13159558	56	0.06	1.49E-04	2.20	4.07E-01	1.18	7.50E-01	1.09	3.00E-01	0.87	4.90E-01	1.92	6	2.14E-03	1.48	0.101	+ + + - +	*LOC345471 & LOC100130177*
6	rs7751774	22	0.13	6.77E-05	0.52	2.06E-01	0.82	5.40E-01	0.88	7.50E-01	0.96	3.32E-01	0.72	7	2.23E-03	0.8	0.049	- - - 0 -	*PECI & LOC100129052*
6	rs1360811	14	0.13	5.80E-05	0.51	2.83E-01	0.84	4.10E-01	0.85	4.40E-01	0.92	4.82E-01	0.79	4	1.50E-03	0.8	0.062	- - - - -	*LOC100129052 & KU-MEL-3*
6	rs9503979	15	0.13	5.80E-05	0.51	2.88E-01	0.85	4.10E-01	0.84	4.10E-01	0.91	4.83E-01	0.79	3	1.13E-03	0.79	0.070	- - - - -	*LOC100129052 & KU-MEL-3*
6	rs6933317	31	0.49	9.44E-05	1.49	5.91E-01	0.95	6.90E-01	1.06	4.80E-01	1.06	8.54E-01	0.96	28	3.09E-02	1.11	0.020	+ - + + -	*PRL & HDGFL1*
6	rs6940071	13	0.41	5.66E-05	1.52	9.38E-01	0.99	6.80E-01	1.06	1.30E-01	1.13	8.05E-01	0.94	9	3.46E-03	1.16	0.036	+ 0 + + -	*PRL & HDGFL1*

CHR	SNP	NELSON-COPD				GenKOLS		COPDGene		ECLIPSE		MESA		Meta-analysis across identification and replication cohorts				Direction of effect	closest gene(s)
		rank	MAF	p	OR	p	OR	p	OR	p	OR	p	OR	rank	p#	OR#	Q		
6	rs12527846	53	0.21	1.42E-04	0.67	8.97E-01	0.99	7.70E-01	0.96	3.70E-01	0.93	8.92E-01	1.04	20	1.36E-02	0.86	0.037	- 0 0 - 0	*PRL & HDGFL1*
6	rs12211633	76	0.42	1.95E-04	0.64	5.54E-01	0.94	7.20E-01	1.06	6.30E-01	1.04	2.18E-01	1.48	38	2.10E-01	0.94	0.006	- - + 0 +	*GRM4**
6	rs2682185	51	0.47	1.38E-04	2.04	7.78E-01	1.05	9.90E-01	1.00	4.40E-01	1.11	4.50E-01	0.73	27	2.69E-02	1.21	0.028	+ + 0 + -	*B3GAT2 & LOC157713*
6	rs164301	8	0.18	3.82E-05	0.64	9.34E-01	1.01	4.20E-01	1.12	8.70E-01	0.99	7.29E-01	1.09	51	5.14E-01	0.94	0.004	- 0 + 0 +	*EPHA7**
6	rs9365242	5	0.21	2.55E-05	0.55	4.29E-01	0.91	5.20E-01	1.12	9.80E-01	1.00	9.84E-01	1.01	29	4.04E-02	0.88	0.006	- - + 0 0	*ZDHHC14**
6	rs12055716	24	0.14	7.26E-05	0.59	5.95E-01	0.94	7.10E-01	1.06	5.40E-01	0.95	7.32E-01	1.11	23	1.97E-02	0.84	0.013	- - + - +	*ZDHHC14**
6	rs9295312	17	0.08	5.96E-05	1.84	7.19E-01	0.95	6.10E-01	0.91	2.90E-01	0.89	7.20E-01	1.13	54	5.64E-01	1.09	0.002	+ - - - +	*PDE10A & C6orf176*
8	rs4875186	42	0.10	1.23E-04	1.91	8.46E-01	0.97	6.80E-01	1.09	2.80E-01	0.87	8.81E-01	0.95	50	4.93E-01	1.12	0.004	+ 0 + - -	*LOC648237 & LOC392180*
8	rs7830870	16	0.21	5.81E-05	1.67	7.27E-01	1.04	1.00E-01	1.32	7.40E-01	1.03	6.98E-01	1.14	12	4.81E-03	1.18	0.024	+ 0 + 0 +	*LOC100133073 & NAT1*
8	rs1864773	7	0.13	2.90E-05	1.88	9.14E-01	1.02	9.80E-01	0.99	8.80E-01	0.98	6.34E-01	1.18	31	4.62E-02	1.15	0.008	+ 0 0 0 +	*KCNQ3**
8	rs7840848	37	0.43	1.10E-04	1.51	6.09E-01	1.05	5.60E-01	1.08	5.20E-01	0.95	4.29E-01	0.82	35	8.90E-02	1.09	0.008	+ + + - -	*ZFAT & LOC286094*
8	rs2289001	46	0.33	1.33E-04	1.53	8.58E-01	1.02	6.80E-01	1.07	3.30E-01	0.92	2.68E-01	1.38	37	1.27E-01	1.08	0.005	+ 0 + - +	*DENND3**
11	rs6483640	75	0.49	1.93E-04	1.47	1.97E-01	1.14	5.80E-01	1.08	8.50E-01	1.02	7.15E-01	1.11	11	4.63E-03	1.15	0.088	+ + + 0 +	*NAV2**
11	rs2217032	54	0.49	1.43E-04	1.51	6.22E-01	1.05	3.00E-01	1.15	1.20E-01	1.13	9.30E-01	0.98	2	1.05E-03	1.18	0.119	+ + + + -	*LOC100130747 & LOC554234*
11	rs2292730	48	0.44	1.36E-04	0.67	8.59E-01	0.98	2.50E-01	0.85	4.60E-01	1.06	7.80E-02	1.61	56	5.89E-01	0.94	0.002	- 0 - + +	*MMP20**
11	rs7935816	18	0.33	6.40E-05	0.63	1.64E-01	1.17	9.10E-01	0.98	1.40E-01	1.13	5.43E-01	0.84	59	6.36E-01	0.94	<0.001	- + 0 + -	*PVRL1 & LOC390255*
12	rs7304675	77	0.41	1.95E-04	0.66	9.16E-01	0.99	8.90E-01	0.98	5.00E-01	1.05	1.13E-02	2.17	75	9.54E-01	0.99	0.001	- 0 0 + +	*DCN & BTG1*
12	rs812512	35	0.39	1.07E-04	1.51	7.33E-01	0.97	7.90E-01	0.96	1.00E-02	0.81	3.94E-01	0.79	76	9.85E-01	1	<0.001	+ - - - -	*CCDC60 & LOC387890*
13	rs495680	6	0.36	2.78E-05	0.63	4.08E-02	1.24	9.60E-01	1.01	6.00E-01	0.96	9.63E-01	1.01	58	6.30E-01	0.94	<0.001	- + 0 0 0	*STARD13**

CHR	SNP	NELSON-COPD				GenKOLS		COPDGene		ECLIPSE		MESA		Meta-analysis across identification and replication cohorts				Direction of effect	closest gene(s)
		rank	MAF	p	OR	p	OR	p	OR	p	OR	p	OR	rank	p#	OR#	Q		
13	rs2858808	4	0.31	1.79E-05	0.60	5.85E-01	1.06	4.10E-01	0.88	7.30E-01	1.03	3.74E-01	1.25	49	4.82E-01	0.92	0.001	- + - 0 +	*STARD13**
13	rs523523	2	0.50	1.32E-05	0.64	3.31E-01	1.10	1.60E-01	1.22	8.70E-01	0.99	8.83E-01	1.04	64	7.49E-01	0.96	<0.001	- + + 0 0	*STARD13**
13	rs2697092	57	0.22	1.49E-04	1.62	3.34E-01	1.12	3.30E-01	0.84	3.80E-01	1.09	9.15E-01	1.03	18	1.13E-02	1.16	0.029	+ + - + 0	*LOC100130462 & FLJ40296*
15	rs8041061	61	0.48	1.60E-04	1.47	8.00E-01	1.03	5.60E-01	1.08	9.40E-01	0.99	2.67E-01	0.76	34	6.83E-02	1.09	0.014	- 0 - 0 +	*RORA**
15	rs809736	62	0.26	1.62E-04	0.64	9.12E-01	1.01	4.20E-01	0.87	8.10E-01	0.98	5.78E-01	1.17	30	4.35E-02	0.89	0.024	- 0 - 0 +	*RORA**
18	rs8088174	72	0.18	1.87E-04	1.64	3.77E-02	0.76	8.30E-01	0.96	4.70E-01	0.93	8.24E-01	1.08	68	8.32E-01	1.03	0.001	+ - 0 - +	*LOC388474 & KC6*
20	rs6085660	10	0.42	4.03E-05	1.55	2.42E-01	0.89	9.10E-01	0.98	1.10E-01	1.13	9.41E-01	0.98	42	3.69E-01	1.1	0.004	+ - 0 + 0	*LOC728383**
20	rs1500545	60	0.41	1.59E-04	1.49	2.86E-01	0.90	9.90E-01	1.00	2.50E-01	1.09	6.86E-01	0.91	33	6.50E-02	1.1	0.010	+ - 0 + -	*LOC728383**
20	rs6055258	58	0.40	1.53E-04	0.67	2.57E-01	0.89	4.00E-02	1.34	2.70E-01	0.92	5.68E-01	1.16	66	7.87E-01	0.96	0.001	- - + - +	*BMP2 & FUSIP1P2*
20	rs969111	45	0.39	1.27E-04	0.67	2.76E-01	0.90	4.00E-02	1.34	2.60E-01	0.92	4.90E-01	1.19	57	5.99E-01	0.94	0.002	- - + - +	*BMP2 & FUSIP1P2*
20	rs1008096	44	0.39	1.26E-04	0.67	2.41E-01	0.89	4.00E-02	1.34	2.70E-01	0.92	4.85E-01	1.20	55	5.89E-01	0.94	0.002	- - + - +	*BMP2 & FUSIP1P2*
20	rs6118681	38	0.34	1.12E-04	1.51	2.46E-01	0.89	4.20E-01	1.13	1.40E-01	0.89	6.16E-01	1.14	52	5.25E-01	1.08	0.001	+ - + - +	*PAK7**
20	rs6141026	9	0.20	3.98E-05	1.69	5.32E-01	0.93	5.60E-01	1.11	4.30E-01	1.08	7.41E-01	1.10	22	1.73E-02	1.16	0.013	+ - + + +	*PAK7**
20	rs6081741	65	0.25	1.71E-04	0.63	9.73E-01	1.00	6.00E-01	1.08	7.80E-01	0.98	6.74E-01	1.14	36	1.05E-01	0.91	0.018	- 0 + 0 +	*LOC100130408**
20	rs6013773	41	0.44	1.18E-04	0.67	8.80E-01	1.02	1.90E-01	1.20	2.40E-01	1.09	6.22E-01	0.88	62	6.94E-01	0.96	0.002	- 0 + + -	*ZNF217 & LOC100128404*
23	rs5927035	32	0.44	9.52E-05	1.78	1.76E-01	0.85			9.10E-01	0.99			53	5.34E-01	1.13	<0.001	+ - x 0 x	*IL1RAPL1**
23	rs2879751	26	0.44	8.10E-05	1.79					9.90E-01	1.00			41	3.24E-01	1.33	0.003	+ x x 0 x	*IL1RAPL1**

CMH is chronic mucus hypersecretion; MAF is minor allele frequency; OR is odds ratio; Q = p-value for heterogeneity; p# = fixed p-value if p-value for heterogeneity > 0.005 and random p-value if p-value for heterogeneity < 0.005; OR#=fixed OR if p-value for heterogeneity > 0.005 and random OR if p-value for heterogeneity < 0.005; Direction of effect in identification and replication cohorts is presented in the following order: NELSON-COPD, GenKOLS, COPDGene, ECLIPSE and MESA; Direction of effect: - = OR ≤ 0.95, 0 = 0.95 ≤ OR ≤ 1.05, 1 = OR ≥ 1.05, x = not applicable; An empty box = SNP was not analyzed in the corresponding replication cohort.

4

Demographic and clinical characteristics of NELSON participants without COPD and the replication cohort LifeLines are presented in Table 3.

Table 3. Characteristics of individuals with and without CMH, in NELSON-non-COPD and in the Lifelines cohort.

	NELSON		LifeLines	
	+ CMH	- CMH	+ CMH	- CMH
N, (%)	342 (25.4)	1,006 (74.6)	130 (5.3)	2,313 (94.7)
Age, yrs	59.6 (5.3)	59.8 (5.3)	47.2 (10.7)	47.4 (9.7)
Female, %	0	0	46.2	53.4
Packyears	38.0 (22-140)	34.2 (20-133)	15.5 (5-84)	13.0 (5-75)
Current smoking, %	70.8	45.2	60.0	43.1
FEV_1, %predicted	105.2 (13.1)	107.6 (13.4)	100.5 (14.2)	103.6 (12.8)
FEV_1/FVC, %	78.0 (4.6)	78.1 (4.5)	77.1 (4.4)	78.0 (4.8)

CMH = chronic mucus hypersecretion; Mean (standard deviation) shown for normally distributed continuous data and median (range) for non-normally distributed continuous data.

Genome wide analyses in NELSON participants without COPD

The same 522,636 SNPs were analyzed in 1,348 NELSON participants without, 342 with and 1,006 without CMH. The QQ-plot confirmed that there was no population stratification ($\lambda = 1.009$). The p-values of this GWA study are presented in the Manhattan plot (Figure 2). There were 79 SNPs associated with CMH with a $p < 2.0 \times 10^{-4}$ (Table 4).

Replication of top SNPs in the general population based LifeLines cohort

Genotypes from 74 of the 79 SNPs with a $p < 2.0 \times 10^{-4}$ were available from the general population based LifeLines cohort, including 130 individuals with CMH and 2,313 without CMH. Ten SNPs showed some association with CMH in LifeLines ($p < 10^{-1}$) and among these, 7 SNPs had effects in the same direction in the NELSON participants without COPD and in LifeLines (Table 6). In the meta-analysis across this NELSON population and LifeLines 4 SNPs were associated with CMH with a $p < 10^{-5}$:

- rs3845529 on chromosome 1q41; p = 3.25 x 10^{-6} (OR = 0.693), located in an intron in the *Usher syndrome 2A gene* (*USH2A*);
- rs1690139 on chromosome 12q; p = 5.91 x 10^{-6} (OR = 1.673), located in a gene desert between LOC100130336 and LOC100131830;
- rs4863687 on chromosome 4q28; p = 7.57 x 10^{-6} (OR = 1.476), located in an intron in the *mastermind-like 3 gene* (*MAML3);*
- rs944899 on chromosome 13q34; p = 8.40 x 10^{-6} (OR = 1.399), located near (< 25 kb) the *SRY (sex determining region Y)-box 1 gene* (*SOX1*).

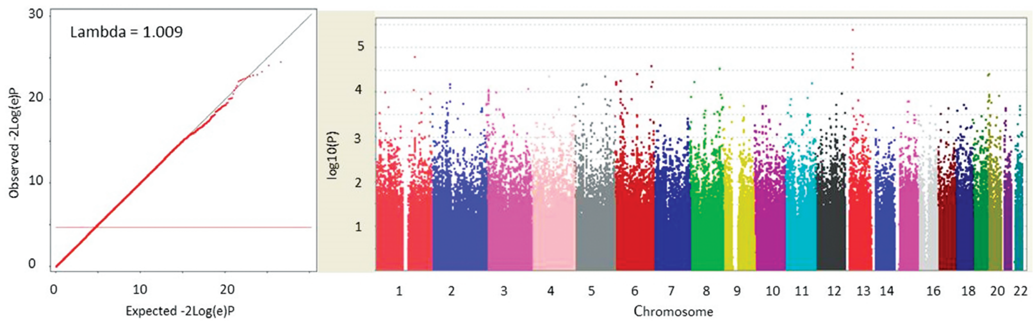


Figure 2. Quantile-quantile plot (left) and Manhattan plot (right) of GWA results for association of SNPs with CMH in NELSON participants without COPD.

Table 4. Association of SNPs with CMH in identification analysis (NELSON-non-COPD) and in replication in LifeLines and subsequent meta-analysis across NELSON-non-COPD and LifeLines.

CHR	SNP	NELSON-non-COPD				LifeLines			META-ANALYSIS			Closest gene(s)
		MAF	rank	P	OR	P	OR	rank	P#	OR#	Q	
1	rs2817896	0.26	59	1.16E-04	1.47	1.09E-01	1.26	8	4.66E-05	1.40	0.362	*EPHB2**
1	rs893961	0.25	66	1.81E-04	1.46	8.86E-02	1.28	9	5.30E-05	1.39	0.445	*EPHB2**
1	rs11208807	0.31	57	1.50E-04	1.43	2.55E-01	1.17	23	1.65E-04	1.34	0.228	*PDE4B**
1	rs2208370	0.39	53	1.98E-04	1.42	7.22E-01	1.07	35	5.51E-04	1.33	0.154	*DNM3**
1	rs3845529	0.42	73	1.96E-04	0.7	4.98E-03	0.67	1	3.25E-06	0.69	0.780	*USH2A**
1	rs629199	0.19	65	1.24E-04	1.54	3.64E-01	1.25	17	1.10E-04	1.48	0.445	*IRF2BP2 & PP2672*
1	rs12028329	0.25	46	2.20E-05	1.55	6.74E-01	1.07	21	1.47E-04	1.39	0.052	*LOC441931 & VN1R5*
2	rs1476151	0.46	19	1.08E-04	1.43	5.37E-01	0.91	62	2.98E-03	1.26	0.010	*CNTP5 & LOC150554*
2	rs13028050	0.42	29	1.25E-04	0.7	7.36E-01	1.05	61	2.71E-03	0.79	0.016	*CNTP5 & LOC150554*
3	rs17776719	0.13	42	6.72E-05	1.64	5.58E-01	0.84	34	5.49E-04	1.49	0.038	*VGLL4**
3	rs2956507	0.35	21	6.61E-05	0.68	7.82E-01	1.04	56	2.06E-03	0.78	0.011	*FBLN2 & WNT7A*
3	rs6792244	0.42	28	5.77E-05	0.68	6.74E-01	1.07	49	1.28E-03	0.77	0.014	*FBLN2 & WNT7A*
3	rs6775581	0.42	16	1.22E-05	0.66	6.80E-01	1.07	30	4.24E-04	0.75	0.009	*FBLN2 & WNT7A*
3	rs6781368	0.43	14	2.02E-05	0.67	8.42E-01	1.03	42	8.12E-04	0.77	0.008	*FBLN2 & WNT7A*
3	rs6794344	0.46	24	8.84E-05	0.7	7.82E-01	1.04	59	2.51E-03	0.80	0.012	*FBLN2 & WNT7A*
3	rs6795216	0.46	41	1.06E-04	0.7	9.03E-01	1.02	47	1.13E-03	0.77	0.035	*FBLN2 & WNT7A*
3	rs2974399	0.45	30	2.89E-05	0.68	7.99E-01	1.04	33	5.38E-04	0.76	0.018	*FBLN2 & WNT7A*
3	rs6768597	0.3	50	7.05E-05	0.66	3.17E-01	0.87	20	1.44E-04	0.73	0.125	*SGOL1 & VENTXP7*
3	rs9682418	0.27	70	9.15E-05	1.48	4.91E-02	1.32	5	1.52E-05	1.43	0.494	*PROK2 & CCDC137P*
3	rs11714053	0.17	37	3.49E-05	1.61	5.06E-01	0.84	28	3.74E-04	1.46	0.026	*CPNE4 & LOC729674*
3	rs9825199	0.06	17	4.83E-05	2.02	4.88E-01	0.81	50	1.38E-03	1.62	0.009	*C3orf21**
3	rs3796160	0.06	22	6.76E-05	2	5.17E-01	0.82	52	1.74E-03	1.60	0.011	*C3orf21**
4	rs17447715	0.19	58	1.94E-04	0.62	1.52E-01	0.78	18	1.16E-04	0.67	0.295	*OR7E94P & GDEP*
4	rs6858670	0.47	32	1.29E-04	1.42	9.08E-01	0.99	57	2.13E-03	1.26	0.022	*LOC100132574 & LOC646316*
4	rs7688325	0.47	35	1.65E-04	1.41	8.99E-01	0.98	60	2.54E-03	1.25	0.024	*LOC100132574 & LOC646316*
4	rs4863687	0.28	72	1.89E-04	1.45	1.22E-02	1.57	3	7.57E-06	1.48	0.688	*MAML3**
4	rs6552407	0.25	1	2.38E-05	1.55	7.85E-02	0.76	73	8.04E-01	1.09	0.000	*LOC391719 & hCG_2025798*
5	rs1816237	0.11	49	1.27E-04	0.53	8.00E-01	0.93	32	5.09E-04	0.61	0.102	*LOC340113 & LOC728553*
5	rs4836527	0.4	33	1.45E-04	1.41	5.38E-01	0.9	54	1.96E-03	1.28	0.022	*PRDM6 & CEP120*
5	rs13183447	0.39	4	9.28E-06	0.65	3.04E-01	1.17	70	6.13E-01	0.86	0.001	*SH3PXD2B & LOC100130394*
5	rs262020	0.39	54	5.78E-05	0.68	8.99E-01	0.97	24	1.68E-04	0.71	0.154	*COL23A1**
6	rs7770889	0.37	60	9.92E-05	1.45	3.65E-01	1.19	13	9.81E-05	1.40	0.368	*FUT9 & KIAA0776*
6	rs9486181	0.36	63	1.30E-04	1.45	2.82E-01	1.22	14	1.03E-04	1.40	0.410	*FUT9 & KIAA0776*
6	rs4425602	0.36	61	1.30E-04	1.45	2.93E-01	1.21	16	1.08E-04	1.39	0.396	*FUT9 & KIAA0776*
6	rs3860243	0.36	62	1.21E-04	1.45	2.79E-01	1.22	12	9.32E-05	1.40	0.402	*FUT9 & KIAA0776*
6	rs12207471	0.36	47	1.30E-04	1.45	9.17E-01	1.02	43	8.20E-04	1.32	0.064	*FUT9 & KIAA0776*
6	rs9398148	0.34	64	1.39E-04	1.45	2.97E-01	1.23	15	1.05E-04	1.40	0.442	*FHL5**
6	rs9375195	0.48	40	1.35E-04	1.42	9.58E-01	1.01	53	1.78E-03	1.26	0.029	*C6orf167 & LOC100129158*
7	rs10499977	0.33	31	4.81E-05	1.48	6.02E-01	0.91	41	7.41E-04	1.34	0.020	*LOC646614 & LOC100128056*

CHR	SNP	NELSON-non-COPD				LifeLines			META-ANALYSIS across NELSON-non-COPD and LifeLines			Closest gene(s)
		MAF	rank	P	OR	P	OR	rank	P#	OR#	Q	
8	rs7007974	0.1	56	1.48E-04	1.69	2.75E-01	1.24	25	1.82E-04	1.53	0.208	*MRPS18CP2 & LOC645960*
8	rs13265648	0.49	2	1.38E-04	0.7	8.67E-02	1.25	72	7.98E-01	0.93	0.000	*TRPA1 & LOC392232*
8	rs16886291	0.12	44	1.90E-04	0.55	6.96E-01	0.92	51	1.46E-03	0.67	0.047	*hCG_1644355 & TRPS1*
9	rs10119913	0.3	3	1.61E-04	0.68	5.54E-02	1.5	74	9.74E-01	0.99	0.001	*LINGO2 & LOC286239*
10	rs10827563	0.48	38	1.04E-04	1.43	5.15E-01	0.88	48	1.14E-03	1.31	0.027	*LOC439954 & PBEF2*
10	rs2696310	0.44	7	1.55E-05	1.5	6.65E-01	0.95	68	4.27E-01	1.20	0.004	*LOC439954 & PBEF2*
10	rs2767073	0.44	8	4.75E-06	1.54	5.86E-01	0.92	26	2.21E-04	1.35	0.006	*LOC439954 & PBEF2*
10	rs1571136	0.44	18	1.57E-05	1.5	6.14E-01	0.92	31	4.56E-04	1.33	0.010	*LOC439954 & PBEF2*
10	rs2804852	0.42	39	8.39E-05	1.44	6.53E-01	0.92	45	1.01E-03	1.31	0.028	*LOC439954 & PBEF2*
11	rs2071461	0.24	26	3.86E-05	1.52	3.12E-01	0.78	37	6.06E-04	1.38	0.013	*GALNTL4**
11	rs3903687	0.37	10	1.40E-04	1.43	4.90E-01	0.91	67	6.03E-03	1.24	0.006	*SLC1A2*
11	rs474158	0.07	36	3.28E-06	2.17	7.05E-01	1.1	7	4.35E-05	1.76	0.024	*GRIA4**
11	rs2288403	0.17	71	1.63E-04	0.6	6.27E-02	0.69	6	3.00E-05	0.63	0.604	*NFRKB**
12	rs10459134	0.18	13	1.47E-04	1.55	5.12E-01	0.89	65	5.21E-03	1.31	0.008	*TMEM16B**
12	rs7959932	0.32	9	2.74E-05	1.49	2.08E-01	0.74	39	6.34E-04	1.35	0.006	*SOX5**
12	rs7308636	0.31	15	3.27E-05	1.48	2.34E-01	0.75	38	6.25E-04	1.35	0.008	*SOX5**
12	rs1690139	0.11	74	1.76E-04	1.67	1.11E-02	1.69	2	5.91E-06	1.67	0.951	*LOC100130336 & LOC100131830*
13	rs4514531	0.29	23	7.12E-05	0.66	6.32E-01	1.08	55	1.99E-03	0.76	0.011	*LOC100130117 & hCG_1795283*
13	rs944899	0.46	69	5.76E-05	1.46	4.05E-02	1.3	4	8.40E-06	1.40	0.476	*SOX1*
15	rs12594495	0.26	6	3.44E-05	0.62	5.49E-01	1.09	69	4.71E-01	0.82	0.002	*CYFIP1**
15	rs8042800	0.3	5	1.36E-04	0.67	2.60E-01	1.17	71	6.39E-01	0.88	0.001	*FAM81A & GCNT3*
15	rs3784350	0.37	11	7.25E-05	0.68	6.38E-01	1.07	63	3.47E-03	0.79	0.006	*ITGA11**
15	rs1348533	0.2	12	1.67E-04	0.63	4.36E-01	1.17	66	5.73E-03	0.75	0.008	*AGBL1*
15	rs8043332	0.3	20	1.85E-05	1.51	3.68E-01	0.82	29	3.84E-04	1.36	0.011	*FAM169B & IGF1R*
16	rs1978316	0.19	67	1.44E-04	1.53	1.85E-01	1.29	11	7.70E-05	1.46	0.448	*A2BP1**
16	rs1344471	0.19	68	1.36E-04	1.53	1.84E-01	1.29	10	7.31E-05	1.47	0.449	*A2BP1**
16	rs12443545	0.19	45	1.31E-04	0.62	5.94E-01	1.18	44	8.58E-04	0.68	0.051	*CDH13**
16	rs12918351	0.2	43	1.30E-04	0.62	9.35E-01	0.98	46	1.12E-03	0.71	0.044	*CDH13**
17	rs1508960	0.3	25	8.74E-05	1.45	7.06E-01	0.95	58	2.36E-03	1.27	0.012	*LOC645163 & LOC645173*
20	rs6042209	0.18	34	3.64E-05	1.59	9.79E-01	1	36	5.69E-04	1.38	0.023	*FKBP1A & NSFL1C*
21	rs2032257	0.39	51	1.30E-04	0.69	3.58E-01	0.88	27	2.78E-04	0.75	0.131	*APP & CYYR1*

CMH is chronic mucus hypersecretion; OR is odds ratio; Q = p-value for heterogeneity; p# = fixed p-value if p-value for heterogeneity > 0.005 and random p-value if p-value for heterogeneity < 0.005; OR# = fixed OR if p-value for heterogeneity > 0.005 and random OR if p-value for heterogeneity < 0.005; Direction of effect in identification and replication cohorts is presented in the following order: NELSON-non-COPD, LifeLines; Direction of effect: - = OR ≤ 0.95, 0 = 0.95 ≤ OR ≤ 1.05, 1 = OR ≥ 1.05, x = not applicable.

Table 5. Comparison of SNPs associated with CMH and p-value < 10^{-2} present in NELSON-COPD and NELSON-non-COPD

CHR	SNP	BP	minor allele	NELSON-COPD				NELSON-non-COPD				Direc-tion of effect	in or close to gene(s)
				MAF	rank	P	OR	MAF	rank	p	OR		
1	rs6677529	160530378	A	0.19	48	7.24E-03	1.42	0.17	10	1.03E-03	1.45	+ +	*NOS1AP**
3	rs12632852	11593682	G	0.40	2	3.20E-04	0.67	0.39	52	8.70E-03	1.28	- +	*VGLL4**
3	rs2574704	11630381	G	0.29	26	3.94E-03	0.72	0.29	4	5.25E-04	1.40	- +	*VGLL4**
3	rs2574720	11635412	C	0.26	7	1.08E-03	0.68	0.26	3	3.97E-04	1.43	- +	*VGLL4**
3	rs2616551	11642123	G	0.18	54	7.91E-03	0.69	0.18	2	3.57E-04	1.50	- +	*VGLL4**
3	rs12374151	16605508	A	0.12	18	2.83E-03	0.61	0.13	48	7.25E-03	1.43	- +	*DAZL**
3	rs9852824	24397993	A	0.46	50	7.51E-03	1.32	0.46	60	9.90E-03	0.79	+ -	*THRB**
3	rs3796150	66584924	A	0.20	55	8.54E-03	0.70	0.17	32	4.73E-03	0.70	- -	*LRIG1**
3	rs7648171	106704936	G	0.20	41	6.16E-03	0.70	0.21	36	6.03E-03	0.73	- -	*ALCAM**
4	rs4306981	80143145	G	0.31	1	4.40E-05	1.57	0.30	6	5.73E-04	1.40	+ +	*PAQR3 & ARD1B*
4	rs10518211	80156089	G	0.48	21	3.50E-03	1.35	0.48	20	1.93E-03	1.33	+ +	*PAQR3 & ARD1B*
4	rs4834752	120275247	A	0.42	12	1.97E-03	0.72	0.44	15	1.30E-03	1.34	- +	*MYOZ2**
4	rs1017710	180937258	A	0.07	5	9.14E-04	1.97	0.07	37	6.23E-03	0.58	+ -	*LOC391719 & hCG_2025798*
4	rs17068194	180952052	A	0.07	6	9.14E-04	1.97	0.07	41	6.71E-03	0.58	+ -	*LOC391719 & hCG_2025798*
5	rs365294	3476838	A	0.38	45	6.74E-03	1.34	0.37	8	7.47E-04	1.38	+ +	*LOC100132531 & IRX1*
5	rs1995385	73415681	G	0.23	4	6.71E-04	0.65	0.23	58	9.39E-03	1.32	- +	*RGNEF & ENC1*
5	rs718164	73417137	G	0.23	3	5.37E-04	0.64	0.23	57	9.37E-03	1.32	- +	*RGNEF & ENC2*
5	rs11738681	176694141	G	0.33	43	6.35E-03	0.74	0.32	43	6.79E-03	0.76	- -	*LMAN2**
5	rs11949401	176698595	G	0.33	36	5.26E-03	0.73	0.31	53	8.76E-03	0.76	- -	*LMAN2**
5	rs9313758	176705697	C	0.33	44	6.35E-03	0.74	0.31	42	6.76E-03	0.76	- -	*LMAN2**
5	rs4532376	176707009	A	0.33	33	4.86E-03	0.73	0.31	33	5.13E-03	0.75	- -	*LMAN2**
5	rs4131289	176713151	A	0.33	40	5.88E-03	0.74	0.31	29	4.15E-03	0.74	- -	*LMAN2 & RGS14*
7	rs40463	40915342	A	0.12	24	3.65E-03	1.55	0.13	51	8.30E-03	0.68	+ -	*C7orf10 &INHBA*
7	rs4729686	100747270	A	0.07	13	2.18E-03	0.50	0.07	22	2.76E-03	1.67	- +	*RABL5**
7	rs2905286	112081312	G	0.48	57	9.04E-03	0.76	0.48	39	6.56E-03	0.78	- -	*NPM1P14 & LOC100128875*
8	rs2055516	769714	C	0.25	11	1.85E-03	1.46	0.25	14	1.27E-03	1.40	+ +	*C8orf68**
8	rs10105558	783149	A	0.25	27	4.04E-03	1.42	0.25	28	3.65E-03	1.35	+ +	*C8orf68**
8	rs13282923	4473969	G	0.29	29	4.10E-03	1.38	0.29	18	1.82E-03	0.72	+ -	*CSMD1**
8	rs13273819	135514435	A	0.23	35	5.25E-03	1.39	0.23	54	9.15E-03	1.32	+ +	*LOC100129104 & ZFAT*
9	rs530582	134354849	G	0.15	17	2.76E-03	0.64	0.17	7	6.63E-04	1.49	- +	*RP11-738I14.8**
10	rs10903396	1208030	G	0.46	28	4.06E-03	0.74	0.46	38	6.26E-03	0.78	- -	*C10orf139 & LOC100130729*
10	rs10905113	7246430	G	0.44	8	1.14E-03	1.41	0.44	50	8.12E-03	0.79	+ -	*SFMBT2**
10	rs17601717	52831431	G	0.23	39	5.38E-03	0.71	0.25	40	6.57E-03	1.32	- +	*PRKG1**
10	rs7902476	72693742	A	0.11	25	3.70E-03	0.60	0.12	26	3.37E-03	0.64	- -	*UNC5B**
11	rs2273688	35295319	A	0.27	31	4.49E-03	0.71	0.28	16	1.56E-03	1.40	- +	*SLC1A2**
11	rs10768129	35319065	A	0.27	47	7.02E-03	0.72	0.28	13	1.21E-03	1.40	- +	*SLC1A2**
11	rs7127824	35330427	A	0.27	22	3.64E-03	0.70	0.28	11	1.14E-03	1.40	- +	*SLC1A2**
11	rs7130967	35330584	A	0.27	23	3.64E-03	0.70	0.28	12	1.14E-03	1.40	- +	*SLC1A2**
11	rs927352	35334090	A	0.30	58	9.36E-03	0.73	0.31	19	1.90E-03	1.36	- +	*SLC1A2**
11	rs11033910	37021958	G	0.28	53	7.82E-03	0.73	0.29	56	9.32E-03	1.30	- +	*C11orf74 & LOC100129825*
11	rs12417575	85832165	G	0.28	37	5.31E-03	0.72	0.27	59	9.85E-03	0.76	- -	*ME3**
11	rs689051	124797700	A	0.16	10	1.43E-03	1.58	0.15	30	4.40E-03	0.67	+ -	*PKNOX2**

CHR	SNP	BP	minor allele	NELSON-COPD				NELSON-non-COPD				Direc-tion of effect	in or close to gene(s)
				MAF	rank	P	OR	MAF	rank	p	OR		
12	rs1894307	11896987	A	0.15	34	4.90E-03	1.49	0.14	9	9.39E-04	1.50	+ +	*ETV6**
12	rs2855708	11904839	G	0.28	30	4.10E-03	1.40	0.27	34	5.40E-03	1.31	+ +	*ETV6**
12	rs1820545	39096860	G	0.41	38	5.32E-03	0.75	0.42	31	4.47E-03	1.29	- +	*LRRK2 & MUC19*
12	rs7306163	39111184	C	0.41	42	6.21E-03	0.75	0.42	35	5.50E-03	1.28	- +	*MUC19**
14	rs8009673	31412453	A	0.14	46	7.00E-03	1.50	0.13	21	2.23E-03	1.49	+ +	*NUBPL & C14orf128*
14	rs7155416	76021126	A	0.12	51	7.72E-03	1.51	0.14	23	3.02E-03	1.46	+ +	*ESRRB**
14	rs9323838	88789353	G	0.37	56	8.68E-03	1.33	0.38	49	7.94E-03	0.78	+ -	*FOXN3**
15	rs1531636	92404552	A	0.36	14	2.36E-03	1.40	0.34	44	7.05E-03	1.28	+ +	*LOC283682 & LOC100129642*
16	rs7202333	67438996	G	0.39	32	4.76E-03	0.73	0.37	47	7.24E-03	0.77	- -	*TMCO7**
16	rs7184633	81379514	A	0.40	19	2.93E-03	0.73	0.40	1	2.67E-04	0.71	- -	*CDH13**
19	rs10411733	62482800	A	0.47	16	2.60E-03	0.73	0.46	25	3.29E-03	1.31	- +	*ZNF460**
20	rs2224326	19689491	A	0.23	9	1.31E-03	0.66	0.24	46	7.15E-03	1.31	- +	*LOC100130408**
20	rs4811610	53652782	G	0.29	60	9.92E-03	1.33	0.31	45	7.11E-03	0.76	+ -	*RPL12P4 &CBLN4*
22	rs2073760	17886456	A	0.40	49	7.33E-03	1.32	0.40	24	3.20E-03	0.76	+ -	*CDC45L**
22	rs467768	28291986	A	0.14	20	3.43E-03	0.64	0.15	55	9.29E-03	0.70	- -	*NIPSNAP1**

**corresponding SNP is present in an intron in this gene*

Functional relevance of identified top SNPs associated with CMH in individuals without COPD

The rs3845529 genotypes showed no significant eQTL effect on *USHA2* mRNA expression levels and rs944899 genotypes not on *SOX1* mRNA expression levels in lung tissue ($p \approx 7 \times 10^{-1}$). In contrast, a strong effect of rs4863687 genotypes (CC = 622, TC = 408, TT = 66) on *MAML3* mRNA expression levels was shown; the CMH associated risk allele T was significantly associated with higher expression of *MAML3* ($p = 2.59 \times 10^{-12}$) (Affymetrix ID: 100146901-TGI-at, Ensemble ID: NM-018717) (Figure 3).
Gene expression profiles of genes close to rs1690139 were not present on the Affymetrix array for the eQTL-analyses.

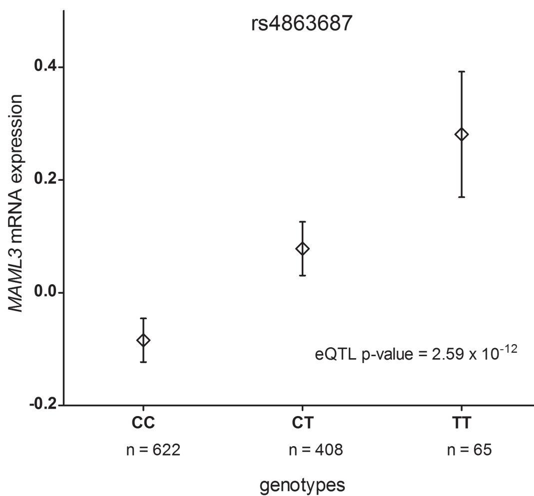


Figure 3. Boxplots of lung gene expression levels for *MAML3* according to genotype groups for SNP rs4868687 in 1,095 individuals.

Overlap of top SNPs associated with CMH in COPD and non-COPD

Comparison of top SNPs in the GWA study in NELSON participants with COPD (5,146 SNPs, $p < 10^{-2}$) and in the GWA study in NELSON participants without COPD (5,186 SNPs, $p < 10^{-2}$) showed 60 overlapping SNPs (Table 5). When only SNPs with a p-value $< 10^{-3}$ were considered, only one overlapping SNP was observed: rs4306981, located close to (64kb) the *progestin and adipoQ receptor family member III gene* (*PAQR3*) on chromosome 4q21.21 ($p = 4.40 \times 10^{-5}$ in individuals with COPD and 5.73×10^{-4} in those without COPD) with effects in the same direction in both analyses (OR = 1.57 and OR = 1.40, respectively). Follow up of this SNP in COPD cohorts did not confirm this association (meta-analysis across NELSON and replication cohorts $p = 4.12 \times 10^{-3}$).

Discussion

In the current study we performed two separate GWA studies on smoking induced CMH, one in individuals with COPD and another in individuals without COPD. We did not find genome wide significance for CMH in either individuals with COPD and without COPD. However, we found suggestive evidence of association of some genes with CMH and differential mRNA expression for some of these genes. Different genes were associated with CMH in smokers with and without COPD. We found one overlapping SNP associated with CMH in NELSON participants with and without COPD with a p-value $< 10^{-3}$, yet this was not replicated in the validation cohorts. Together our data raise the possibility that the pathogenetic development of CMH is differentially regulated in individuals with and without COPD.

In the analysis of CMH performed in individuals with COPD, we found one SNP, rs10461985, in *GDNF-AS1* which has a lower p-value in the replication cohorts compared with the identification analysis (p = 5.43×10^{-5} and p = 1.82×10^{-4} respectively), the SNP showing the same direction of effect in all cohorts except one separately. Antisense RNAs are transcribed to prevent translation of a complementary mRNA by base pairing to it and blocking translation [25]. In this way *GDNF-AS1* prevents expression of *GDNF*. As *GDNF* expression was significantly lower in bronchial biopsies of COPD patients with CMH than without CMH, this is suggestive for the hypothesis that expression of *GDNF-AS1* attenuates CMH. Unfortunately, we were not able to perform a relevant study to assess the expression of *GDNF-AS1* in bronchial biopsies of COPD-patients with and without CMH, since *GDNF-AS1* was not present on the Affymetrix chip used to investigate mRNA expression in COPD patients (GLUCOLD). *GDNF* is a neurotrophic factor that can induce plasticity in sensory neurons innervating the respiratory tract and is involved in lung development [26-28]. These data suggest that GDNF is a biologically plausible candidate gene for both COPD and CMH. However, the gene has not been identified in previous GWA studies of lung function or COPD, making it more likely that it is a gene related to CMH in those who have COPD or a gene that interacts with genes associated with COPD. We did not have sufficient power to further investigate the latter possibility.

The SNP rs4863687 which is located in the *MAML3* gene on chromosome 4, a transcriptional co-activator for Notch signaling, was associated with CMH in individuals without COPD. It has been suggested that *MAM* interacts functionally with different transcription factors, including β-catenin and NF-κB both associated with lung inflammation [29]. We

found a strong effect of rs4863687 genotype on *MAML3* mRNA expression levels; the risk allele T was significantly associated with higher expression of *MAML3*. These data suggest that *MAML3* affects risk for CMH by influencing inflammation. Additionally, it was shown in mice that coordinated cooperation between Wnt signaling and Notch signaling in intestinal epithelium is necessary for the maintenance of proliferative cells and that disruption of the Notch signaling pathway induces goblet cell conversion of crypt proliferative cells [30]. It is conceivable that the role of the Notch signaling pathway is also important in the airway epithelium and that *MAML3* may play a role in goblet cell hyperplasia and consequently CMH.
Rs944899, associated with CMH in individuals without COPD, is located close to the *SOX1* gene that belongs to a family of transcription factors involved in many tissues and developmental processes. SOX proteins have unique functions in different cell types, and different functions within the same cell type. The specificity of these functions is regulated by protein-protein interactions [31]. SOX proteins also regulate the Wnt signaling pathway, required for the specification and differentiation of lung epithelial cells, by interacting with β-catenin [31]. Since *SOX* and *MAML3* are both associated with β-catenin it is conceivable that there is a link between these genes and CMH.

There are limitations to the study. The power of each identification analysis (338 cases and 511 controls in COPD and 342 cases and 1,006 controls in non-COPD) is rather limited, possibly explaining the lack of genome-wide significant findings. Moreover, also some replication cohorts were underpowered and CMH is rather a rough estimate. However, we found suggestive evidence for a genetic contribution to CMH in the full population without stratification for COPD, thus suggesting that power would be more of a problem than the definition of CMH [15]. When we analyzed whether our previously reported gene *SATB1* was associated with CMH in individuals with and without COPD, we also found that the significance was considerably reduced, p-values of rs6577641 being $2.52\ 10^{-2}$ and $5.69\ 10^{-2}$ respectively.

In summary, we found no significant overlap in genes associated with CMH in individuals with COPD and in individuals without COPD. In COPD lower *GDNF* mRNA expression in bronchial biopsies was significantly associated with CMH, possibly by the altered action of *GDNF-AS1*, our top gene. Furthermore, in individuals without COPD, a top SNP in *MAML3* that nominally replicated in the non-COPD cohort was an eQTL in lung tissue. Our results suggest genetic heterogeneity of CMH in individuals with and without COPD and indicate that it is worthwhile to repeat this study in much larger cohorts.

References

1 de Oca MM, Halbert RJ, Lopez MV et al. The chronic bronchitis phenotype in subjects with and without COPD: the PLATINO study. The European respiratory journal: official journal of the European Society for Clinical Respiratory Physiology. 2012;40(1):28-36

2 Kim V and Criner GJ. Chronic bronchitis and chronic obstructive pulmonary disease. American journal of respiratory and critical care medicine. 2013;187(3):228-237

3 Vestbo J. Epidemiological studies in mucus hypersecretion. Novartis Foundation symposium. 2002;248(1528-2511; 1528-2511):3-12

4 Agusti A, Calverley PM, Celli B et al. Characterisation of COPD heterogeneity in the ECLIPSE cohort. Respiratory research. 2010;11:122-9921-11-122

5 Lange P, Parner J, Prescott E and Vestbo J. Chronic bronchitis in an elderly population. Age and Ageing. 2003;32(0002-0729; 6):636-642

6 Patel IS, Seemungal TA, Wilks M, Lloyd-Owen SJ, Donaldson GC and Wedzicha JA. Relationship between bacterial colonisation and the frequency, character, and severity of COPD exacerbations. Thorax. 2002;57(9):759-764

7 Vestbo J. Chronic bronchitis: should it worry us? Chron Respir Dis. 2004;1(1479-9723; 1479-9723; 3):173-176

8 Forey BA, Thornton AJ and Lee PN. Systematic review with meta-analysis of the epidemiological evidence relating smoking to COPD, chronic bronchitis and emphysema. BMC pulmonary medicine. 2011;11:36-2466-11-36

9 Dijkstra AE, de Jong K, Boezen HM et al. Risk factors for chronic mucus hypersecretion in individuals with and without COPD: influence of smoking and job exposure on CMH. Occupational and environmental medicine. 2014

10 Viegi G, Carrozzi L, Di Pede F et al. Risk factors for chronic obstructive pulmonary disease in a north Italian rural area. European journal of epidemiology. 1994;10(6):725-731

11 Silverman EK, Chapman HA, Drazen JM et al. Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. Risk to relatives for airflow obstruction and chronic bronchitis. American journal of respiratory and critical care medicine. 1998;157(6 Pt 1):1770-1778

12 Hallberg J, Dominicus A, Eriksson UK et al. Interaction between smoking and genetic factors in the development of chronic bronchitis. American journal of respiratory and critical care medicine. 2008;177(5):486-490

13 Zhu G, Agusti A, Gulsvik A et al. CTLA4 gene polymorphisms are associated with chronic bronchitis. Eur. Respir.J. 2009;(1399-3003)

14 van Klaveren RJ, Oudkerk M, Prokop M et al. Management of lung nodules detected by volume CT scanning. *The New England journal of medicine.* 2009;361(23):2221-2229

15 Dijkstra AE, Smolonska J, van den Berge M et al. Susceptibility to chronic mucus hypersecretion, a genome wide association study. *PloS one.* 2014;9(4):e91621

16 Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R and Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *The European respiratory journal.Supplement.* 1993;16:5-40

17 Cho MH, Castaldi PJ, Wan ES et al. A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19q13. *Human molecular genetics.* 2011

18 Regan EA, Hokanson JE, Murphy JR et al. Genetic epidemiology of COPD (COPDGene) study design. *Copd.* 2010;7(1):32-43

19 Vestbo J, Anderson W, Coxson HO et al. Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). *The European respiratory journal: official journal of the European Society for Clinical Respiratory Physiology.* 2008;31(4):869-873

20 Bild DE, Bluemke DA, Burke GL et al. Multi-ethnic study of atherosclerosis: objectives and design. *American Journal of Epidemiology.* 2002;156(9):871-881

21 Hao K, Bosse Y, Nickle DC et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS genetics.* 2012;8(11):e1003029

22 Lapperre TS, Snoeck-Stroband JB, Gosman MM et al. Effect of fluticasone with and without salmeterol on pulmonary outcomes in chronic obstructive pulmonary disease: a randomized trial. *Annals of Internal Medicine.* 2009;151(8):517-527

23 van den Berge M, Steiling K, Timens W et al. Airway gene expression in COPD is dynamic with inhaled corticosteroid treatment and reflects biological pathways associated with disease activity. *Thorax.* 2014;69(1):14-23

24 Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics.* 2007;81(0002-9297; 0002-9297; 3):559-575

25 Weiss B, Davidkova G and Zhou LW. Antisense RNA gene therapy for studying and modulating biological processes. C*ellular and molecular life sciences :* CMLS. 1999;55(3):334-358

26 Tollet J, Everett AW and Sparrow MP. Development of neural tissue and airway smooth muscle in fetal mouse lung explants: a role for glial-derived neurotrophic factor in lung innervation. *American journal of respiratory cell and molecular biology.* 2002;26(4):420-429

27 Karihaloo A, Nickel C and Cantley LG. Signals which build a tubule. *Nephron.Experimental nephrology.* 2005;100(1):e40-5

28 Lieu TM, Myers AC, Meeker S and Undem BJ. TRPV1 induction in airway vagal low-threshold mechanosensory neurons by allergen challenge and neurotrophic factors. *American journal of physiology.Lung cellular and molecular physiology.* 2012;302(9):L941-8 /

29 Oyama T, Harigaya K, Sasaki N et al. Mastermind-like 1 (MamL1) and mastermind-like 3 (MamL3) are essential for Notch signaling in vivo. *Development (Cambridge, England).* 2011;138(23):5235-5246

30 van Es JH, van Gijn ME, Riccio O et al. Notch/gamma-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature.* 2005;435(7044):959-963

31 Bernard P and Harley VR. Acquisition of SOX transcription factor specificity through protein-protein interaction, modulation of Wnt signalling and post-translational modification. *The international journal of biochemistry & cell biology.* 2010;42(3):400-410

Chapter 5

Low-dose CT measurements of airway dimensions and emphysema associated with airflow limitation in heavy smokers: a cross sectional study

AE Dijkstra, DS Postma, N ten Hacken, JM Vonk, M Oudkerk, PMA van Ooijen, P Zanen, FA Mohamed Hoesein, B van Ginneken, M Schmidt, HJM Groen

Respiratory Research. 2013 January

Abstract

Background

Increased airway wall thickness (AWT) and parenchymal lung destruction both contribute to airflow limitation. Advances in computed tomography (CT) post-processing imaging allow to quantify these features. The aim of this Dutch population study is to assess the relationships between AWT, lung function, emphysema and respiratory symptoms.

Methods

AWT and emphysema were assessed by low-dose CT in 500 male heavy smokers, randomly selected from a lung cancer screening population. AWT was measured in each lung lobe in cross-sectionally reformatted images with an automated imaging program at locations with an internal diameter of 3.5mm, and validated in smaller cohorts of patients. The 15th percentile method (Perc15) was used to assess the severity of emphysema. Information about respiratory symptoms and smoking behavior was collected by questionnaires and lung function by spirometry.

Results

Median AWT in airways with an internal diameter of 3.5mm ($AWT_{3.5}$) was 0.57 (0.44-0.74)mm. Median AWT in subjects without symptoms was 0.52 (0.41-0.66) and in those with dyspnea and/or wheezing 0.65 (0.52-0.81)mm ($p < 0.001$). In the multivariate analysis only $AWT_{3.5}$ and emphysema independently explained 31.1% and 9.5% of the variance in FEV_1%predicted, respectively, after adjustment for smoking behavior.

Conclusions

Post processing standardization of airway wall measurements provides a reliable and useful method to assess airway wall thickness. Increased airway wall thickness contributes more to airflow limitation than emphysema in a smoking male population even after adjustment for smoking behavior.

Introduction

The quantification of airway dimensions by CT has become feasible with the development of multi detector computed tomography (CT) and new software tools for image analysis [1, 2]. Assessment of airway dimensions by CT has been studied particularly in relation to asthma, smoking and chronic obstructive pulmonary disease (COPD) [3-8], diseases generally associated with chronic or intermittent airflow limitation. So far, airway wall thickness (AWT) measurements have been performed by selecting well quantifiable airways [9-12] or by standardizing dimensions to airways with a 10mm internal lumen perimeter (pi10, equivalent to about 3.2mm internal airway lumen diameter) derived from a small number of airways [5, 6, 13-15].

More recently, low-dose multi slice CT has become available, a technique with a good quantifying performance that is preferred for monitoring of pulmonary and airway abnormalities as compared to the high radiation exposures with high-resolution computed tomography (high-dose CT). The cumulative radiation dose exposure with low-dose CT remains very low, even when individuals are exposed multiple times [16]. Airway dimensions measured with low-dose CT have only been reported in few studies using the same diversity in analytic approaches as applied with high-dose CT measurements [7, 17-19].
Airway dimensions and extent of emphysema are known to be associated with airflow limitation [6, 9-14, 19, 20], although the influence of smoking behavior or signs of airway disease such as cough, dyspnea, wheezing or mucus overproduction on airflow limitation is not clear [5, 7, 21, 22].

The aims of this study are to quantify airway dimensions of the lung in multiple airway sections of each lobe in a novel manner and the extent of emphysema by using low-dose CT. These measurements were associated with the influence of airflow limitation, respiratory symptoms and corrected for smoking behavior.

Methods

Population

We randomly selected 500 current and former smokers participating in the Groningen cohort of a male population-based multi-centre lung cancer screening study (NELSON). The Dutch ministry of health and the Medical Ethics Committee of the hospital approved the study protocol. Informed consent was obtained from all participants. Detailed inclusion criteria and characteristics have been described elsewhere 23. In short, subjects with a smoking history of at least 20 packyears were included. Information on the presence or absence of respiratory symptoms and smoking (packyears and former or current smoking) was obtained by questionnaires. The question used to record respiratory symptoms was "do you have experienced the following symptoms cough, sputum expectoration, wheezing or dyspnea for at least 3 months during the past year, even when you did not have a cold?"

Lung function

All participants performed standardized spirometry according to the European Respiratory Society guidelines [24], including forced expiratory volume in 1 sec (FEV_1) and forced vital capacity (FVC) at the start of the study. In this population- based study we did not administer a bronchodilator.

CT scanning

Low-dose CT images of the chest were acquired at full inspiration after appropriate instruction on one CT scanner (Sensation 16, Siemens Medical Solutions, Forchheim, Germany) [23] according to the following protocol: spiral acquisition at 120kV, 20mAs, rotation time 0.5s, pitch 1.5 and collimation 16 × 0.6mm, field of view 300 to 350mm, slice thickness 1mm and slice increment 0.7mm. The effective radiation dose was less than 0.8mSv. Contrast medium was not used. The images were reconstructed to a pixel matrix of 512 × 512 using B30f kernel. Thus, the spatial resolution was 0.59 to 0.68mm in x-y plane, and 0.7mm in z plane. The CT system was calibrated routinely.

Quantification of AWT

AWT was measured in cross-sectionally reformatted images with an automated research software prototype MEVIS Airway Examiner v1.0 (release 2009, Fraunhofer MEVIS, Bremen, Germany) based on an algorithm by Weinheimer at locations with a fixed internal diameter of 3.5mm in each lung lobe [25]. This software automatically extracts

airway centerlines, re-samples images perpendicular to the airway direction at equally spaced positions along the centerline and detects inner and outer airway wall borders in these images. The outer wall border is detectable when no adjacent tissue with similar CT density is present and is taken into account when the wall is detected in at least 25% of the perimeter at a location. AWT and the fraction of perimeter where the outer wall border was identified (Assessed Perimeter Fraction, APF) are calculated for each location. Wall thickness quantification accounts for partial volume effects by integrating Hounsfield units across the wall. Accuracy and reproducibility of this algorithm was tested previously under clinical conditions using a similar protocol as used in our study [2]. Average wall thickness and cumulative APF of all detectable airway locations with a fixed lumen diameter is reported per lobe and for the whole lung. The borders of the lung lobes were automatically calculated by the software in a standard way. All low-dose CT scans were visually evaluated for appropriate segmentation.

Quantification of emphysema and lung volume

Quantification of emphysema was based on density differences and measured with a software tool called ImageXplorer (Image Sciences Institute, Utrecht, the Netherlands) [16, 26]. This software produces automatically the lung volume. The extent of emphysema was automatically performed at the 15th percentile (Perc15) of the Hounsfield density distribution. Perc15 is the threshold density value where 15% of all voxels has a lower density [27]. A lower Perc15, i.e. closer to -1000 HU, means that more emphysema is present. All scans were reconstructed with a soft reconstruction filter (Philips B, Siemens B30f). Airways were automatically excluded to assess density of lung parenchyma exclusively and HU densities of the entire scan were recalibrated using automatically measured average densities in the trachea and shifting the HU values of the entire scan so that air density in the trachea became -1000 HU. Additionally, the percentage of low attenuation area, defined as the proportion of low-density voxels below -950 HU (%LAA-950HU) was used. %LAA-950HU was log-transformed because of skewed distribution.

Explorative studies

Prior to the research described above we have

1) established the optimal internal airway diameter, i.e. the internal airway diameter that allows the highest number of cumulatively Assessed Perimeter Fractions (APF) for the whole lung. Therefore we measured APF on 20 selected NELSON CT's in airways with a lumen diameter of 2.5, 3.0, 3.5, 4.0, 4.5 and 5.0mm (± 0.25mm) divided into 3 groups: no emphysema and normal lung function (n = 8, p^{15} > -920 and FEV_1/FVC > 85%),

moderate emphysema and normal lung function (n = 4, -940 < p^{15} < -960 and FEV_1/FVC > 70%) and no emphysema and severe airflow limitation (n = 8, p^{15} > -920 and FEV_1/FVC < 50%).

2) compared the mean $AWT_{3.5}$, using the same method as described above, at the established internal lumen size with high- and low-dose CT in 8 NELSON subjects from whom high- and low-dose CT were available. These CT data were obtained in spiral mode with 16 × 0.75mm collimation and in full inspiration with the same scanner (Sensation-16 Siemens Medical Solutions, Forchheim, Germany). Axial images were reconstructed with 1.0mm thickness at 0.7mm increments. All scans were reconstructed with a soft reconstruction filter (Siemens B30f) at a 512 × 512 matrix.

3) determined the generation where airways with the established optimal internal lumen size are present. AWT measurements at 3.5mm internal lumen size were performed in 57 randomly selected low-dose CTs of NELSON subjects. A multi-planar reconstruction (MPR) was made in each case of the apical upper lobe bronchus (B1) and the posterior lower lobe bronchus (B10). Subsequently the location was projected on the segmentation image. Three-dimensional image moving created the opportunity to observe airways from various directions and to check bifurcations and count airway generations according to the method of Boyden [28].

Statistical analysis

Data are reported as mean ± standard deviation (SD) or median (25th - 75th percentile) values as appropriate. The mean AWT at 3.5mm internal lumen size ($AWT_{3.5}$) of all five lobes per case was calculated taking into account the APF per lobe by the following formulae: ((AWT left upper lobe x APF left upper lobe) + (AWT left lower lobe x APF left lower lobe) + (AWT right upper lobe x APF right upper lobe) + (AWT right middle lobe x APF right middle lobe) + (AWT left upper lobe x APF left upper lobe)) / sum of APF of all lung lobes. $AWT_{3.5}$ for the whole population was skewed distributed, therefore we report median AWT and range, and log-transformed AWT was used in the analyses.

Differences between various categories were explored using chi-square tests (dichotomous data), 2-tailed unpaired Student's t-tests for normally distributed continuous data and Mann-Whitney U-tests for not normally distributed continuous data. The difference in airway wall thickness between lung lobes was assessed with a Wilcoxon signed rank test. Univariate linear regression analyses was used to study the relationships between clinical variables and AWT, and those variables with FEV_1%predicted. Next, multivariate linear

regression analyses were performed on those clinical variables showing significance in the univariate regression analyses. Outcomes of these analyses have been described with beta's and p-values. Bland-Altman plot was used to analyze the agreement between AWT by high- and low-dose CT [29]. All statistical analyses were performed using SPSS 18.0 for Windows; P-values below 0.05 were considered statistically significant.

Results

Population characteristics

After visual evaluation 8 out of the 500 randomly selected subjects were excluded because of (partial) missing of airway segmentation on CT. The mean age of the cohort was 59.4 (± 5.2) years, approximately 59% were current smokers and median packyears was 34.0 (28.0 - 45.6). More than 51 % of the population reported at least one respiratory symptom (Table 1).

Table 1. Clinical and demographic characteristics of heavy smokers from the general population cohort

N	**492**
Age, years	59.4 ± 5.2
Packyears smoking	34.0 (28.0 - 45.6)
Current smoking, %	59.1
FEV_1, liter	3.45 ± 0.76
FEV_1, %predicted	98.2 ± 19.7
FEV_1/FVC, %	70.0 ± 10.7
Emphysema; Perc15, HU	-920 (-930 to -907)
Emphysema; %LAA -950 HU	2.5 (1.3 - 4.3)
Emphysema; >5 %LAA -950 HU, %	19.7
Lung volume on CT, liter	6.5 ± 1.4
Chronic Mucus Hypersecretion, %	29.7
Cough, %	32.7
Dyspnea, %	28.4
Wheezing, %	25.3
No respiratory symptoms, %	48.6

Mean ± standard deviation shown for continuous data and median (interquartile range) for non-parametric distribution. Definition of abbreviations: HU = Hounsfield Units; Perc15 = the threshold density value where 15% of all voxels has a lower density; %LAA -950 = percentage low attenuation areas < -950 HU; >5 %LAA -950 HU, % = percentage of the population having >5% low attenuation areas < -950HU.

Airway wall thickness, Exploratory analyses

The highest numbers of cumulatively assessed perimeter fractions (APF) of airways were reached at an internal lumen perimeter of 3.5 ± 0.25mm (Figure 1); therefore this diameter was selected for further analysis.

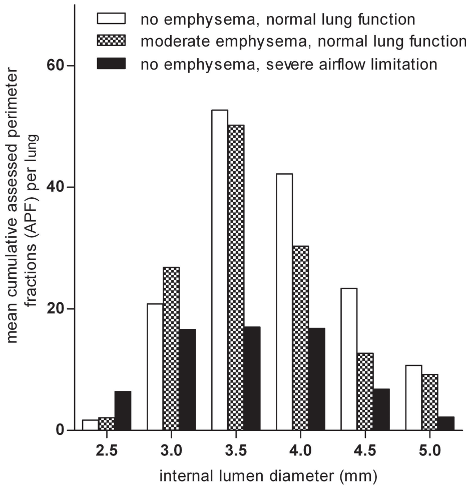


Figure 1. Determination of the optimal airway size.

Mean cumulative assessed perimeter fractions in the total lung, for different groups of patients at different internal airway lumen diameters. APF was measured on low-dose CT of 20 selected NELSON subjects divided into 3 groups: subjects without emphysema and with normal lung function (n = 8, perc15 > -910 and FEV_1/FVC > 85%), with moderate emphysema and normal lung function (n = 4, -940 < perc15 < -960 and FEV_1/FVC > 70%) and without emphysema and having severe airflow limitation (n = 8, perc15 > -920 and FEV_1/FVC < 50%), in airways with a lumen diameter of 2.5, 3.0, 3.5, 4.0, 4.5 and 5.0mm (± 0.25mm).

Median $AWT_{3.5}$ on low-dose CT was comparable with median $AWT_{3.5}$ on high-dose CT, respectively 0.57 (0.48-0.74)mm and 0.55 (0.47-0.73)mm (p = 0.89, n = 8). This demonstrates that MEVIS software analysis of data from low-dose CT gives similar results as from high-dose CT (Figure 2).

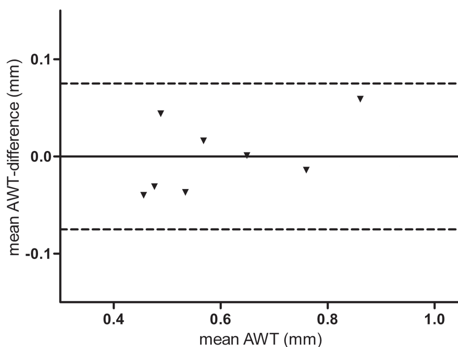


Figure 2. Comparison of airway wall thickness on high- and low-dose CT.

$AWT_{3.5}$ was measured on high and low-dose CT of 8 NELSON subjects. Bland & Altman plot shows agreement between mean $AWT_{3.5}$ determined by high- and low-dose CT. Dashed lines depict the 95% confidence interval.

Airways with an internal diameter of 3.5mm appeared mainly in the 5th generation (range, 3rd-7th generation) in the upper lobe bronchus, and mainly in the 8th-10th generation (range, 6th-12th generation) in the lower lobe bronchus. This distribution was observed

irrespective of smoking, presence of airflow limitation (defined as FEV_1/FVC <0.70) or chronic mucus hypersecretion (CMH) (Figure 3).

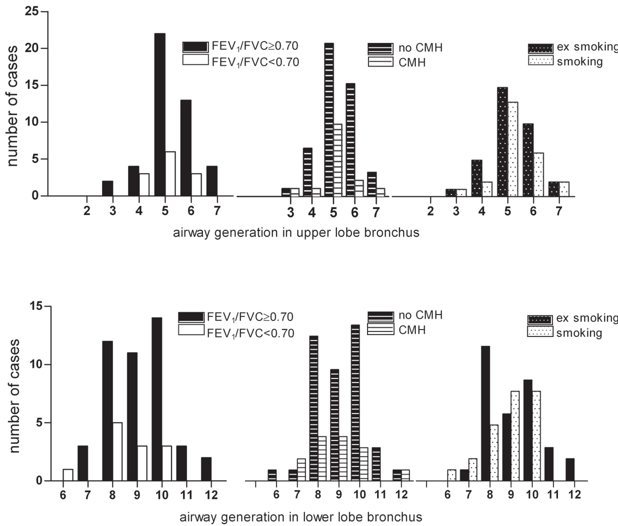


Figure 3. Distribution of 3.5mm sized airways over airway generations.

The distribution of 3.5mm internal lumen sized airways over the 2nd till 7th airway generations in the apical upper lobe bronchus (B1) and over the 6th till 12th generation in the lower lobe bronchus (B10) in the right lung, for subjects with and without airway obstruction, for subjects with and without Chronic Mucus Hypersecretion (CMH) and for current and former smokers. The number of cases (e.g. no CMH and CMH) in each group was limited but the distribution was similar.

Airway wall thickness in the population

In the whole population the median (25th-75th percentile) $AWT_{3.5}$ was 0.57 (0.44-0.74) mm. The APF in the whole lung varied from 142 to 295 (median 215). The results per lung lobe are presented in the supplement, Table 1.

Airway wall thickness and clinical characteristics

Significantly higher $AWT_{3.5}$ values were observed in subjects with dyspnea and/or wheezing (n = 181, median $AWT_{3.5}$ 0.66mm) or with cough and/or CMH (n = 201, median $AWT_{3.5}$ 0.63mm) compared to subjects without dyspnea and/or wheezing (n = 309, median $AWT_{3.5}$ 0.53mm, $p < 0.001$) or without cough and/or CMH (n = 291, median AWT 0.53mm, $p < 0.001$). Current smokers and former smokers had comparable median $AWT_{3.5}$ values, i.e. 0.58mm and 0.56mm (p = 0.113) respectively (Table 2).

Table 2: Airway wall thickness in subjects with and without respiratory symptoms and in current and former smokers.

	n	$AWT_{3.5}$ (mm)		n	$AWT_{3.5}$ (mm)	p-value
CMH	146	0.62 (0.49-0.80)	no CMH	346	0.56 (0.44-0.71)	0.002*
Cough	161	0.64 (0.50-0.84)	no cough	331	0.54 (0.43-0.70)	<0.001*
Dyspnea	140	0.65 (0.51-0.80)	no dyspnea	352	0.54 (0.43-0.69)	<0.001*
Wheezing	125	0.66 (0.52-0.85)	no wheezing	367	0.53 (0.43-0.70)	<0.001*
CMH and/or cough	201	0.63 (0.49-0.79)	no CMH or cough	291	0.53 (0.43-0.70)	<0.001*
Dyspnea and/or wheezing	181	0.66 (0.52-0.83)	no dyspnea or wheezing	311	0.53 (0.42-0.67)	<0.001*
Cough, CMH, dyspnea and wheezing	49	0.69 (0.51-0.89)	no cough, CMH, dyspnea or wheezing	239	0.52 (0.41-0.66)	<0.001*
Current smoking	291	0.58 (0.45-0.75)	Former smoking	201	0.56 (0.43-0.71)	0.113

*Data presented as median (interquartile range) values. * is p-value < 0.05.*
Definition of abbreviations: $AWT_{3.5}$ = airway wall thickness at 3.5mm internal lumen size; CMH = chronic mucus hypersecretion.

Univariate linear regression analysis showed inverse associations between log-$AWT_{3.5}$ and FEV_1 ($b = -0.233$, $p < 0.001$), FEV_1/FVC ($b = -0.015$, $p < 0.001$), FEV_1%predicted ($b = -0.010$, $p < 0.001$) and lung volume ($b = -0.055$, $p < 0.001$) and positive associations between log-$AWT_{3.5}$, Perc15 and number of packyears smoking, respectively ($b = 0.003$, $p < 0.001$) and ($b = 0.003$, $p = 0.003$) (Table 3).

Table 3. Univariate associations between (A) log-transformed $AWT_{3.5}$ and (B) FEV_1% predicted, and clinical characteristics

Dependent variable	A. Log-$AWT_{3.5}$		B. FEV_1, % predicted	
	Beta	p-value	Beta	p-value
Log-$AWT_{3.5}$			-31.21	**<0.001**
FEV_1, % predicted	-0.010	**<0.001**		
FEV_1, liter	-0.233	**<0.001**	23.56	**<0.001**
FEV_1/FVC, %	-0.015	**<0.001**	1.441	**<0.001**
FVC, % predicted	-0.009	**<0.001**	0.971	**<0.001**
Emphysema; Perc15	0.003	**<0.001**	0.204	**<0.001**
Emphysema; Log-%LAA -950 HU	-0.041	**0.023**	-5.872	**<0.001**
Lung volume on CT, Liter	-0.055	**<0.001**	-0.257	0.691
Packyears	0.003	**0.003**	-0.153	**0.004**
Smoking (former/current)	0.059	0.065	-3.257	0.071
Age, years	-0.002	0.458	-0.289	0.092
Height, cm	-0.002	0.444	-0.121	0.425
Cough	0.156	**<0.001**	-9.234	**<0.001**
CMH	0.114	**0.001**	-8.171	**<0.001**
Dyspnea	0.158	**<0.001**	-11.99	**<0.001**
Wheezing	0.202	**<0.001**	-13.26	**<0.001**

Definition of abbreviations: Log-$AWT_{3.5}$ = log transformed airway wall thickness at 3.5mm diameter; CMH = chronic mucus hypersecretion; Log-%LAA -950 HU = log transformed percentage of low attenuation areas < -950 HU; bold is p-value < 0.05.

Multivariate analysis on all clinical variables significantly associated with AWT in univariate analyses revealed that log-$AWT_{3.5}$ was independently associated with lower FEV_1%predicted ($b = -0.010$, $p < 0.001$), higher Perc15 ($b = 0.005$, $p < 0.001$), and lung volume ($b = -0.037$, $p = 0.005$) respectively (Table 4).

Contribution of airway wall thickness and emphysema to airflow limitation

To study the contribution of $AWT_{3.5}$, emphysema (Perc15), and clinical variables to airflow limitation, FEV_1%predicted was taken as dependent variable. A significant negative association was found between FEV_1%predicted and log-$AWT_{3.5}$ ($b = -31.21$, $p < 0.001$), packyears ($b = -0.153$, $p = 0.004$), cough, CMH, dyspnea and wheezing (b = -9.234, -8.17, -11.99, -13.26 respectively, all p-values < 0.001), and a positive association between FEV%predicted and Perc15 ($b = 0.204$, $p < 0.001$) (Table 3).

Table 4. Multivariate linear regression: dependent variable is (A) log transformed $AWT_{3.5}$ and (B) FEV_1% predicted

Dependent variable	**A. Log-$AWT_{3.5}$**		**B. FEV_1, % predicted**	
	Beta	**p-value**	**Beta**	**p-value**
FEV_1,% predicted	-0.010	<0.001		
Log-$AWT_{3.5}$			-31.28	<0.001
Emphysema; Perc15	0.005	<0.001	0.342	<0.001
Lung volume	-0.037	0.005	0.209	0.773
Packyears	0.002	0.031	-0.029	0.499
Smoking (former/current)	-0.011	0.700	-2.756	0.086
Cough	0.030	0.404	-0.408	0.834
CMH	-0.026	0.447	-2.098	0.262
Dyspnea	-0.007	0.828	-2.945	0.116
Wheezing	0.058	0.124	-4.004	0.051

Definition of abbreviations: Log-$AWT_{3.5}$ = log transformed airway wall thickness at 3.5mm diameter; CMH = Chronic Mucus Hypersecretion; bold is p-value < 0.05.

Multivariate analysis on all clinical variables that were associated with FEV_1 %predicted in univariate analyses showed that higher FEV_1 %predicted was significantly associated with lower $AWT_{3.5}$ values ($b = -31.3$, $p < 0.001$) and with higher Perc15 ($b = 0.342$, $p < 0.001$) (Table 4). $AWT_{3.5}$ and Perc15 explained 31.1% and 9.5% of the variance of FEV_1 %predicted, respectively. The results of the multivariate regression analysis with emphysema expressed as %LAA-950HU as independent variable, are presented in the supplement, Table 2.

Discussion

Low-dose CT is an appealing approach to quantify simultaneously pulmonary and airway abnormalities. Our study shows that the use of low-dose CT combined with modern post processing software tools provides reliable information on airway wall thickness and the extent of emphysema in a heavy smoking male population. Although CT does not provide dynamic measurements, airway wall thickening and emphysema explained respectively 31.1% and 9.5% of the variance in FEV_1%predicted, the most commonly used variable of airflow limitation. Changes in AWT of more than 0.1mm reflecting lumen surface changes over 8% measured at one air lumen level were observed between cases with and without respiratory symptoms.

Our study confirmed that increased AWT is associated with lower FEV_1%predicted. This lower FEV_1%predicted depends on the airway size in which the measurements of AWT are being performed [12, 30] and on the characteristics of the study population [7, 20]. Our population consisted of rather healthy elderly males from a randomly recruited Dutch heavy smoking population and still we were able to find significant associations between thicker airway walls and more severe airflow limitation. In contrast with the study of Nakano we found a significant negative association between AWT and FEV_1/FVC illustrating the sensitivity of our method [9].

The significant but weak negative association between airway wall thickness and emphysema has also been reported in other studies 7, 13, 31 but was not found in the study by Nakano 9. Loss of elastic recoil may contribute to collapse of the airways resulting in a more proximal localization of airways with 3.5mm internal lumen diameter. As these more proximal airways have thicker airway walls this phenomenon contributes to the weak negative association. Another possible explanation for this negative association may be that there are subjects with predominantly airway wall thickening and others with predominantly emphysematous changes. Particularly subjects with relatively more airway wall thickening are responsible for the negative association and subjects with predominantly emphysematous changes do hardly contribute. Apparently, in our population of subjects with normal lung function and with mild airflow limitation, the bronchitic phenotype is already present in the very early stages of smoking-induced lung disease. Discrepancies between the study of Nakano [9] and our study may be due to the composition and size of the study populations, respectively predominantly emphysema versus predominantly healthy smokers with respiratory symptoms.

Importantly, we observed that the contribution of $AWT_{3.5}$ to airflow limitation was larger than the emphysema component. Moreover, $AWT_{3.5}$ and emphysema together only explained about 40% of the variance in FEV_1%predicted in this smoking male population. This unexpected low contribution of $AWT_{3.5}$ and emphysema to FEV_1%predicted may be that the CT images were obtained at full inspiration, while FEV_1 reflects expiratory airflow limitation. One explanation for this observation is that airflow limitation is not only due to reduced airway diameter at one level but should be evaluated as an integral of all airways at all lumen diameters. This is difficult to achieve and therefore we took the smallest measurable lumen diameter that provides the largest contribution to airflow limitation. A more obvious physiological explanation may be the presence of the heterogeneity in airway ventilation interrupting the symmetry in parallel airways leading to large clusters of poorly ventilated lung units [32].

In the univariate analysis, increased $AWT_{3.5}$ was associated with respiratory symptoms. However, $AWT_{3.5}$ was not associated with the presence of any respiratory symptom in the multivariate analysis after adjustment for FEV_1%predicted, emphysema and smoking behavior. This finding corresponds with other studies [7] and is consistent with the idea that inflammation and airway remodeling, associated with chronic bronchitis, is located in the more central airways [33]. The study of Martinez et al. showed a positive association between airway dimensions and questionnaires, the BODE index [34] and the St. George's respiratory questionnaire [35] including questions about BMI, respiratory symptoms, exercise capacity and lung function. Also Camiciottoli et al. found a positive association between BODE and airway wall thickness [36]. Our study also showed that including respiratory symptoms in the multivariate model with $AWT_{3.5}$ and emphysema has no impact on airflow limitation.
Lung volume depends on height, weight and sex and as a consequence each person has different airway dimensions. Therefore, airway dimensions should be corrected for lung volume. Actually, volume-corrected AWT is the best parameter to use. In this study lung volume does not change the multivariate model because FEV1%predicted is already corrected for lung volume by correcting for patient height.

It has been shown that the automated imaging program (MEVIS Airway Examiner) based on a method by Weinheimer et al. performed much better than the often used "full-width-at-half-maximum"method in a silicon tube phantom, regarding the blurring effect of CT [25, 37]. Usually it is better to use sharper kernels for airway quantification. However, it was shown in the study by Schmidt et al. that the MEVIS airway examiner provides reproducible quantitative results across different reconstruction kernels (B30f

and B50f) and repeated acquisitions [2]. Moreover, the "full-width-at-half-maximum" technique systematically overestimates AWT, particularly in small airways [37]. Because low-dose CT and the automated imaging program (MEVIS Airway Examiner) had not been used previously in clinical studies, we firstly optimized our post processing measurements in smaller cohorts of patients before applying it in the population study. We demonstrated that the highest number of AWT measurements could be performed on airways with an internal diameter of 3.5mm, irrespective of the presence of airflow limitation or emphysema. In addition we demonstrated that differences in $AWT_{3.5}$ are not explained by differences in airway generations. Finally, we demonstrated that low-dose CT imaging provided similar AWT results as high-dose CT imaging.

In a non-biased way we were able to evaluate 230 cumulatively assessed perimeter fractions (APF) per CT, ranging from 27 - 641 APF. In contrast to the commonly used pi10 method, in which a secondary derived variable from few, mostly 6 selected airways is used to estimate the airway wall thickness [5, 6, 13-15], we obtained many direct airway wall measurements. To our opinion direct measurements assessed over all lobes provide a better overall reflection of AWT than a limited number of secondary AWT measurements.

A limitation of this study is, inherent to general population-based studies, that only male smokers with mild emphysema and/or airflow limitation were included. Strengths of our study is the non-biased way of analyzing a high number of airway sections in all lobes that makes our approach more suitable for combined airway wall thickness and emphysema measurements on one low-dose CT scan. Such approach allows monitoring of intervention effects on both parameters. This is important when new treatment modalities will become available for clinical testing.

In the future further developments may involve measurements of thickness of airway walls along the full length of the bronchial tree at in- and expiration scans. Possibly more airflow variability will be explained.

In conclusion, post processing standardization of large numbers of airway wall measurements in all lung lobes is feasible, reliable and an useful method to assess airway wall thickness. We have demonstrated that increased airway wall thickness contributes more to airflow limitation than emphysema in a smoking male population even after adjustment for smoking behavior.

References

1 Takasugi JE and Godwin JD. Radiology of chronic obstructive pulmonary disease. Radiologic clinics of North America. 1998;36(1):29-55

2 Michael Schmidt, Jan-Martin Kuhnigk, Stefan Krass, Michael Owsijewitsch, Bartjan de Hoop and Heinz-Otto Peitgen. Reproducibility of airway wall thickness measurements. Proc SPIE. Mar 09, 2010;762410

3 Kurashima K, Hoshi T, Takayanagi N et al. Airway Dimensions and Pulmonary Function in Chronic Obstructive Pulmonary Disease and Bronchial Asthma. Respirology (Carlton, Vic.). 2011

4 Matsumoto H, Niimi A, Tabuena RP et al. Airway wall thickening in patients with cough variant asthma and nonasthmatic chronic cough. Chest. 2007;131(4):1042-1049

5 Grydeland TB, Dirksen A, Coxson HO et al. Quantitative computed tomography: emphysema and airway wall thickness by sex, age and smoking. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology. 2009;34(1399-3003; 4):858-865

6 Kim WJ, Silverman EK, Hoffman E et al. CT metrics of airway disease and emphysema in severe COPD. Chest. 2009;136(2):396-404

7 Mair G, Maclay J, Miller JJ et al. Airway dimensions in COPD: Relationships with clinical variables. Respiratory medicine. 2010;(1532-3064; 0954-6111)

8 Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS and GOLD Scientific Committee. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. American journal of respiratory and critical care medicine. 2001;163(5):1256-1276

9 Nakano Y, Muro S, Sakai H et al. Computed tomographic measurements of airway dimensions and emphysema in smokers. Correlation with lung function. American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association. 2000;162(1073-449; 3):1102-1108

10 Matsuoka S, Kurihara Y, Yagihashi K, Hoshino M and Nakajima Y. Airway dimensions at inspiratory and expiratory multisection CT in chronic obstructive pulmonary disease: correlation with airflow limitation. Radiology. 2008;248(1527-1315; 3):1042-1049

11 Nishimura M. Application of three-dimensional airway algorithms in a clinical study. Proc Am Thorac Soc. 2008;5(1546-3222; 9):910-914

12 Hasegawa M, Nasuhara Y, Onodera Y et al. Airflow limitation and airway dimensions in chronic obstructive pulmonary disease. American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association. 2006;173(1073-449; 12):1309-1315

13 Patel BD, Coxson HO, Pillai SG et al. Airway wall thickening and emphysema show independent familial aggregation in chronic obstructive pulmonary disease. American Journal of Respiratory & Critical Care Medicine. 2008;178(5):500-505

14 Lee YK, Oh YM, Lee JH et al. Quantitative assessment of emphysema, air trapping, and airway thickening on computed tomography. Lung. 2008;186(0341-2040; 0341-2040; 3):157-165

15 Martinez CH, Chen YH, Westgate PM et al. Relationship between quantitative CT metrics and health status and BODE in chronic obstructive pulmonary disease. Thorax. 2012;67(5):399-406

16 Mets OM, de Jong PA, van Ginneken B, Gietema HA and Lammers JW. Quantitative computed tomography in COPD: possibilities and limitations. Lung. 2012;190(2):133-145

17 Yamashiro T, Matsuoka S, Estepar RS et al. Quantitative assessment of bronchial wall attenuation with thin-section CT: An indicator of airflow limitation in chronic obstructive pulmonary disease. AJR.American journal of roentgenology. 2010;195(2):363-369

18 Washko GR, Dransfield MT, Estepar RS et al. Airway wall attenuation: a biomarker of airway disease in subjects with COPD. Journal of applied physiology (Bethesda, Md.: 1985). 2009;107(1):185-191

19 Bon JM, Leader JK, Weissfeld JL et al. The influence of radiographic phenotype and smoking status on peripheral blood biomarker patterns in chronic obstructive pulmonary disease. PloS one. 2009;4(8):e6865

20 Achenbach T, Weinheimer O, Biedermann A et al. MDCT assessment of airway wall thickness in COPD patients using a new method: correlations with pulmonary function tests. European radiology. 2008;18(1432-1084; 0938-7994; 12):2731-2738

21 Orlandi I, Moroni C, Camiciottoli G et al. Chronic obstructive pulmonary disease: thin-section CT measurement of airway wall thickness and lung attenuation. Radiology. 2005;234(0033-8419; 2):604-610

22 Grydeland TB, Dirksen A, Coxson HO et al. Quantitative CT Measures of Emphysema and Airway Wall Thickness Are Related to Respiratory Symptoms. American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association. 2009;(1535-4970; 1535-4970)

23 van Klaveren RJ, Oudkerk M, Prokop M et al. Management of lung nodules detected by volume CT scanning. The New England journal of medicine. 2009;361(23):2221-2229

24 Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R and Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. The European respiratory journal. Supplement. 1993;16:5-40

25 Weinheimer O, Achenbach T, Bletz C, Duber C, Kauczor HU and Heussel CP. About objective 3-d analysis of airway geometry in computerized tomography. IEEE Transactions on Medical Imaging. 2008;27(0278-0062; 0278-0062; 1):64-74

26 Gietema HA, Schilham AM, van Ginneken B, van Klaveren RJ, Lammers JW and Prokop M. Monitoring of smoking-induced emphysema with CT in a lung cancer screening setting: detection of real increase in extent of emphysema. Radiology. 2007;244(0033-8419; 0033-8419; 3):890-897

27 Parr DG, Sevenoaks M, Deng C, Stoel BC and Stockley RA. Detection of emphysema progression in alpha 1-antitrypsin deficiency using CT densitometry; methodological advances. Respiratory research. 2008;9:21

28 EA B. A synthesis of the prevailing patterns of the bronchopulmonary segments in the light of their variations. (0096-0217; 0096-0217)

29 Bland JM and Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307-310

30 Yamashiro T, Matsuoka S, Estepar RS et al. Quantitative assessment of bronchial wall attenuation with thin-section CT: An indicator of airflow limitation in chronic obstructive pulmonary disease. AJR.American journal of roentgenology. 2010;195(2):363-369

31 Kim S, Choi IS, KimYJ et al. Airway responsiveness to inhaled aspirin is influenced by airway hyperresponsiveness in asthmatic patients. The Korean journal of internal medicine. 2010;25(3):309-316

32 Venegas JG, Winkler T, Musch G et al. Self-organized patchiness in asthma as a prelude to catastrophic shifts. Nature. 2005;434(7034):777-782

33 Hogg JC, Chu F, Utokaparch S et al. The nature of small-airway obstruction in chronic obstructive pulmonary disease. The New England journal of medicine. 2004;350(1533-4406; 0028-4793; 26):2645-2653

34 Ong KC, Earnest A and Lu SJ. A multidimensional grading system (BODE index) as predictor of hospitalization for COPD. Chest. 2005;128(6):3810-3816

35 Jones PW, Quirk FH, Baveystock CM and Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. The American Review of Respiratory Disease. 1992;145(6):1321-1327

36 Camiciottoli G, Bigazzi F, Bartolucci M et al. BODE-index, modified BODE-index and ADO-score in chronic obstructive pulmonary disease: relationship with COPD phenotypes and CT lung density changes. Copd. 2012;9(3):297-304

37 Mets OM, Buckens CF, Zanen P et al. Identification of chronic obstructive pulmonary disease in lung cancer screening computed tomographic scans. JAMA : the journal of the American Medical Association. 2011;306(16):1775-1781

Chapter 5

Supplement

Table 1. Median $AWT_{3.5}$ and cumulatively assessed perimeter fractions (APF) per lung lobe

N=492	$AWT_{3.5}$ (mm)	APF (n)
Right upper lobe	0.58 (0.43-0.78)	40.0 (25.6-59.0)
Right middle lobe	0.55 (0.44-0.72)	19.9 (11.5-33.1)
Right lower lobe	0.56 (0.43-0.78)	59.0 (33.6-88.8)
Left upper lobe	0.58 (0.44-0.74)	43.7 (28.8-61.3)
Left lower lobe	0.61 (0.45-0.81)	43.5 (23.2-69.0)
All lobes	0.57 (0.44-0.74)	214.8 (142.4-295.4)

Data presented as median (interquartile range) values.
Definition of abbreviations: $AWT_{3.5}$ *= airway wall thickness at 3.5 mm internal lumen size; APF = cumulatively assessed perimeter fractions of airway walls.*

The thickest airway walls were present in the left lower lung lobe, i.e. median 0.61 (0.45 - 0.81) mm, and were significantly thicker compared to the airway walls in the other lung lobes (all p-values <0.001). The thinnest airway walls were present in the middle lobe, i.e. median 0.55 (0.44 - 0.72) mm. The highest APF at 3.5 mm internal lumen diameter was observed in the right lower lung lobe, i.e. 59 (34 - 89) and the lowest in the right middle lobe, i.e. 20 (12 - 33).

Table 2. Multivariate linear regression: dependent variable is (A) log transformed $AWT_{3.5}$ and (B) FEV_1% predicted. LAA%-950 HU was used to quantify the extent of emphysema in this analysis.

Dependent variable	A. Log-$AWT_{3.5}$		B. FEV_1, % predicted	
	Beta	p-value	Beta	p-value
FEV_1,% predicted	-0.010	<0.001		
Log-$AWT_{3.5}$			-30.241	<0.001
Emphysema; log-%LAA -950 HU	-0.078	<0.001	-7.429	<0.001
Lung volume	-0.048	<0.001	0.006	0.993
Packyears	0.002	0.023	-0.020	0.636
Smoking (former/current)	0.001	0.986	-2.650	0.096
Cough	0.041	0.253	0.532	0.784
CMH	-0.026	0.446	-2.055	0.271
Dyspnea	-0.007	0.839	-2.538	0.175
Wheezing	0.061	0.107	-4.145	0.993

Definition of abbreviations: Log-$AWT_{3.5}$ = log transformed airway wall thickness at 3.5 mm diameter; CMH = chronic mucus hypersecretion; Log-LAA% -950 HU = log transformed percentage of low attenuation areas < -950 HU

Chapter 6

Chronic respiratory symptoms associate with airway wall thickening measured by thin-slice low-dose CT

X Xie, AE Dijkstra, JM Vonk, M Oudkerk, R Vliegenthart, HJM Groen

American Journal of Roentgenology, 2014

Abstract

Background

The prevalence of chronic respiratory symptoms is high in heavy smokers in lung cancer screening. The study-purpose is to compare CT-derived airway wall measurements between male smokers with and without chronic respiratory symptoms.

Materials and methods

50 heavy male smokers with chronic respiratory symptoms (cough, mucus, dyspnea and wheezing) and 50 without any respiratory symptom were randomly selected from the NELSON trial. Images on thin-slice low-dose CT were evaluated using dedicated software for airway measurements. Wall area percentage (WA%) and airway wall thickness (AWT) were measured from trachea to bronchi in five different pulmonary lobes of airways with a luminal diameter ≥ 5mm. Association between airway wall measurements and respiratory symptoms was analyzed using multiple linear regression, adjusted for age, body mass index, smoking status, emphysema and pulmonary function.

Results

After adjusting for relevant factors, a significant positive association between airway wall measurements and respiratory symptoms was found in airways with a 5-10mm luminal diameter ($p < 0.01$), but not in ≥ 10mm airways ($p > 0.05$). In the 5-10mm airway level, mean WA% was 51.5 ± 7.9% and 48.1 ± 7.7% in the symptomatic and the asymptomatic group ($p < 0.01$), respectively. AWT was 1.54 ± 0.39mm and 1.37 ± 0.35mm ($p < 0.001$), respectively.

Conclusion

Heavy male smokers with chronic respiratory symptoms in lung cancer screening, who are at high-risk for chronic bronchitis, have bronchial wall thickening in airways with luminal diameter from 5 to 10mm, but not in larger airways.

Introduction

Nearly 50% of smokers in lung cancer screening have chronic respiratory symptoms, i.e., chronic hyper secretion of mucus, combined with chronic cough, often accompanied by dyspnea and wheezing [1]. Smokers with these symptoms are of risk to develop chronic bronchitis, a disease associated with accelerated decline in pulmonary function – an important risk factor for chronic obstructive pulmonary disease (COPD) and all-cause mortality [2, 3]. During their lifetime, over 40% of smokers develop chronic bronchitis. The clinical diagnosis of chronic bronchitis is commonly based on a combination of medical history, physical examination, spirometry and laboratory test [4]. Despite the high prevalence, chronic bronchitis was often under-diagnosed or late-diagnosed [5].

Chronic bronchitis is histopathologically found in a range of airways, commonly in large airways [6]. The morphological basis of chronic bronchitis is bronchial wall thickening and airway luminal narrowing, which subsequently result in airflow limitation [7]. The observation of morphological changes is important to understand pathogenesis and effect of therapeutic interventions for chronic bronchitis [8]. Recent development of thin-slice multi-detector CT and dedicated software, allows accurate non-invasive quantification of airway dimensions in large bronchi [9-11].
In participants in a lung cancer screening study, it is important to know whether there are morphological changes of airway walls in subjects with chronic respiratory symptoms, since early treatment would be performed in case of airway remodeling of chronic bronchitis in this high-risk population. Airway wall thickening associated with some relevant factors, such as age, body mass index, smoking status, emphysema and pulmonary function [12, 13]. Orlandi et al has shown the relationship between airway wall thickening on CT and chronic bronchitis in COPD population [14]. However, in male heavy smokers, the representative population in lung cancer screening, the adjusted association between CT-derived airway wall quantification and respiratory symptoms is still unclear. Therefore, the purpose of our study was to retrospectively compare the airway wall thickness along the respiratory pathway between male heavy smokers with and without chronic respiratory symptoms, adjusted for relevant factors.

Materials and methods

Sample

The study sample was randomly selected from the baseline round in Groningen center of a population-based multi-centric Dutch-Belgian Randomized Lung Cancer Screening Trial (NELSON) (Figure 1). The symptomatic group contained 50 male heavy smokers (who smoked ≥ 15 cigarettes/day during ≥ 25 years or ≥ 10 cigarettes/day during ≥ 30 years) with four chronic respiratory symptoms, including chronic cough, chronic mucus hyper-secretion, dyspnea and wheezing lasting for at least three months during the last year before inclusion. The asymptomatic group contained 50 male smokers without any respiratory symptoms. The sample in each group was randomly selected by statistical software (SPSS version 20, IBM, New York, US). It was reported that wall area percentages (WA%) in large airways are 53.5% and 48.0% between symptomatic and asymptomatic groups, and standard deviation (SD) is 8.0 15. Assuming the results as similar as the reported data, a 50/50 sample size would provide a statistical power of > 0.90 with the confidence level of 95% in t-test when assessing WA% in large airways between these two groups. The NELSON trial was approved by the Dutch Minister of Health and the ethics board in the participating center. All participants gave written informed consent. Detailed inclusion and exclusion criteria and characteristics in this male population-based trial have been described elsewhere 16. In short, current or former heavy smokers between 50 and 75 years old were included. Individuals in moderate or poor health with inability to climb two flights of stairs were excluded. Information about the presence of respiratory symptoms and smoking behavior (current or former smoker, packyears, etc.) was obtained by questionnaires. The question asked to the participants to record their respiratory symptoms was "do you have experienced the following symptoms cough, sputum expectoration, wheezing or dyspnea for at least 3 months during the past year, even when you did not have a cold?"

CT imaging

A 16-row multi-detector CT (Sensation 16, Siemens, Forchheim, Germany) was utilized with a low-dose acquisition protocol [17]. The protocol was: spiral acquisition at 120kV, 20mAs, rotation time 0.5s, pitch 1.5 and collimation 16 × 0.6mm, field of view 350mm, slice thickness 1mm and slice increment 0.7mm. The effective radiation dose is less than 0.8mSv. Contrast media was not used. The images were reconstructed to a pixel matrix of 512 × 512 using a medium-smooth B30f kernel. The CT system was calibrated routinely. The image acquisition was performed during one breath-holding at full inspiration after appropriate instruction.

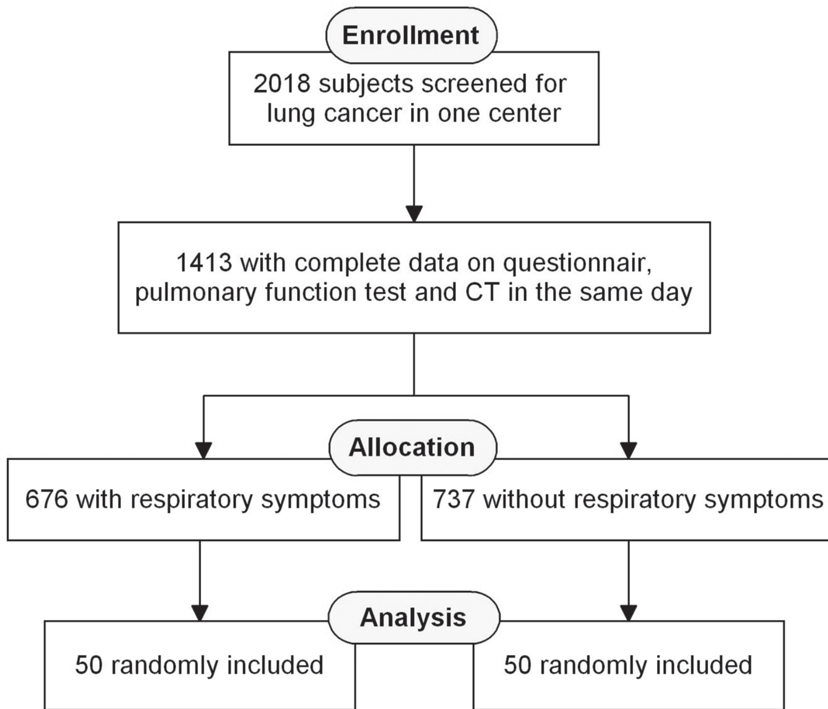


Figure 1. Flow diagram of subject selection

6

Pulmonary function testing

Standard pulmonary function testing was performed according to the European Respiratory Society guidelines on the same day as the CT acquisition [18]. In this population based trial, a bronchodilator was not administered. Forced expiratory volume in one second (FEV_1) and forced vital capacity (FVC) were assessed. FEV_1 was presented as a percentage of predicted (FEV_1, %pred).

Airway selection

We measured the airways from the tracheal, main and lobar bronchi, down to the segmental and sub-segmental bronchi with an internal luminal diameter ≥ 5mm, which were subsequently categorized into three categories (luminal diameter ≥ 15mm, between 10 and 15mm, and between 5 and 10mm). We selected five bronchi of segmental level, each representing one pulmonary lobe: the apical bronchus (RB1) of the right upper lobe, the lateral bronchus (RB4) of the right middle lobe, the posterior basal bronchus (RB10) of the right lower lobe, the apicoposterior bronchus (LB1+2) of the left upper lobe, and posterior basal bronchus (LB10) of the left lower lobe. Those bronchi were selected because they are relatively free from cardiac motion artifacts.

In the bronchial tree below segmental level, when there was more than one bronchus was present in each airway generation, we evaluated only one bronchus in each generation. Thereafter, the bronchial pathway was evaluated from trachea down to airway level of 5mm luminal diameter.
After randomization of all included subjects, one radiologist with nine years' experiences in thoracic diagnostic radiology selected airways and evaluated the images, blinded to subject information on basic characteristics and clinical data during evaluation. Time duration of whole assessment processes for each subject was approximately three minutes.

Quantitative Image Analysis

The images were evaluated using dedicated software for airway measurement (Airway Examiner 1.0, Fraunhofer MEVIS, Bremen, Germany), which was based on a three-dimensional (3D) algorithm of airway geometry, instead of the traditional full-width-at-half-maximum (FWHM) algorithm [11], Briefly, wall thickness is estimated in 3D space using this 3D algorithm, rather than 2D plane in FWHM method [11]. For airway wall thickness as small as 1mm, the 3D algorithm showed much better accuracy and reproducibility than the FWHW method in phantom studies [11, 19]. This software tool follows two principal steps. Firstly, the software automatically segments the bronchial tree. Secondly, after clicking a bronchus in the bronchial tree, the software automatically quantifies airway dimensions along trachea to the chosen bronchus, presenting data per 1mm along this respiratory pathway. A representative figure generated by the software package was shown in Figure 2. To avoid potential misleading segmentation of airway walls in branching points, the airways within 5mm from the branching point were not included. We collected three airway quantitative measurements, including WA%, airway wall thickness (AWT) and airway luminal diameter. WA% was defined as (wall area) / (wall area + lumen area) × 100%. Emphysema quantification was performed using dedicated software (ImageXplorer, Image Sciences Institute, Utrecht, The Netherlands). This software automatically implies lung segmentation, image noise reduction, and CT density calibration to improve accuracy and reproducibility of evaluation [20, 21]. The images were re-calibrated by shifting CT density so that the density inside the trachea became -1000 Hounsfield unit (HU) [22]. We collected three CT quantification parameters, including 15 percentile point of lung density (Perc15), percentage of lung attenuation area under -950HU (%LAA-950) and lung volume. Larger emphysematous tissues are indicated by lower Perc15 or higher %LAA-950.

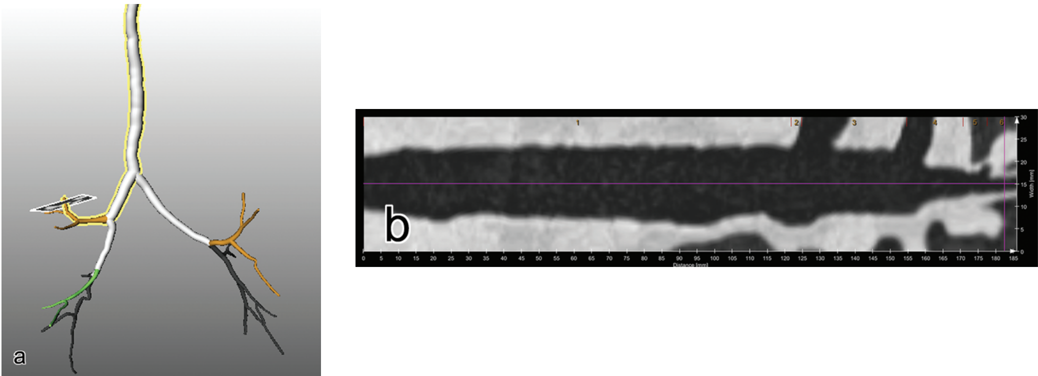


Figure 2. Reformatted images generated by the dedicated software. (a) The bronchial tree was automatically segmented. (b) A bronchial pathway from trachea to a selected bronchus was converted to a stretched multiple planar reconstruction (MPR) image.

Statistics

Data are reported as mean ± standard deviation (SD) for normally distributed data or median (25th, 75th percentile) for non-normally distributed data. Differences in characteristics between the symptomatic and asymptomatic group were assessed by independent-samples *t*-test for normally distributed continuous data, by Mann-Whitney U-tests for non-normally distributed continuous data, and by Chi-square test for nominal data. The association between airway wall measurements (WA% and AWT) and potentially associated factors was evaluated by univariate linear regression. The association between airway wall measurements and respiratory symptoms was analyzed using multiple linear regression, adjusted for age, BMI, smoking status, Perc15, and FEV_1, %pred.

A $p < 0.05$ was considered as statistically significant. Statistical analyses were performed using SPSS version 20 (IBM, New York, US).

Results

Sample characteristics

Characteristics of the symptomatic and asymptomatic group are presented in Table 1. All subjects were male, with a mean age of 56.5 ± 5.4 years (range from 50 to 69 years). No diseases (pneumonia and atelectasis, etc.) affecting airway wall measurements were observed through reviewing CT images. No pulmonary fibrosis affecting the observation of large bronchi wall was found. No obstructive diseases (airway mass/tumor, external compression and bronchial stricture, etc.) causing airway wall thickening were found either. Airway walls were successfully evaluated in all 100 subjects, in which, 65,070 cross sections of airways were measured. WA% increased from proximal to distal airway, while AWT and luminal diameter decreased. WA% ranged from 14.6 to 75.5%. AWT ranged from 0.7 to 3.2mm. Airway luminal diameter ranged from 5.0 to 22.9mm.

Table 1 Characteristics of included subjects with and without chronic respiratory symptoms

	Symptomatic	Asymptomatic	p-value
Sample size, *n*	50	50	
Basic characteristics			
Male, *n*	50	50	
Age, years	56.0 ± 5.1	57.3 ± 5.7	0.190
Weight, kg	86.9 ± 13.6	83.9 ± 12.5	0.253
Height, m	1.79 ± 0.05	1.79 ± 0.06	0.628
Body mass index, kg/m^2	27.0 ± 3.9	26.2 ± 3.3	0.282
Smoking behavior			
Current/former smoker, *n*	37/13	23/27	<0.05
Packyears	42.1 ± 15.5	39.9 ± 14.6	0.411
Smoking duration, years	8.3 ± 1.0	8.0 ± 1.1	0.130
CT quantification			
Wall thickness, mm	1.55 ± 0.44	1.42 ± 0.40	<0.001
Wall area percentage, %	47.0 ± 12.1	43.3 ± 11.1	<0.001
Lung volume, L	7.0 (6.3, 8.0)	6.6 (5.7, 7.5)	<0.001
Perc15, HU	-922 (-933, -912)	-915 (-928, -898)	<0.001
%LAA-950, %	3.2 (2.1, 5.5)	2.3 (1.0, 4.0)	<0.001
Pulmonary function test			
FEV_1, %pred	80.3 (65.4, 110.4)	102.9 (95.6, 108.2)	<0.001
FEV_1/FVC, %	65.5 (51.9, 72.4)	75.3 (69.6, 80.8)	<0.001
COPD, *n*	31	13	<0.001

Data are reported as mean ± standard deviation for normally distributed data or median (25th, 75th percentile) for non-normally distributed data. HU = Hounsfield unit; Perc15 = 15 percentile point of lung density; %LAA-950 = percentage of lung attenuation area under -950HU; FEV_1, %pred = forced expiratory volume in the first second as percentage from predicted; FEV_1/FVC = FEV_1 divided by forced vital capacity.

The symptomatic group had more current smokers than asymptomatic group (74% vs. 46%, $p < 0.05$). No significant differences were found for age, BMI, packyears and smoking duration between those two groups ($p > 0.05$). The symptomatic group had significantly worse pulmonary function (FEV_1, %pred and FEV_1/FVC) than the asymptomatic group ($p < 0.001$).

On CT emphysema quantification, the symptomatic group had significantly lower Perc15 and higher %LAA-950 than the asymptomatic group ($p < 0.001$), indicating more emphysematous tissues in the symptomatic group. Median %LAA-950 was 3.2 (25th, 75th percentile: 2.1, 5.5) and 2.3 (1.0, 4.0) in the symptomatic and asymptomatic group, respectively, indicating mild emphysema in those two groups [9].

Factors associated with airway wall measurements

Univariate linear regression analysis showed that thicker airway walls positively associated with the presence of respiratory symptoms, higher BMI, current smoking, higher packyears and longer smoking duration ($p<0.05$) (Table 2). Conversely, thinner airway walls positively associated with higher age, longer duration of smoking cessation and better pulmonary function (FEV_1, %pred and FEV_1/FVC) ($p < 0.01$).

6

Table 2. Univariate regression analysis for factors associated with airway wall measurements

	Airway wall thickness		Wall area percentage	
	B	**p-value**	**B**	**p-value**
Basic characteristics				
Age, years	-0.003	**<0.05**	-0.132	**<0.01**
Weight, kg	0	0.508	0.041	0.281
Height, m	0.151	0.373	1.940	0.667
Body mass index, kg/m²	0.012	**<0.05**	0.160	**<0.05**
Smoking behavior				
Current smoking	0.102	**<0.001**	2.418	**<0.001**
Packyears	0.004	**<0.001**	0.084	**<0.001**
Smoking duration, years	0.021	**<0.001**	0.233	**<0.01**
Duration of smoking cessation, years	-0.029	**<0.001**	-0.523	**<0.001**
Chronic respiratory symptoms				
Presence of respiratory symptoms	0.124	**<0.001**	3.742	**<0.001**
CT emphysema quantification				
Lung volume, L	0.021	0.244	-0.163	0.361
Perc15, HU	0	0.259	0.021	0.132
%LAA-950, %	0.002	0.242	-0.041	0.430
Pulmonary function test				
FEV_1, %pred	-0.005	**<0.001**	-0.120	**<0.001**
FEV_1/FVC, %	-0.007	**<0.001**	-0.164	**<0.001**

HU = Hounsfield unit; Perc15 = 15 percentile point of lung density; %LAA-950 = percentage of lung attenuation area under -950HU; FEV_1, %pred = forced expiratory volume in the first second as percentage from predicted; FEV_1/FVC = FEV_1 divided by forced vital capacity; bold is p-value < 0.05.

Airway wall thickness differences between the symptomatic and asymptomatic group

Without adjustment for relevant factors, the symptomatic group showed overall higher WA% (47.0 ± 12.1% vs. 43.3 ± 11.1%, $p < 0.001$) and higher AWT (1.55 ± 0.44mm vs. 1.42 ± 0.40mm, $p < 0.001$) than the asymptomatic group. In detail, in all the three categories of large airways (luminal diameter ≥ 15mm, between 10 and 15mm, and between 5 and 10mm), the symptomatic group showed significantly higher WA% and AWT ($p < 0.01$) (Figure 3). The representative images of those two groups are shown in Figure 4.

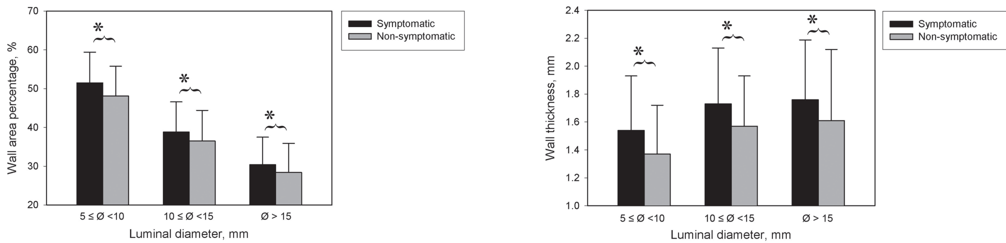


Figure 3. Airway wall area percentage (**a**) and airway wall thickness (**b**) of symptomatic and asymptomatic subjects with chronic respiratory symptoms, without adjusting for relevant factors. Significant difference between symptomatic and asymptomatic group was indicated as "*".

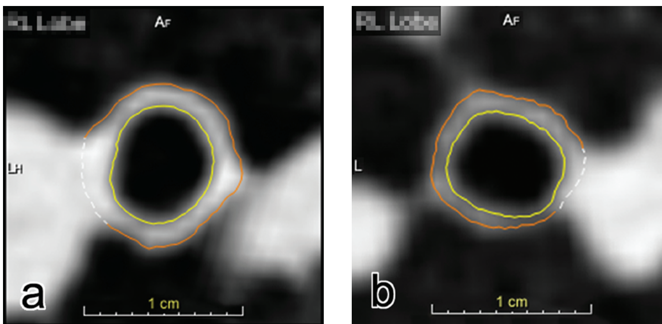


Figure 4. Cross-section images perpendicular to long axis of the bronchi. (a) An image of a symptomatic subject. Wall area percentage (WA%) = 51%, airway wall thickness (AWT) = 1.4 mm and luminal diameter = 7 mm. (b) An image of an asymptomatic subject. WA% = 43%, AWT = 1.1 mm and luminal diameter = 7 mm.

However, when using multiple linear regression, after adjustment for age, BMI, smoking status, Perc15, and FEV_1, %pred, a significant positive association between thicker airway walls (WA% and AWT) and the presence of respiratory symptoms was found only in airways with a luminal diameter from 5 to 10mm ($p < 0.01$). In the airway level from 5 to 10mm, mean WA% was 51.5 ± 7.9% and 48.1 ± 7.7% in the symptomatic and the asymptomatic group ($p < 0.01$), respectively. AWT was 1.54 ± 0.39mm and 1.37 ± 0.35mm ($p < 0.001$), respectively. No significant associations were found in airways with a luminal diameter ≥ 10mm ($p > 0.05$) (Table 3).

Table 3. Multiple linear regression for the association between airway wall measurements and chronic respiratory symptoms, adjusted for age, BMI, current smoking status, Perc15, and FEV_1, %pred.

		Airway luminal diameter, mm					
		5 ≤ diameter < 10		**10 ≤ diameter < 15**		**diameter ≥ 15**	
		B	**p-value**	**B**	**p-value**	**B**	**p-value**
Wall thickness	Presence of respiratory symptoms	0.091	**<0.001**	0.051	0.170	0.052	0.339
	Age	0.004	0.075	0.003	0.409	-0.013	**<0.01**
	BMI	0.014	**<0.001**	0.002	0.646	-0.009	0.186
	Current smoking	0.019	0.461	0.159	**<0.001**	-0.014	0.793
	Perc15	0.003	**<0.001**	0	0.618	-0.001	0.228
	FEV_1, %pred	-0.006	**<0.001**	-0.004	**<0.001**	-0.003	**<0.01**
Wall area percentage	Presence of respiratory symptoms	1.647	**<0.01**	0.054	0.944	0.551	0.500
	Age	0.042	0.348	-0.065	0.347	-0.164	<0.05
	BMI	0.333	**<0.001**	0.015	0.881	-0.131	0.219
	Current smoking	0.509	0.330	3.084	**<0.001**	-0.202	0.801
	Perc15	0.077	**<0.001**	-0.017	0.332	-0.019	0.305
	FEV_1, %pred	-0.130	**<0.001**	-0.076	**<0.001**	-0.059	**<0.001**

BMI = body mass index; Perc15 = 15 percentile point of lung density; FEV1, %pred = forced expiratory volume in the 1st second as percentage from predicted; bold is p-value < 0.05.

Discussion

Thin-slice CT and automated software are promising tools to quantify airway walls. Using these techniques, after adjustment for relevant factors, we showed that heavy smokers with chronic respiratory symptoms had significant thicker airway walls in airways with luminal diameter from 5 to 10mm, but not in the larger airways. If not adjusted for, the thicker airway walls were in all ≥ 5mm airways.

The symptomatic group showed general thicker airway walls, up to the trachea. The common causes of bronchial wall thickening are inflammatory, congenital (e.g., cystic fibrosis, alpha-1 antitrypsin deficiency, etc.) and obstructive bronchial diseases [23]. Inflammation of the mucous membrane directly results in hyper-secretion of mucus, leading to respiratory symptoms, such as cough, dyspnea and wheezing [24]. Our sample was from a population-based trial, and the prevalence of these congenital bronchial diseases is rare [25]. An experienced radiologist reviewed the CT images, and did not observe obstructive bronchial diseases. Thus, the primary cause of bronchial wall thickening in our study is likely inflammatory. Chronic bronchitis is a disease associated with long-term inflammatory stimulation [6].

The histological evidence indicates that inflammation and airway remodeling, associated with chronic bronchitis, is located in the more central airways [26]. Using CT quantification in a COPD population instead of heavy smokers, Patel, et al. observed a significant positive association between airway wall thickness and respiratory symptoms in approximately 6mm airways in luminal diameter [27]. Mair, et al. found a significant positive association in proximal airways (> 11mm in luminal diameter), but not in distal airways (approximately 2 to 4mm) [15]. Thus, we investigated bronchial walls with a luminal diameter of ≥ 5mm. On the other hand, thickening of smaller airway walls is important for the pathogenesis of COPD and asthma [28, 29]. Increased CT-derived airway wall thickness of more peripheral airways, as small as 2 to 4mm in luminal diameter, strongly correlated with airflow limitations in those diseases [22, 30, 31].

Importantly, we adjusted for five relevant factors to determine the adjusted association between thicker airway walls and chronic respiratory symptoms. In a study at the same age group as our study, smokers had thicker bronchial walls than non-smokers [32]. Smoking often causes more airway inflammation, and is an important potential confounder when investigating airway wall thickening [33]. Thus, it is essential to adjust for smoking behavior. Next, in accordance with the previous studies where age, BMI and pulmonary function were associated with airway wall thickening [13, 34, 35], we also found that thicker airway walls were significantly associated with current smoking, younger age, higher

BMI and worse pulmonary function by univariate linear regression. Inconsistent with prior studies where emphysema was associated with airway wall thickening [13, 36], we found a non-significant association between CT emphysema quantification and airway wall thickening. That inconsistent finding might be explained by the presence of only mild emphysema in our sample, selected from a population-based screening trial. After adjustment for these mentioned potential confounders, we found significant thicker airway walls in airways with luminal diameter from 5 to 10mm, but not in the larger airways.

A recently introduced three-dimensional algorithm was used to assess airway walls in this study. With this algorithm, wall thickness is approximated by an integral based closed-form solution, based on the volume conservation property of convolution [11]. In contrast, the traditionally utilized "full-width-at-half-maximum (FWHM)" algorithm calculates the x-ray attenuation values along rays placed from the lumen center to outward directions in two-dimensional cross section [10]. The former algorithm has shown much better accuracy and reproducibility than the latter, for wall thickness as small as 1mm in phantom studies [11, 19]. Accuracy of bronchial wall quantification in CT depends on other factors, such as image noise. Lutey et al. showed that partial volume effects had more influence on bronchial wall measurements in smaller airways than larger airways [37]. Diaz et al. showed that emphysema had more influence in smaller airways [38]. In a phantom study based on the three-dimensional algorithm in thin-slice low-dose CT (slice thickness 0.9–1.25mm, tube current as low as 18mAs), the average wall thickness error for a tube (luminal diameter 3.4 mm, wall thickness 0.9 mm) was 4.4% [19]. This phantom study simulated thin-slice low-dose settings with considerable image noises, and the airways as small as 3.4mm in luminal diameter, which were similar to our methodology, thus we expected that our measurements were accurate. Our results were based on 16-row multi-detector CT scanners with images of 512 × 512 voxels. This image matrix is widely available in current clinical practice. The latest CT technique provides higher spatial resolution thus improve airway wall assessment [39]. For example, image matrix of 1024 × 1024 voxels might result in more accurate measurements.

CT quantification is associated with pathophysiological changes of airway remodeling. A number of CT-derived measurements have been used for airway quantification, such as thickness, area, perimeter, CT density and visual score [12, 27, 40]. We measured WA% and AWT, because these two measurements directly indicate, and pathologically reflect airway wall thickening [29]. Moreover, we utilized automated dedicated software to evaluate airway dimensions per 1mm. Some prior studies used manual methods to quantify airway walls, where airway walls were non-continuously measured on cross-sections in-between a large gap of up to 20mm [14, 27]. Mair, et al. quantified airway walls,

as a function to airway generation [15]. In addition to that study, we evaluated airway walls as a function to airway luminal diameter, since the same airway generation in two bronchi might be of different size.

Clinical implications

In lung cancer screening trials, the common participants are heavy smokers, with a high prevalence of chronic bronchitis. Despite the high prevalence, chronic bronchitis was often under-diagnosed or late-diagnosed [5]. We expect to use CT bronchial wall quantification among smokers in screening, which is the important population for early diagnosis and treatment of chronic bronchitis, such as the participants in the NELSON trial. In this screening trial, individuals in moderate or poor health were excluded, because participants need to have enough cardiopulmonary reserve to undergo surgery. Thus, the population in the NELSON trial is representative for general heavy smokers. One morphological manifestation of chronic bronchitis is bronchial wall thickening caused by chronic inflammatory stimulation [6]. We found that heavy smokers with chronic respiratory symptoms had significant thicker airway walls, which represents airway remodeling in an inflammatory process. Our study shows that this airway remodeling can be detected using thin-slice CT and dedicated software, thus this method has potential benefit for early diagnosis of chronic bronchitis.

Currently, clinical symptoms and spirometry are commonly used for diagnosis and surveillance of chronic bronchitis [41]. Non-invasive CT quantification of airway walls has shown the potential for regional and morphological evaluation of the therapeutic response in treatment of chronic bronchitis [8]. In our study, significantly thicker airway walls are especially identified in large airways of 5 to 10 mm in luminal diameter in the respiratory symptomatic group. Thus, CT quantification of airway walls may provide additional morphological information beyond clinical symptoms and spirometry. When assessing airway wall thickening in symptomatic individuals, the airways of 5 to 10mm in luminal diameter optimally reflect the presence of respiratory symptoms, but not the larger airways.

The absolute increase of airway wall thickening in symptomatic subjects is commonly within a millimeter that is difficult to be perceptible to human eyes on CT images. Kim et al showed significant association between visual and quantitative assessment [42]. In screening, a quantitative method is important across different observers and during follow-up of patients, since the small changes have to be recorded accurately and reproducibly. Hence, dedicated software is a good candidate to screen airway walls.

Limitations

Firstly, inherent to a population-based lung cancer screening trial, histopathological results of bronchial wall are very hard to be available. At least, our results suggested that the chronic respiratory symptoms associated with airway remodeling caused by an inflammatory process, which is a pathological basis for chronic bronchitis. Also, participants were only heavy male smokers with mild emphysema. Gender was not adjusted for in this study. Worldwide it is estimated that men smoke nearly five times as much as women [43]. Lung cancer shows higher incidence and mortality rates of the disease among men than women [44]. Due to higher prevalence of heavy smoking and lung cancer, male smokers are more often screened and examined by CT. Our study was only performed in male. Whether the results are generalizable to female should be investigated in the future.

Secondly, a post-bronchodilator pulmonary function test was not performed to assess reversibility in airflow limitation, which is a criterion to exclude bronchial asthma [28]. Our sample was from a population-based trial, and the prevalence of asthma in elderly men is approximately 2% in the Netherlands [45]. Thus, our results might not be substantially influenced by that limitation.

Thirdly, five bronchi from different pulmonary lobes were evaluated because they are relatively free from cardiac motion artifacts, instead of bronchi from all the pulmonary lobes. A large number of airway cross sections (650 per subject) were measured. We expected that five bronchi with a large number of measurements could represent the quantification of airway dimensions.

Conclusions

After adjustment for relevant factors, male heavy smokers with chronic respiratory symptoms from a population-based lung cancer screening trial have significantly thicker bronchial walls than asymptomatic smokers in airways with a luminal diameter from 5 to 10mm, but not in larger airways. Thus, male heavy smokers with chronic respiratory symptoms do have airway remodeling. Thin-slice CT and dedicated software showed the potential to evaluate airway remodeling in smokers with chronic respiratory symptoms.

References

1 Mets OM, Buckens CF, Zanen P et al. Identification of chronic obstructive pulmonary disease in lung cancer screening computed tomographic scans. JAMA: the journal of the American Medical Association. 2011;306(16):1775-1781

2 Pelkonen M. Smoking: relationship to chronic bronchitis, chronic obstructive pulmonary disease and mortality. *Current opinion in pulmonary medicine*. 2008;14(2):105-109

3 Guerra S, Sherrill DL, Venker C, Ceccato CM, Halonen M and Martinez FD. Chronic bronchitis before age 50 years predicts incident airflow limitation and mortality risk. *Thorax*. 2009;64(10):894-900

4 Laitinen LA and Koskela K. Chronic bronchitis and chronic obstructive pulmonary disease: Finnish National Guidelines for Prevention and Treatment 1998-2007. *Respiratory medicine*. 1999;93(5):297-332

5 Huchon GJ, Vergnenegre A, Neukirch F, Brami G, Roche N and Preux PM. Chronic bronchitis among French adults: high prevalence and underdiagnosis. *The European respiratory journal*. 2002;20(4):806-812

6 Hogg JC. Pathophysiology of airflow limitation in chronic obstructive pulmonary disease. *Lancet*. 2004;364(1474-547; 9435):709-721

7 Hogg JC and Timens W. The pathology of chronic obstructive pulmonary disease. *Annual review of pathology*. 2009;4:435-459

8 Hamada S, Ito Y, Imai S, Oguma T, Niimi A and Mishima M. Effect of inhaled corticosteroid therapy on CT scan-estimated airway dimensions in a patient with chronic bronchitis related to ulcerative colitis. *Chest*. 2011;139(4):930-932

9 Shaker SB, Dirksen A, Bach KS and Mortensen J. Imaging in chronic obstructive pulmonary disease. *COPD*. 2007;4(2):143-161

10 Mets OM, de Jong PA, van Ginneken B, Gietema HA and Lammers JW. Quantitative computed tomography in COPD: possibilities and limitations. *Lung*. 2012;190(2):133-145

11 Weinheimer O, Achenbach T, Bletz C, Duber C, Kauczor HU and Heussel CP. About objective 3-d analysis of airway geometry in computerized tomography. *IEEE Transactions on Medical Imaging*. 2008;27(0278-0062; 0278-0062; 1):64-74

12 Xie X, de Jong PA, Oudkerk M et al. Morphological measurements in computed tomography correlate with airflow obstruction in chronic obstructive pulmonary disease: systematic review and meta-analysis. *European radiology*. 2012;22(10):2085-2093

13 Kim WJ, Silverman EK, Hoffman E et al. CT metrics of airway disease and emphysema in severe COPD. *Chest*. 2009;136(2):396-404

14 Orlandi I, Moroni C, Camiciottoli G et al. Chronic obstructive pulmonary disease: thin-section CT measurement of airway wall thickness and lung attenuation. *Radiology*. 2005;234(0033-8419; 2):604-610

15 Mair G, Maclay J, Miller JJ et al. Airway dimensions in COPD: Relationships with clinical variables. *Respiratory medicine*. 2010;(1532-3064; 0954-6111)

16 Xie X, Heuvelmans MA, van Ooijen PM, Oudkerk M and Vliegenthart R. A practical approach to radiological evaluation of CT lung cancer screening examinations. *Cancer imaging : the official publication of the International Cancer Imaging Society*. 2013;13(3):391-399

17 Xie X, Zhao Y, Snijder RA et al. Sensitivity and accuracy of volumetry of pulmonary nodules on low-dose 16- and 64-row multi-detector CT: an anthropomorphic phantom study. *European radiology*. 2013;23(1):139-147

18 Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R and Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *The European respiratory journal.Supplement*. 1993;16:5-40

19 Michael Schmidt, Jan-Martin Kuhnigk, Stefan Krass, Michael Owsijewitsch, Bartjan de Hoop and Heinz-Otto Peitgen. Reproducibility of airway wall thickness measurements. *Proc SPIE*. Mar 09, 2010;762410

20 Schilham AM, van Ginneken B, Gietema H and Prokop M. Local noise weighted filtering for emphysema scoring of low-dose CT images. *IEEE Transactions on Medical Imaging*. 2006;25(4):451-463

21 van Rikxoort EM, de Hoop B, Viergever MA, Prokop M and van Ginneken B. Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. *Medical physics*. 2009;36(7):2934-2947

22 Dijkstra AE, Postma DS, Ten Hacken N et al. Low-dose CT measurements of airway dimensions and emphysema associated with airflow limitation in heavy smokers: a cross sectional study. *Respiratory research*. 2013;14(1):11

23 Hyde DM, King TE,Jr, McDermott T et al. Idiopathic pulmonary fibrosis. Quantitative assessment of lung pathology. Comparison of a semiquantitative and a morphometric histopathologic scoring system. *The American Review of Respiratory Disease*. 1992;146(4):1042-1047

24 Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. American Thoracic Society. *American journal of respiratory and critical care medicine*. 1995;152(5 Pt 2):S77-121

25 Rahaghi FF, Sandhaus RA, Brantly ML et al. The prevalence of alpha-1 antitrypsin deficiency among patients found to have airflow obstruction. *Copd*. 2012;9(4):352-358

26 Hogg JC, Chu F, Utokaparch S et al. The nature of small-airway obstruction in chronic obstructive pulmonary disease. *The New England journal of medicine*. 2004;350(1533-4406; 0028-4793; 26):2645-2653

27 Patel BD, Coxson HO, Pillai SG et al. Airway wall thickening and emphysema show independent familial aggregation in chronic obstructive pulmonary disease. *American Journal of Respiratory & Critical Care Medicine*. 2008;178(5):500-505

28 Rabe KF, Hurd S, Anzueto A et al. Global strategy for the diagnosis, management, and prevention of COPD: GOLD executive summary. *American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association*. 2007;176(1073-449; 6)

29 Aysola RS, Hoffman EA, Gierada D et al. Airway remodeling measured by multidetector CT is increased in severe asthma and correlates with pathology. *Chest*. 2008;134(6):1183-1191

30 Matsuoka S, Kurihara Y, Yagihashi K, Hoshino M and Nakajima Y. Airway dimensions at inspiratory and expiratory multisection CT in chronic obstructive pulmonary disease: correlation with airflow limitation. *Radiology*. 2008;248(1527-1315; 3):1042-1049

31 Hasegawa M, Nasuhara Y, Onodera Y et al. Airflow limitation and airway dimensions in chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association*. 2006;173(1073-449; 12):1309-1315

32 Lederlin M, Laurent F, Dromer C, Cochet H, Berger P and Montaudon M. Mean bronchial wall attenuation value in chronic obstructive pulmonary disease: comparison with standard bronchial parameters and correlation with function. *AJR.American journal of roentgenology*. 2012;198(4):800-808

33 Battaglia S, Mauad T, van Schadewijk AM et al. Differential distribution of inflammatory cells in large and small airways in smokers. *Journal of clinical pathology*. 2007;60(8):907-911

34 Donohue KM, Hoffman EA, Baumhauer H et al. Asthma and lung structure on computed tomography: the Multi-Ethnic Study of Atherosclerosis Lung Study. *The Journal of allergy and clinical immunology*. 2013;131(2):361-8.

35 Zach JA, Newell JD,Jr, Schroeder J et al. Quantitative computed tomography of the lungs and airways in healthy nonsmoking adults. *Investigative radiology*. 2012;47(10):596-602

36 Kim YI, Schroeder J, Lynch D et al. Gender differences of airway dimensions in anatomically matched sites on CT in smokers. *Copd*. 2011;8(4):285-292

37 Lutey BA, Conradi SH, Atkinson JJ et al. Accurate Measurement of Small Airways on Low Dose Thoracic CT Scans in Smokers. *Chest*. 2012

38 Diaz AA, Han MK, Come CE et al. Effect of emphysema on CT scan measures of airway dimensions in smokers. *Chest*. 2013;143(3):687-693

39 Obenauer S, Dullin C, Alves F, Missbach-Guentner J, Grabbe E and Heuser M. Flat-panel-detector-based volumetric CT: performance evaluation of imaging for skeletal structures of small animals in comparison to multislice CT. *Clinical imaging*. 2007;31(1):18-22

40 Yamashiro T, Matsuoka S, Estepar RS et al. Quantitative assessment of bronchial wall attenuation with thin-section CT: An indicator of airflow limitation in chronic obstructive pulmonary disease. *AJR.American journal of roentgenology*. 2010;195(2):363-369

41 Blasi F, Ewig S, Torres A and Huchon G. A review of guidelines for antibacterial use in acute exacerbations of chronic bronchitis. *Pulmonary pharmacology & therapeutics*. 2006;19(5):361-369

42 Kim SS, Seo JB, Lee HY et al. Chronic obstructive pulmonary disease: lobe-based visual assessment of volumetric CT by Using standard images--comparison with quantitative CT and pulmonary function test in the COPDGene study. *Radiology*. 2013;266(2):626-635

43 WHO. Gender empowerment and female-to-male smoking prevalence ratios. *http://www.whoint/bulletin/volumes/89/3/10-079905/en/* Accessed on December 1, 2013

44 WHO. *http://www.whoint/gender/documents/tobacco/9241592524/en/* Accessed on December 1, 2013

45 Nationaal Kompas Volksgezondheid [National Public Health Compass]. *http://www.nationaalkompasnl/gezondheid-en-ziekte/ziekten-en-aandoeningen/ademhalingswegen/astma/omvang* Accessed on June 1, 2013

Chapter 7

Novel genes for airway wall thickness identified with combined genome wide association and expression analyses

AE Dijkstra, DS Postma, B van Ginneken, MO Wielpütz, M Schmidt, N Becker, M Owsijewitsch, HU,Kauczor, HJ de Koning, JWJ Lammers, M Oudkerk, CA Brandsma, Y Bossé, DC Nickle, DD Sin, PS Hiemstra, C Wijmenga, J Smolonska, P Zanen, JM Vonk, M van den Berge, HM Boezen, HJM Groen

Submitted

Abstract

Rationale

Airway wall thickness (AWT) is affected by both environmental and genetic factors and is strongly associated with airflow limitation in smaller airways.

Objectives

The aim of our study was to investigate its genetic component.

Methods

AWT was measured on low-dose CT-scans in male heavy smokers participating in a lung cancer screening study (n = 2,640). Genome wide association studies on AWT were performed under an additive model using linear regression (adjusted for packyears, lung volume), followed by meta-analysis. An independent cohort was used for validation of the most strongly associated single nucleotide polymorphisms (SNPs). The functional relevance of significant SNPs was evaluated.

Measurements and main results

Three significant loci on chromosomes 2q (rs734556, $p = 6.2 \times 10^{-7}$) and 10q (rs10794108, $p = 8.6 \times 10^{-8}$; rs7078439, $p = 2.3 \times 10^{-7}$) were associated with AWT and confirmed in the meta-analysis in cohorts with comparable lung function: p-values 4.6×10^{-8}, 7.4×10^{-8} and 7.5×10^{-8}, respectively. SNP rs734556 was associated with decreased lung tissue expression of *SERPINE2*, a susceptibility gene for emphysema. Two nominally significant SNPs showed effects with similar direction: rs10251504 in *MAGI2* ($p = 5.8 \times 10^{-7}$) and rs4796712 in *NT5C3B* ($p = 3.1 \times 10^{-6}$). Higher *MAGI2* expression in bronchial biopsies of COPD patients was significantly associated with lower inflammatory cell numbers, lower *NT5C3B* expression with worse lung function. The *NT5C3B* risk allele was associated with higher lung tissue expression ($p = 1.09 \times 10^{-41}$).

Conclusions

Genetic variants contribute to AWT. Amongst others, the identified genes are involved in emphysema, airway obstruction and bronchial inflammation.

Introduction

Airway wall thickening can occur over the total length of the respiratory tract and is associated with chronic mucus hypersecretion (CMH) in larger airways and with airway obstruction in smaller airways [1]. The pathologic process underlying airway wall thickening is chronic inflammation and remodeling of the airway wall due to external factors like cigarette smoke and occupational exposures.
Not every heavy smoker develops airway wall thickening and subsequent airway obstruction. Therefore, a genetic predisposition is likely to play a role in the origin of this phenomenon. This is supported by a familial aggregation study [2] and by the association of several COPD candidate genes with airway wall thickening in another study [3].

In the past, knowledge on the process of airway wall thickening was mainly obtained through autopsy and bronchoscopic biopsy studies. Nowadays, computed tomography (CT) can be used to more accurately measure the dimensions of the airway wall. Previous studies using low-dose CT have assessed airway dimensions, particularly in relation to airflow limitation, respiratory symptoms, emphysema and smoking habits [4]. Research by Nakano et al. revealed that CT measurements of larger airways could be used to estimate the dimensions of the small conducting airways [1]. Therefore, airway wall thickness measurements may reflect the dimensions of smaller airways, the main source of airway obstruction in COPD [1].

The aim of the present study was to identify which genetic variants are associated with increased airway wall thickness measured with low-dose CT in a cohort of male heavy current and ex-smokers participating in the Dutch-Belgian lung cancer screening trial (NELSON). We subsequently validated our findings in the German lung cancer screening intervention trial (LUSI), thereby obtaining better insights in the origins of airway wall thickening that contributes to the development of COPD [5, 6].

Methods

Ethics Statement

The Dutch Ministry of Health and the Medical Ethics Committee of the hospital approved the study protocol for all Dutch centers. Ethics approval and written informed consent was obtained from all participants in the studies.

Population

Male participants from Groningen and Utrecht were recruited from the Dutch NELSON study, a heavy smoking population-based lung cancer screening trial. Detailed inclusion criteria and characteristics have been described elsewhere [5]. In short, individuals with a smoking history of ≥ 20 packyears obtained by a standardized questionnaire were included. In order to confirm the results of the analyses performed in participants of in the NELSON study, additional analyses were performed in subjects participating in the German lung cancer screening intervention trial (LUSI), an epidemiological study among men and women with a history of heavy smoking (≥ 20 packyears) [6].

CT-scanning and defining groups

Low-dose CTs of the chest were acquired in full inspiration after appropriate instruction without using contrast medium. CT images were attained with 16-row detector scanners (Sensation 16, Siemens Medical Solutions, Forchheim, Germany) (Groningen NELSON population, group I) or Brilliance 16P (Philips Medical Systems, Cleveland, OH, USA) (Utrecht NELSON population, group II). CT acquisition for the LUSI trial was performed from 2007 to 2010 with a 16-row scanner (Aquilion 16, Toshiba Corp., Tokyo, Japan) (LUSI, group III), and from 2010 - 2012 with a 128-row detector scanner (Somatom Definition Flash, Siemens Medical Solutions, Forchheim, Germany) (LUSI, group IV). All CT systems were calibrated routinely. CT scanning settings in NELSON and LUSI were previously described [5, 6].

Lung function

Spirometry according to the European Respiratory Society guidelines [7], including forced expiratory volume in 1 sec (FEV_1) and forced vital capacity (FVC) was performed at the start of the study.

Quantification of airway dimensions and lung volume

Airway wall thickness (AWT) was measured in cross-sectionally reformatted images with an automated research software prototype MEVIS Airway Examiner v1.0 (release 2009, Fraunhofer MEVIS, Bremen, Germany) at locations with an internal diameter of 3.5 mm in each lung lobe as described previously [4]. More detailed information is provided in the supplement.

Quantification of lung volume was based on automatic lung segmentation provided by a software tool called ImageXplorer (Image Sciences Institute, Utrecht, the Netherlands) [8]. CT scans were evaluated for appropriate segmentation. The mean AWT at 3.5 mm internal lumen size ($AWT_{3.5}$) of all five lobes per case was calculated taking into account the fraction of perimeter where the outer wall border was identified (Assessed Perimeter Fraction, APF) per lobe by the following formulae: ((AWT left upper lobe x APF left upper lobe) + (AWT left lower lobe x APF left lower lobe) + (AWT right upper lobe x APF right upper lobe) + (AWT right middle lobe x APF right middle lobe) + (AWT right lower lobe x APF right lower lobe)) / sum of APF of all lung lobes, as published previously [4]. $AWT_{3.5}$ for the whole population is not normally distributed, therefore we report median AWT and range, and log-transformed AWT was used in the analyses.

GWA study in the identification cohort

Group I and II individuals were genotyped using the Illumina Quad 610 array containing > 620,000 single nucleotide polymorphisms (SNPs). A GWA study of AWT was performed separately in group I and group II, in order to correct for differences between used CT-scanners. Subsequently, results of these analyses were meta-analyzed.

Replication of top SNPs in an independent cohort

Forty-eight SNPs ($p < 10^{-4}$) not in strong linkage disequilibrium ($r^2 \geq 0.80$) with other top SNPs were genotyped in groups III and IV using a custom made VeraCode assay (Illumina). As two different scanners were used in groups III and IV, two separate replication analyses were performed using a similar model with additional adjustment for sex (as also females were included in this cohort). Finally, a meta-analysis was performed on top SNPs across groups I, II, III and IV.

In order to replicate the findings in homogeneous populations, the analysis was repeated by
- excluding group III and IV women
- selecting individuals of groups III and IV with lung function values of FEV_1/FVC < 80%, comparable to the lung function values in groups I and II.

Functional relevance of the identified top SNPs

We assessed whether the identified top SNPs were expression quantitative trait loci (eQTLs) by analyzing the association of gene expression levels with SNP-genotypes in lung tissue from three independent cohorts recruited from Laval University, University of British Columbia, and University of Groningen as described previously [9].

Additionally, we assessed whether lung function (FEV_1 % predicted) and bronchial biopsy inflammatory markers were associated with mRNA expression in airway wall biopsies from 79 COPD participants in the GLUCOLD-study [10, 11].

Details on the methods of the functional studies are given in the online supplement.

Statistical analysis

General characteristics of the participants and differences between the cohorts were calculated with SPSS 20.0.

Quality control (QC), regression- and meta-analyses were performed with PLINK 1.07 [12]. SNPs with call rate < 95%, Minor Allele Frequency (MAF) < 0.05, proportion of individuals for which no genotype was called (mind) < 0.95 and Hardy Weinberg equilibrium (HWE) < 0.0001 were excluded. Ethnic outliers, duplicates and relatives were removed. In LUSI, QC was similar except for mind, which was set to < 0.9 (exclusion of individuals with < 90% of genotypes).

Linear regression analysis under an additive genetic model, with adjustment for packyears and lung volume, was used to identify SNPs associated with AWT. SNPs were included for replication if there was a strong association with AWT (top SNP; p-value < 10^{-4}). When two top SNPs were in strong linkage disequilibrium ($r^2 \geq 0.8$), the SNP with the lowest p-value was followed up. Meta-analysis was performed using a fixed effect model.

Results

Study populations

Characteristics of the identification and replication populations are presented in Table 1. After QC, 1,513 individuals in group I and 1,127 individuals in group II, and 522,636 SNPs were included in the analyses. In the replication analyses 714 individuals, 488 in group III and 226 in group IV, were included after QC. Median AWT on CTs was comparable in all cohorts studied. In groups I and II 82% of cases had an FEV_1/FVC < 80% and 52% of cases in groups III and IV.

For comparison with group I and II we provide the characteristics of subgroups of III and IV with males only, or including individuals with an FEV_1/FVC < 80% in the supplement, Table 1.

Table 1. Demographic and clinical characteristics of the groups I and II NELSON identification and groups III and IV LUSI replication populations

	NELSON		LUSI	
	Group I	**Group II**	**Group III**	**Group IV**
N	1,513	1,127	488	226
Characteristics				
Age, years	59.9 (5.4)	60.8 (5.5)	57.1 (5.2)	53.1 (4.4)
Height, cm	178.9 (6.1)	178.2 (6.6)	173.3 (8.6)	173.2 (8.6)
Male gender, %	100	100	66.0	59.9
Smoking				
Packyears smoking	38 (21-140)	40 (22-140)	37 (19-146)	33 (19-104)
Current smoking, %	57.2	55.0	60.2	67.6
Lung function				
FEV_1, liter	3.46 (0.74)	3.28 (0.71)	2.88 (0.78)	2.93 (0.78)
FEV_1/FVC, %	71.9 (9.8)	71.4 (9.4)	78.6 (9.5)	79.3 (11.7)
FEV_1, % predicted	99.2 (18.7)	95.8 (17.7)	91.6 (19.6)	91.2 (17.0)
CT measurements				
Median AWT, mm	0.57 (0.28-1.72)	0.60 (0.28-1.76)	0.57 (0.31-1.36)	0.60 (0.34-1.20)
Lung volume, liter	6.71 (1.2)	6.83 (1.4)	5.52 (1.8)	5.80 (1.3)

Mean (± standard deviation) shown for continuous data and median (range) for non-parametric distribution. AWT = airway wall thickness at 3.5 mm internal lumen size.

Identification and replication analysis

Genome wide analysis in groups I and II and the subsequent meta-analysis showed 69 SNPs to be associated with AWT (p-value $\leq 10^{-4}$). The QQ-plot provided no evidence of population stratification (λ = 1.012) (Figure 1). Genome wide association for AWT ordered by chromosome is shown in the Manhattan plot (Figure 1).

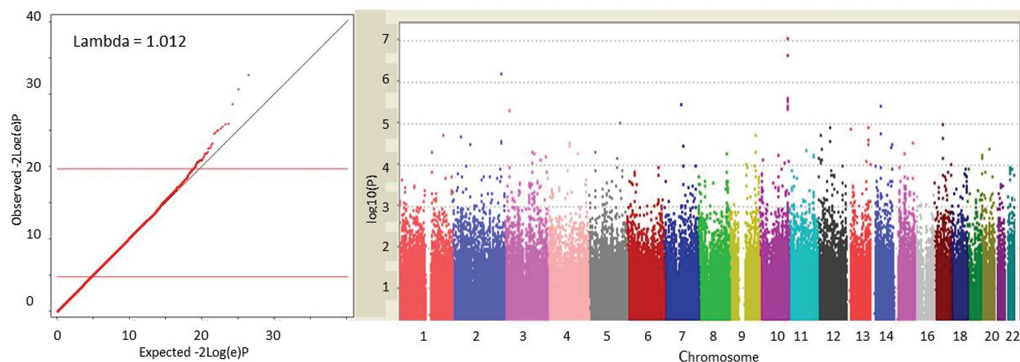


Figure 1. Quantile-quantile plot (left) and Manhattan plot (right) of GWA results for association of SNPs with AWT in the meta-analysis in NELSON.

The lowest p-value was found for rs10794108 on chromosome 10q (p-value = 8.60×10^{-8} located between the *Chromosome 10 open reading frame 90 (C10orf90*, distance 200kb) and the *Dedicated Of Cytokinesis gene* (*DOCK1*, distance 355kb). Table 2 displays the 12 top SNPs with a p-value < 10^{-5}. Table 2 in the supplement displays SNPs in association with AWT with a p-value < 10^{-4}.

Table 2. Top 12 of SNPs associated with AWT in the meta-analysis in NELSON groups I and II.

CHR	BP	SNP	Minor allele	p	B	Q
10	128413863	rs10794108	A	8.60E-08	0.051	4.12E-01
10	128409974	rs7078439	A	2.27E-07	0.046	5.33E-01
2	224269573	rs734556	G	6.23E-07	0.042	8.63E-01
10	128425326	rs10794113	A	2.39E-06	0.044	6.62E-01
10	128426429	rs4962605	A	2.49E-06	0.044	6.58E-01
10	128420360	rs10901682	A	3.02E-06	0.041	7.01E-01
7	77527824	rs10251504	G	3.36E-06	0.038	1.18E-01
14	47679625	rs1959775	C	3.78E-06	0.037	1.61E-01
10	128421433	rs4494239	G	3.90E-06	0.043	7.36E-01
10	128415036	rs11245122	G	4.44E-06	0.048	3.88E-01
3	21611547	rs1382167	C	4.82E-06	0.051	9.63E-01
5	147826008	rs3995090	C	9.40E-06	-0.037	5.00E-03

BP = Base pair; Q = p-value for heterogeneity

Based on the statistical significance of the association with AWT and presence of linkage disequilibrium (LD) between SNPs, 48 SNPs were selected for replication in groups III and IV. Out of these SNPs, one SNP (rs507098) did not pass QC. The other 47 SNPs were associated with AWT as measured in groups I and II and followed by meta-analysis. The 12 top SNPs from this analysis are shown in Table 3. (All 47 SNPs are shown in the supplement, Table 3).

The meta-analysis in groups I, II, III and IV provided 6 SNPs with a p-value < 10^{-5} including 2 SNPs with effects in the same direction in all 4 cohorts analyzed:

1. rs10251504 on chromosome 7q21, an intronic SNP in the *membrane associated guanylate kinase WW and PDZ domain containing* 2 gene (*MAGI2*), p-value 5.79×10^{-7}, b = 0.035;
2. rs4796712 on chromosome 17q21.2, an intronic SNP located in an intron in the *5′-nucleotidase, cytosolic IIIB (NT5C3B)* gene, p-value 3.11×10^{-6}, b = 0.052.

Table 3. Top 12 of 47 top SNPs associated with AWT identified in NELSON (groups I and II) followed by replication in LUSI (groups III and IV) and meta-analysis in NELSON and LUSI.

CHR	SNP	minor allele	NELSON (groups I and II)			LUSI (group III)		LUSI (group IV)		Meta-analysis in NELSON and LUSI			Direction of effect	Closest gene(s)
			MAF	p	B	p	B	p	B	p	B	Q		
7	rs10251504	G	0.435	3.36E-06	0.038	4.29E-02	0.031	5.88E-01	0.013	5.79E-07	0.035	3.24E-01	+ + + +	*MAGI2**
17	rs4796712	A	0.104	1.04E-05	0.057	2.37E-01	0.031	1.93E-01	0.057	3.11E-06	0.052	8.44E-01	+ + + +	*NT5C3B**
10	rs7078439	A	0.281	2.27E-07	0.046	9.77E-01	0.000	7.72E-01	0.007	7.70E-06	0.034	6.42E-02	+ + 0 +	*C10orf90 & DOCK1*
15	rs11070836	G	0.403	5.56E-05	0.033	8.10E-01	0.004	2.36E-03	0.070	7.81E-06	0.032	1.20E-01	+ + + +	*TNFAIP8L3*
2	rs734556	G	0.350	6.23E-07	0.042	3.22E-01	0.016	5.38E-01	-0.014	8.72E-06	0.032	9.11E-02	+ + + -	*SCG2 & AP1S3*
10	rs10794108	A	0.235	8.60E-08	0.051	7.80E-01	-0.005	9.10E-01	0.003	9.43E-06	0.035	2.03E-02	+ + - +	*C10orf90 & DOCK1*
3	rs1382167	C	0.163	4.82E-06	0.051	6.79E-01	0.010	5.55E-01	0.020	1.59E-05	0.042	4.23E-01	+ + + +	*ZNF385D**
3	rs925440	A	0.22	6.53E-05	-0.040	7.49E-02	-0.031	7.81E-01	-0.007	2.38E-05	-0.035	5.03E-01	- - - -	*LEPREL1*
12	rs1391708	G	0.102	1.30E-05	0.060	1.62E-01	0.034	6.89E-01	-0.013	4.21E-05	0.046	2.20E-01	+ + + -	*LOC100128944*
10	rs11259285	A	0.403	7.31E-05	0.033	4.19E-01	0.013	2.48E-01	0.027	4.85E-05	0.029	2.42E-01	+ + + +	*FAM107B**
14	rs2029614	A	0.102	3.09E-05	0.056	3.61E-01	0.022	9.07E-01	0.005	7.40E-05	0.045	2.38E-01	+ + + +	*RPL3P4 & BCL11B*
20	rs1291101	A	0.259	4.18E-05	-0.038	7.60E-01	-0.005	2.32E-01	-0.030	7.79E-05	-0.031	8.45E-02	- - - -	*C20orf117 & C20orf118*
10	rs11018027	A	0.104	8.03E-05	0.053	9.33E-01	0.002	5.17E-02	0.072	8.00E-05	0.045	2.42E-01	+ + + +	*TCERG1L & FLJ46300*
13	rs2065550	A	0.111	1.24E-05	0.056	9.42E-01	-0.002	4.27E-01	0.032	8.92E-05	0.043	2.20E-01	+ + - +	*LOC728183 & DAOA*

*BP = base pair; MAF = minor allele frequency in groups I and II; Q = p-value for heterogeneity; Direction of effect per cohort: each sign reflects one cohort, direction of effect is presented by: + = (B > 0), – = (B < 0), and 0 = no effect; *Corresponding SNP is located in an intron in this gene.*

Replication analyses including males from groups III and IV only (n = 457) and subsequent meta-analysis in groups I, II, III, IV showed comparable results (supplement, Table 4).

Replication analyses in individuals from groups III and IV with an FEV_1/FVC < 80% (n = 374) and subsequent meta-analysis in groups I, II, III and IV showed stronger associations between several SNPs and AWT compared to the initial analysis, counting three SNPs with genome wide significant associations, despite the smaller sample size (reducing from n = 3,354 to n = 3,014):

1) rs734556, a SNP located between the *secretogranin II* gene (*SCG2*) and the *adaptor-related protein complex 1* (sigma 3 subunit) gene (*AP1S3*) and close to the *WD repeat and FYVE domain containing 1 (WDFY1)* gene, the *mitochondrial ribosomal protein L44 (MRPL44)* gene and the *serpin peptidase inhibitor, clade E member 2 (SERPINE2)* gene on chromosome 2q (supplement, Figure 1), p-value 4.60×10^{-8}, b = 0.043;
2) rs7078439 and rs10794108, located between *C10orf90* and *DOCK1* (distance 3,889 kb, moderate linkage disequilibrium, $r^2 = 0.82$) on chromosome 10q, p-values 7.44×10^{-8}, b = 0.047 and 7.53×10^{-8}, b = 0.044 respectively.

   The top 12 SNPs from this analysis are shown in Table 4, all 47 replicated SNPs in the supplement Table 5.

Table 4. Association analyses on airway wall thickening of 12 top SNPs identified in NELSON followed by replication in LUSI (groups III and IV) with FEV_1/FVC < 80% and meta-analysis in NELSON and LUSI.

CHR	SNP	NELSON (groups I and II)		FEV1/FVC < 80% LUSI (group III)		LUSI (group IV)		Meta-analysis in NELSON and LUSI FEV_1/FVC < 80%			Direction of effect	Closest gene(s)
		p	B	p	B	p	B	p	B	Q		
2	rs734556	6.23E-07	0.042	3.29E-02	0.050	4.02E-01	0.029	4.60E-08	0.043	9.60E-01	+ + + +	SCG2 & AP1S3
10	rs7078439	2.27E-07	0.046	4.64E-01	0.018	8.57E-02	0.063	7.44E-08	0.044	6.20E-01	+ + + +	C10orf90 & DOCK1
10	rs10794108	8.60E-08	0.051	5.74E-01	0.015	2.07E-01	0.048	7.53E-08	0.047	5.05E-01	+ + + +	C10orf90 & DOCK1
17	rs4796712	1.04E-05	0.057	1.28E-01	0.058	3.21E-01	0.067	1.86E-06	0.057	9.96E-01	+ + + +	NT5C3B*
7	rs10251504	3.36E-06	0.038	2.75E-02	0.051	2.40E-01	-0.041	2.14E-06	0.036	4.85E-02	+ + + -	MAGI2*
15	rs11070836	5.56E-05	0.033	6.33E-01	0.011	6.13E-04	0.109	2.46E-06	0.036	8.09E-02	+ + + +	TNFAIP8L3 & CYP19A1
10	rs10794113	2.39E-06	0.044	8.40E-01	0.005	3.48E-01	0.036	4.19E-06	0.039	5.32E-01	+ + + +	C10orf90 & DOCK1
10	rs11245122	4.44E-06	0.048	6.29E-01	-0.015	3.65E-02	0.086	5.12E-06	0.044	1.26E-01	+ + - +	C10orf90 & DOCK1
2	rs10172774	3.02E-05	0.036	5.29E-02	0.047	7.44E-01	0.011	5.70E-06	0.036	8.52E-01	+ + + +	SCG2 & AP1S3
14	rs2029614	3.09E-05	0.056	8.10E-02	0.060	6.38E-01	0.029	6.16E-06	0.055	5.85E-01	+ + + +	RPL3P4 & BCL11B
2	rs10176854	2.78E-05	0.038	1.11E-01	0.041	6.14E-01	0.020	7.14E-06	0.037	9.05E-01	+ + + +	SCG2 & AP1S3
10	rs11259285	7.31E-05	0.033	3.03E-01	0.025	6.44E-02	0.058	9.97E-06	0.034	2.92E-01	+ + + +	FAM107B*

*Bold is decreasing p-value in the final meta-analysis compared to p-value in the identification analysis; Direction of effect per cohort: each sign reflects one cohort, direction of effect is presented by: + = (B > 0), – = (B < 0); *Corresponding SNP is located in an intron in this gene.*

Functional analyses on SNPs and corresponding genes identified in the initial analysis

We found a strong significant association between rs4796712 on the lung mRNA expression levels of *NT5C3B* (Affymetrix ID: 100128528-TGI-at, Ensemble ID: NM_052935). The (susceptibility) T allele associated significantly with a higher *NT5C3B* expression (genotypes: CC = 910, TC = 172, TT = 13, p = 1.09 x 10^{-41}, b = -0.910). There was no significant association between rs10251504 and *MAGI2* expression (Figure 2).

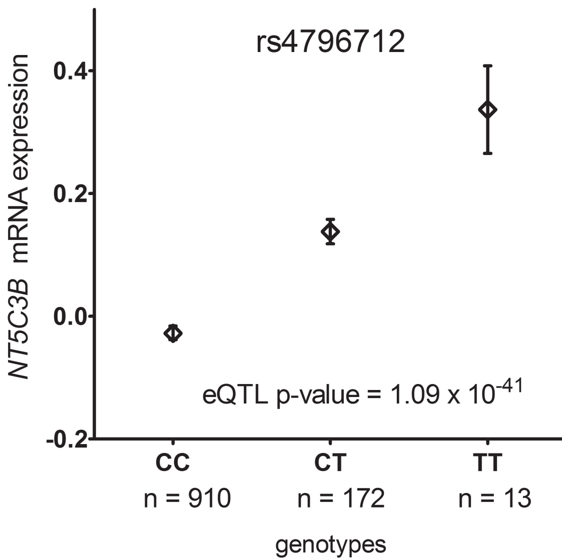


Figure 2. Boxplots of lung gene expression levels for NT5C3B according to genotype groups for SNP rs4868687 in 1,095 individuals.

Subsequently, *MAGI2 and NT5C3B* expression were assessed in airway wall biopsies of COPD patients who did not use inhaled corticosteroids in GLUCOLD. mRNA expression levels were correlated with lung function and bronchial biopsy inflammatory markers [10, 11]. Higher *MAGI2* mRNA expression was significantly associated with lower numbers of inflammatory markers: macrophages (p = 1.08 x 10^{-2}), CD3 lymphocytes (p = 1.57 x 10^{-2}), CD4 lymphocytes (p = 4.08 x 10^{-4}), CD8 lymphocytes (p = 4.22 x 10^{-3}) and % intact epithelial cells (p = 1.24 x 10^{-3}) but not with eosinophils (p = 0.72) or mast cells (p = 9.7 x 10^{-2}). *MAGI2* mRNA expression was not associated with the level of lung function (post bronchodilator FEV_1% predicted, p = 0.42). Lower *NT5C3B* mRNA expression was significantly associated with lower lung function (post bronchodilator FEV_1% predicted; p = 3.17 x 10^{-2}) but not with any of the inflammatory markers (macrophages, mast cells, CD3-, CD4-, CD8- lymphocytes or % intact epithelial cells, p-values 0.14, 0.33, 0.66, 0.66, 0.83, 0.76 respectively).

Functional analyses on SNPs and corresponding genes identified in individuals with comparable lung function

We found a significant association between rs734556 and *SERPINE2* expression (Affymetrix ID: 100307061_TGI_at, Ensemble ID: BQ876560). The (susceptibility) T allele was significantly associated with lower *SERPINE2* expression (genotypes: TT = 489, TG = 463, GG = 143; p = 3.21 x 10^{-4}, b = 0.153) (Figure 3). We did not find significant associations of rs10794108 and rs7078439 genotypes with *C10orf90* and *DOCK1* expression levels in lung tissue, nor of rs734556 genotypes with *SCG2*, *AP1S3* and *WDFY1* expression levels.

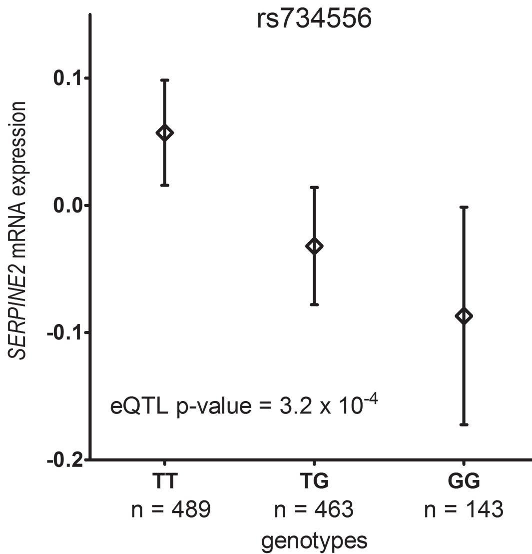


Figure 3. Boxplots of lung gene expression levels for *SERPINE2* according to genotype groups for SNP rs4868687 in 1,095 individuals.

Discussion

To our knowledge, this is the first genome wide association study for CT-quantified airway wall thickness. We provide evidence for genetic origins of AWT, an important contributing factor for airway obstruction and development of COPD. We identified two SNPs (rs10251504 and rs4796712) associated with AWT showing effects in the same direction in both the identification and replication cohort. Moreover, when selecting individuals in the replication cohort (LUSI) with comparable level of lung function as those in the identification cohort (NELSON), three SNPs (rs734556, rs7078439 and rs10794108) reached genome wide significance in the meta-analysis in the cohorts studied.

In the identification analysis we discovered two SNPs, rs10794108 and rs7078439 located near each other in a "desert" between 2 genes (*C10orf90 & DOCK1*) strongly associated with AWT. The SNP rs10794108 was found previously to be associated with severity of airflow obstruction (FEV_1/FVC < 90% predicted and FEV_1 < 80% predicted) in a GWAs performed in the Framingham Heart Study [13]. When replicating this SNP (rs10794108) in a cohort with comparable lung function this association was confirmed. We have previously shown a significant relation between airflow obstruction and airway wall thickening, both known features of COPD, in the NELSON cohort [4]. The current finding that rs10794108 is associated with airflow obstruction as well as with AWT supports our earlier finding.
The SNP rs734556 was not only significantly associated with AWT but the risk allele (G) also was significantly associated with lower expression of *SERPINE2* in lung tissue. This is of interest, since *SERPINE2* was identified previously as a susceptibility gene for COPD and particularly emphysema [3, 14, 15]. SERPINE2 has been shown to inhibit extracellular matrix destruction [16]. SNPs in this gene may influence alterations in repair of smoking-induced airway wall damage and our data suggest that this SNP may be involved in a common pathway of the origin of emphysema and AWT.

Another SNP showing a strong primary association with AWT was located in the *guanylate kinase WW and PDZ domain containing 2 gene* (*MAGI2*), a large gene that encodes a scaffolding protein involved in the epithelial tight junction pathway [17]. Cell membranes of epithelial cells join together forming a virtually impermeable barrier. Tight junctions are the most apically located of the intercellular junctions and play a critical role in epithelial barrier function. Therefore, variants in tight junction genes may affect this barrier function in the airways. We found higher expression of *MAGI2* associated

with lower numbers of inflammatory markers in bronchial biopsies of COPD patients. It could be speculated that the protective function of the bronchial epithelial layer is weakened by decreased expression of *MAGI2* thereby allowing inhaled particles like cigarette smoke to penetrate more easily in the underlying tissue, causing inflammation and increased levels of markers associated with inflammation. This inflammation may subsequently lead to remodeling in the respiratory tract and thickening of the airways, particularly when this process takes place in the smaller airways, as present e.g. in COPD. Interestingly, SNPs in *MAGI2* are also associated with inflammatory bowel disease supporting the hypothesis that diseases in which the integrity of the epithelium is affected share common underlying genetics [18, 19].

The SNP, rs4796712 in *NT5C3B* was second in rank of significance encoding the enzyme cytosolic 5′-nucleotidase 3. The functional studies showed that *NT5C3B* is involved in airway wall thickening and thereby with airway obstruction as 1) rs4796712 has a significant eQTL effect on *NT5C3B* lung tissue expression, the risk allele (T) was associated with significantly higher expression of *NT5C3B* while 2) lower *NT5C3B* mRNA expression is associated with worse lung function in COPD patients and 3) the association of rs4796712 with AWT is stronger when specifically individuals with comparable worse lung function are included. *NT5C3B* encodes a 5′-nucleotidase, a member of the 5′-nucleotidase family of enzymes that catalyze the de-phosphorylation of nucleoside monophosphates. These enzymes are involved in various functions, such as cell-cell communication, nucleic acid repair, purine salvage pathway (for the synthesis of nucleotides), signal transduction and membrane transport [20].

In summary, our study provides the first evidence that airway wall thickness in airways of 3.5 mm internal diameter is associated with a genetic predisposition. When analyzing cohorts with a comparable level of airway obstruction a SNP (rs10794108) that was previously found to be associated with the level of FEV_1 was also strongly associated with AWT. Another SNP (rs734556) that strongly associated with AWT was an eQTL that decreased lung tissue expression of *SERPINE2*, a gene previously identified as a susceptibility gene for COPD, particularly emphysema. A top SNP in *MAGI2*, a gene involved in epithelial integrity, was replicated and higher *MAGI2* expression was associated with lower numbers of bronchial inflammation markers. A top SNP in *NT5C3B*, was nominally replicated and an eQTL in lung tissue, the risk allele being associated with higher expression. Furthermore, lower *NT5C3B* mRNA expression was significantly associated with worse lung function. Thus, our data support the notion that AWT is associated with genes involved in emphysema, bronchial inflammation, and lung function.

References

1 Nakano Y, Wong JC, de Jong PA et al. The prediction of small airway dimensions using computed tomography. *American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association*. 2005;171(1073-449; 2):142-146

2 Patel BD, Coxson HO, Pillai SG et al. Airway wall thickening and emphysema show independent familial aggregation in chronic obstructive pulmonary disease. *American Journal of Respiratory & Critical Care Medicine*. 2008;178(5):500-505

3 Kim WJ, Hoffman E, Reilly J et al. Association of COPD candidate genes with computed tomography emphysema and airway phenotypes in severe COPD. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*. 2011;37(1):39-43

4 Dijkstra AE, Postma DS, Ten Hacken N et al. Low-dose CT measurements of airway dimensions and emphysema associated with airflow limitation in heavy smokers: a cross sectional study. *Respiratory research*. 2013;14(1):11

5 van Klaveren RJ, Oudkerk M, Prokop M et al. Management of lung nodules detected by volume CT scanning. *The New England journal of medicine*. 2009;361(23):2221-2229

6 Becker N, Motsch E, Gross ML et al. Randomized study on early detection of lung cancer with MSCT in Germany: study design and results of the first screening round. *Journal of cancer research and clinical oncology*. 2012;138(9):1475-1486

7 Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R and Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *The European respiratory journal.Supplement*. 1993;16:5-40

8 Gietema HA, Schilham AM, van Ginneken B, van Klaveren RJ, Lammers JW and Prokop M. Monitoring of smoking-induced emphysema with CT in a lung cancer screening setting: detection of real increase in extent of emphysema. *Radiology*. 2007;244(0033-8419; 0033-8419; 3):890-897

9 Hao K, Bosse Y, Nickle DC et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS genetics*. 2012;8(11):e1003029

10 Lapperre TS, Snoeck-Stroband JB, Gosman MM et al. Effect of fluticasone with and without salmeterol on pulmonary outcomes in chronic obstructive pulmonary disease: a randomized trial. *Annals of Internal Medicine*. 2009;151(8):517-527

11 van den Berge M, Steiling K, Timens W et al. Airway gene expression in COPD is dynamic with inhaled corticosteroid treatment and reflects biological pathways associated with disease activity. *Thorax*. 2014;69(1):14-23

12 Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*. 2007;81(0002-9297; 0002-9297; 3):559-575

13 Wilk JB, Chen TH, Gottlieb DJ et al. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. *PLoS genetics*. 2009;5(3):e1000429

14 DeMeo D, Mariani T, Lange C et al. The SERPINE2 gene is associated with chronic obstructive pulmonary disease. *Proceedings of the American Thoracic Society*. 2006;3(6):502

15 Kukkonen MK, Tiili E, Hamalainen S et al. SERPINE2 haplotype as a risk factor for panlobular type of emphysema. *BMC medical genetics*. 2011;12:157-2350-12-157

16 Bergman BL, Scott RW, Bajpai A, Watts S and Baker JB. Inhibition of tumor-cell-mediated extracellular matrix destruction by a fibroblast proteinase inhibitor, protease nexin I. *Proceedings of the National Academy of Sciences of the United States of America*. 1986;83(4):996-1000

17 Wu X, Hepner K, Castelino-Prabhu S et al. Evidence for regulation of the PTEN tumor suppressor by a membrane-localized multi-PDZ domain containing scaffold protein MAGI-2. *Proceedings of the National Academy of Sciences of the United States of America*. 2000;97(8):4233-4238

18 Wapenaar MC, Monsuur AJ, van Bodegraven AA et al. Associations with tight junction genes PARD3 and MAGI2 in Dutch patients point to a common barrier defect for coeliac disease and ulcerative colitis. *Gut*. 2008;57(4):463-467

19 McGovern DP, Taylor KD, Landers C et al. MAGI2 genetic variation and inflammatory bowel disease. *Inflammatory bowel diseases*. 2009;15(1):75-83

20 Hunsucker SA, Mitchell BS and Spychala J. The 5′-nucleotidases as regulators of nucleotide and drug metabolism. *Pharmacology & therapeutics*. 2005;107(1):1-30

Chapter 7

Supplement

MeVis software

This software automatically extracts airway centerlines, re-samples images perpendicular to the airway direction at equally spaced positions along the centerline and detects inner and outer airway wall borders in these images. The outer wall border is detectable when no adjacent tissue with similar CT density is present and is taken into account when the wall is detected in at least 25% of the perimeter at a location. AWT and the fraction of perimeter where the outer wall border was identified (Assessed Perimeter Fraction, APF) are calculated for each location. Airway wall thickness quantification accounts for partial volume effects by integrating Hounsfield units across the wall.

The lung eQTL study

Non-tumor lung tissues were collected from patients who underwent lung resection surgery at three participating sites: Laval University (Quebec City, Canada), University of Groningen (Groningen, the Netherlands), and University of British Columbia (Vancouver, Canada). Whole-genome gene expression and genotyping data were obtained from these specimens. Gene expression profiling was performed using an Affymetrix custom array testing 51,627 non-control probe sets and normalized using RMA [1]. Genotyping was performed using the Illumina Human1M-Duo BeadChip array. At Laval university, lung specimens were collected from patients undergoing lung cancer surgery and stored at the "Institut universitaire de cardiologie et de pneumologie de Québec" (IUCPQ) site of the Respiratory Health Network Tissue Bank of the "Fonds de recherche du Québec - Santé" (www.tissuebank.ca). Written informed consent was obtained from all subjects and the study was approved by the IUCPQ ethics committee. At Groningen University, lung specimens were provided by the local tissue bank of the Department of Pathology and the study protocol was according to the Research Code of the University Medical Center Groningen and Dutch national ethical and professional guidelines ("Code of conduct; Dutch federation of biomedical scientific societies"; http://www.federa.org). At the University of British Columbia, the lung specimens were provided by the James Hogg Research Center Biobank at St Paul's Hospital and subjects provided written informed consent. The study was approved by the ethics committees at the UBC-Providence Health Care Research Institute Ethics Board. The lung eQTL analysis was performed as described by Fehrmann and Hao [2, 3].

Gene expression analysis in GLUCOLD 4

RNA isolation and size fractionation

Out of 114 COPD subjects in GLUCOLD, 89 individuals had endobronchial biopsies from normal airway wall tissue which had been immediately snap-frozen and stored at -80°C. RNA was extracted and fractionated into low (<200 nt) and high molecular weight (>200 nt) fractions using the miRNeasy mini kit (QIAGEN) according to manufacturer's protocol. The purity of RNA fractions was assessed using a NanoDrop 1000 UV-Vis spectrophotometer. The integrity of the large RNA fraction was assessed by using the RNA Pico assay in the Agilent 2100 BioAnalyzer.

RNA processing and microarray hybridization

All procedures were performed at Boston University Microarray Resource Facility as described in GeneChip® Whole Transcript (WT) Sense Target Labeling Assay Manual (Affymetrix, Santa Clara, CA, current version available at www.affymetrix.com). The Qiagen miRNeasy Mini Kit and RNeasy MinElute Cleanup Kit were used to isolate high and low molecular weight RNA. 200 ng of high molecular weight large RNA was reverse transcribed using the Whole Transcript cDNA Synthesis kit (Affymetrix, Santa Clara, CA). The obtained cDNA was used as a template for in vitro transcription using the Whole Transcript cDNA Amplification Kit (Affymetrix, Santa Clara, CA). The obtained antisense cRNA was purified using GeneChip Sample Cleanup Module (Affymetrix, Santa Clara, CA), and used as a template for reverse transcription (Whole Transcript cDNA Synthesis kit, Affymetrix, Santa Clara, CA) to produce single-stranded DNA in the sense orientation. During this step, dUTP was incorporated. The DNA was then fragmented using uracil DNA glycosylase (UDG) and apurinic/apyrimidinic endonuclease 1 (APE 1), and labeled with DNA Labeling Reagent that was covalently linked to biotin using terminal deoxynucleotidyl transferase (TdT, Whole Transcript Terminal Labeling kit, Affymetrix, Santa Clara, CA). IVT and cDNA fragmentation quality controls were carried out by running an mRNA Nano assay in the Agilent 2100 Bioanalyzer. The labeled fragmented DNA was hybridized to the Affymetrix Human Gene 1.0 ST Arrays for 16-18 hours in GeneChip Hybridization oven 640 at 45°C using 60 rpm rotation. The hybridized samples were washed and stained using Affymetrix fluidics station 450. The first stain with streptavidin-R-phycoerythrin (SAPE) was followed by signal amplification using a biotinilated goat anti-streptavidin antibody and another SAPE staining (Hybridization, Washing and Staining Kit, Affymetrix, Santa Clara, CA). Microarrays were immediately scanned using Affymetrix GeneArray Scanner 3000 7G Plus (Affymetrix, Santa Clara, CA).

Data acquisition, probeset summarization and normalization, and data preprocessing. Normalization was performed with Affymetrix Expression Console software using Affymetrix default Robust Multichip Analysis (RMA) sketch algorithm workflow on all samples except one due to low quality of the microarray data.
Microarray data quality was assessed using relative log expression (RLE) plots, normalized unscaled standard error (NUSE) plots, and principle component analysis (PCA). Based on the RLE and NUSE plots, a total of 9 microarrays were excluded, leaving 79 microarrays for subsequent analysis.
Association of the number of bronchial inflammatory cells with *MAGI2* and *NT5C3B* and the association of lung function outcome with *MAGI2* and *NT5C3B* mRNA-expression levels was analyzed with linear regression including adjustment for current smoking and RNA integrity score.

Table 1. Demographic and clinical characteristics of groups III and IV LUSI replication populations. 1) groups III and IV including only males; 2) groups III and IV including participants with FEV_1/FVC < 80%

	1) males (n = 457)		2) FEV_1/FVC < 80% (n = 374)	
	III	IV	III	IV
N	322	135	253	121
Characteristics				
Age, years	57.4 (5.4)	53.3 (4.6)	57.7 (5.4)	53.5 (4.7)
Height, cm	177.6 (6.2)	177.7 (7.1)	173.2 (8.4)	173.9 (9.1)
Male gender, %	100	100	64.6	63.6
Smoking				
Pack-years smoking	37 (19-116)	37 (19-104)	37.0 (19-116)	33.0 (19-104)
Current smoking, %	57.8	64.4	63.0	69.4
Lung function				
FEV1 , liter	3.17 (0.73)	3.29 (0.69)	2.7 (0.73)	2.8 (0.75)
FEV1/FVC, %	78.4 (10.0)	78.3 (11.6)	71.8 (7.6)	70.8 (8.4)
FEV_1, % predicted	91.1 (18.7)	91.0 (16.5)	86.4 (18.4)	84.9 (16.9)
CT measurements				
Median AWT, mm	0.58 (0.39-1.36)	0.62 (0.34-1.20)	0.60 (0.32-1.36)	0.67 (0.38-1.20)
Lung vol. on CT, liter	6.02 (1.35)	6.35 (1.29)	5.62 (1.3)	6.0 (1.5)

Group III: LUSI-Toshiba population; Group IV: LUSI-Siemens population; Mean (± standard deviation) shown for continuous data and median (range) for non-parametric distribution; AWT = airway wall thickness at 3.5 mm internal lumen size.

Table 2. SNPs associated with AWT (p < 10^{-4}) in the meta-analysis in NELSON groups I and II.

CHR	BP	SNP	Minor allele	p fixed	p random	B fixed	B random	Q
10	128413863	rs10794108	A	8.60E-08	8.60E-08	0.051	0.051	4.12E-01
10	128409974	rs7078439	A	2.27E-07	2.27E-07	0.046	0.046	5.33E-01
2	224269573	rs734556	G	6.23E-07	6.23E-07	0.042	0.042	8.63E-01
10	128425326	rs10794113	A	2.39E-06	2.39E-06	0.044	0.044	6.62E-01
10	128426429	rs4962605	A	2.49E-06	2.49E-06	0.044	0.044	6.58E-01
10	128420360	rs10901682	A	3.02E-06	3.02E-06	0.041	0.041	7.01E-01
7	77527824	rs10251504	G	3.36E-06	2.58E-03	0.038	0.039	1.18E-01
14	47679625	rs1959775	C	3.78E-06	8.91E-04	0.037	0.038	1.61E-01
10	128421433	rs4494239	G	3.90E-06	3.90E-06	0.043	0.043	7.36E-01
10	128415036	rs11245122	G	4.44E-06	4.44E-06	0.048	0.048	3.88E-01
3	21611547	rs1382167	C	4.82E-06	4.82E-06	0.051	0.051	9.63E-01
5	147826008	rs3995090	C	9.40E-06	1.34E-01	-0.037	-0.035	5.00E-03
17	37240656	rs4796712	A	1.04E-05	1.04E-05	0.057	0.057	8.53E-01
13	102902377	rs2065550	A	1.24E-05	1.24E-05	0.056	0.056	8.73E-01
12	55591847	rs1391708	G	1.30E-05	1.30E-05	0.060	0.060	8.02E-01
13	22650396	rs574662	C	1.33E-05	2.23E-02	-0.038	-0.039	5.10E-02
14	47653933	rs1992691	G	1.74E-05	1.74E-05	0.035	0.035	3.34E-01
12	13852010	rs220583	A	1.85E-05	7.72E-02	0.040	0.038	2.20E-02
9	118828951	rs7047287	A	1.87E-05	1.87E-05	0.040	0.040	9.55E-01
1	203875633	rs12039255	A	1.90E-05	1.90E-05	0.073	0.073	9.10E-01
2	40544972	rs11679585	A	2.06E-05	2.06E-05	-0.044	-0.044	8.68E-01
17	37341634	rs9303323	A	2.29E-05	2.29E-05	0.057	0.057	8.17E-01
13	101633794	rs554393	G	2.51E-05	9.67E-05	0.037	0.037	2.79E-01
2	224284646	rs10176854	G	2.78E-05	2.78E-05	0.038	0.038	5.55E-01
12	55543511	rs11832720	A	2.79E-05	2.79E-05	0.061	0.061	7.66E-01
15	87660998	rs3087374	A	3.01E-05	5.96E-04	-0.068	-0.068	2.23E-01
2	224249717	rs10172774	A	3.02E-05	3.02E-05	0.036	0.036	7.84E-01
4	101918405	rs2089540	G	3.03E-05	3.03E-05	0.038	0.038	4.44E-01
13	102898787	rs10508110	G	3.07E-05	3.07E-05	0.053	0.053	8.94E-01
14	98617020	rs2029614	A	3.09E-05	1.71E-03	0.056	0.056	1.89E-01
4	101921280	rs4571333	A	3.25E-05	3.25E-05	0.038	0.036	4.66E-01
2	80581631	rs7576475	G	3.28E-05	3.28E-05	-0.037	-0.037	5.54E-01
4	101903022	rs1508421	A	3.48E-05	3.48E-05	0.038	0.038	4.53E-01
4	101913861	rs12503265	A	3.48E-05	3.48E-05	0.038	0.038	4.53E-01
7	88935516	rs10232434	C	3.49E-05	3.49E-05	-0.038	-0.038	5.36E-01
14	92420129	rs4905002	A	3.73E-05	6.18E-03	0.036	0.036	1.27E-01
12	20385638	rs7303397	G	3.97E-05	3.97E-05	0.035	0.035	3.32E-01
20	34927500	rs1291101	A	4.18E-05	3.20E-02	-0.038	-0.040	4.80E-02
11	81350924	rs12794636	A	4.44E-05	4.44E-05	0.067	0.067	9.90E-01
1	152408532	rs4845364	G	4.86E-05	4.86E-05	-0.034	-0.034	6.90E-01
5	31806568	rs919336	A	4.91E-05	4.91E-05	-0.034	-0.034	9.39E-01
3	129496335	rs2811518	G	4.97E-05	4.97E-05	-0.049	-0.049	7.36E-01
9	118822192	rs1334090	G	5.10E-05	5.10E-05	0.038	0.038	9.60E-01
3	135980476	rs10154906	C	5.22E-05	5.22E-05	-0.040	-0.040	3.76E-01
4	138143227	rs679959	A	5.37E-05	1.34E-02	0.035	0.035	1.07E-01
8	131791979	rs1396976	A	5.56E-05	2.32E-03	0.043	0.044	1.81E-01
15	49223689	rs11070836	G	5.56E-05	5.56E-05	0.033	0.033	8.26E-01
20	2497008	rs4815352	G	5.61E-05	5.61E-05	0.033	0.033	4.13E-01
10	85538632	rs509948	A	5.76E-05	5.76E-05	-0.033	-0.033	7.99E-01
11	113947313	rs550897	G	5.95E-05	6.57E-03	0.033	0.033	1.34E-01
11	113928530	rs531824	G	6.14E-05	6.55E-03	0.033	0.033	1.34E-01
3	191325669	rs925440	A	6.53E-05	6.83E-05	-0.040	-0.039	3.16E-01
5	132431332	rs4705990	G	6.62E-05	6.62E-05	0.037	0.037	8.66E-01
20	2498579	rs767739	A	6.70E-05	5.57E-02	0.033	0.034	3.30E-02
17	37398449	rs12952314	A	6.82E-05	6.82E-05	0.053	0.053	4.81E-01
12	3534965	rs10774156	A	6.90E-05	6.90E-05	0.037	0.037	7.60E-01
5	132463471	rs4367292	A	7.02E-05	7.02E-05	0.037	0.037	8.56E-01
10	14782881	rs11259285	A	7.31E-05	1.82E-02	0.033	0.034	8.50E-02
12	20402460	rs2009084	G	7.35E-05	1.34E-03	0.033	0.033	2.20E-01
12	13880137	rs2284424	A	7.49E-05	4.74E-02	0.035	0.034	7.00E-02
3	168672301	rs17246389	C	7.60E-05	7.60E-05	0.037	0.037	6.67E-01
10	133229112	rs11018027	A	8.03E-05	8.03E-05	0.053	0.053	5.47E-01
13	93841289	rs1407999	C	8.22E-05	6.73E-03	0.032	0.032	1.53E-01
3	129438962	rs2999081	A	8.63E-05	8.63E-05	-0.047	-0.047	6.95E-01
10	116052500	rs507098	G	9.00E-05	9.00E-05	0.050	0.050	3.78E-01
12	13880286	rs2284425	A	9.03E-05	4.48E-02	0.035	0.033	6.20E-02
9	75056788	rs10429583	G	9.33E-05	9.33E-05	-0.034	-0.034	9.47E-01
13	112791769	rs534298	A	9.34E-05	3.54E-03	-0.039	-0.039	1.77E-01

Q = p-value for heterogeneity

Table 3. Association analysis on airway wall thickening of 47 top-SNPs (identified in NELSON) followed by replication in LUSI and subsequent meta-analysis in NELSON and LUSI.

CHR	BP	BNP	minor allele	MAF	NELSON		LUSI group III		LUSI group IV		Meta-analysis			Direction of effect	Closest gene(s)
					p	B	p	B	p	B	p	B	Q		
7	rs10251504	77527824	G	0.435	3.36E-06	0.038	4.29E-02	0.031	5.88E-01	0.013	5.79E-07	0.035	3.24E-01	+ + + +	*MAGI2**
17	rs4796712	37240656	A	0.104	1.04E-05	0.057	2.37E-01	0.031	1.93E-01	0.057	3.11E-06	0.052	8.44E-01	+ + + +	*NT5C3B**
10	rs7078439	128409974	A	0.281	2.27E-07	0.046	9.77E-01	0.000	7.72E-01	0.007	7.70E-06	0.034	6.42E-02	+ + 0 +	*C10orf90 & DOCK1*
15	rs11070836	49223689	G	0.403	5.56E-05	0.033	8.10E-01	0.004	2.36E-03	0.070	7.81E-06	0.032	1.20E-01	+ + + +	*TNFAIP8L3*
2	rs734556	224269573	G	0.350	6.23E-07	0.042	3.22E-01	0.016	5.38E-01	-0.014	8.72E-06	0.032	9.11E-02	+ + + -	*SCG2 & AP1S3*
10	rs10794108	128413863	A	0.235	8.60E-08	0.051	7.80E-01	-0.005	9.10E-01	0.003	9.43E-06	0.035	2.03E-02	+ + - +	*C10orf90 & DOCK1*
3	rs1382167	21611547	C	0.163	4.82E-06	0.051	6.79E-01	0.010	5.55E-01	0.020	1.59E-05	0.042	4.23E-01	+ + + +	*ZNF385D**
3	rs925440	191325669	A	0.22	6.53E-05	-0.040	7.49E-02	-0.031	7.81E-01	-0.007	2.38E-05	-0.035	5.03E-01	- - - -	*LEPREL1*
12	rs1391708	55591847	G	0.102	1.30E-05	0.060	1.62E-01	0.034	6.89E-01	-0.013	4.21E-05	0.046	2.20E-01	+ + + -	*LOC100128944*
10	rs11259285	14782881	A	0.403	7.31E-05	0.033	4.19E-01	0.013	2.48E-01	0.027	4.85E-05	0.029	2.42E-01	+ + + +	*FAM107B**
14	rs2029614	98617020	A	0.102	3.09E-05	0.056	3.61E-01	0.022	9.07E-01	0.005	7.40E-05	0.045	2.38E-01	+ + + +	*RPL3P4 & BCL11B*
20	rs1291101	34927500	A	0.259	4.18E-05	-0.038	7.60E-01	-0.005	2.32E-01	-0.030	7.79E-05	-0.031	8.45E-02	- - - -	*C20orf117 & C20orf118*
10	rs11018027	133229112	A	0.104	8.03E-05	0.053	9.33E-01	0.002	5.17E-02	0.072	8.00E-05	0.045	2.42E-01	+ + + +	*TCERG1L & FLJ46300*
13	rs2065550	102902377	A	0.111	1.24E-05	0.056	9.42E-01	-0.002	4.27E-01	0.032	8.92E-05	0.043	2.20E-01	+ + - +	*LOC728183 & DAOA*
14	rs1959775	47679625	C	0.441	3.78E-06	0.037	8.73E-01	0.003	6.52E-01	-0.010	1.03E-04	0.027	3.68E-02	+ + + -	*RPS15AP3 & RPL18P1*
1	rs12039255	203875633	A	0.061	1.90E-05	0.073	9.45E-02	0.049	1.26E-01	-0.069	1.27E-04	0.054	3.23E-02	+ + + -	*ELK4 & SLC45A3*
2	rs10172774	224249717	A	0.340	3.02E-05	0.036	2.86E-01	0.017	4.96E-01	-0.016	1.64E-04	0.027	1.80E-01	+ + + -	*SCG2 & AP1S3*
2	rs7576475	80581631	G	0.312	3.28E-05	-0.037	3.96E-01	0.015	4.95E-02	-0.055	1.72E-04	-0.029	4.39E-02	- - + -	*CTN2**
12	rs10774156	3534965	A	0.287	6.90E-05	0.037	4.96E-01	0.012	8.90E-01	0.004	1.96E-04	0.029	4.67E-01	+ + + +	*PRMT8**
10	rs10794113	128425326	A	0.252	2.39E-06	0.044	6.23E-01	-0.009	8.77E-01	-0.004	2.09E-04	0.029	2.98E-02	+ + - -	*C10orf90 & DOCK1*
2	rs10176854	224284646	G	0.284	2.78E-05	0.038	2.24E-01	0.021	1.64E-01	-0.036	2.56E-04	0.028	4.65E-02	+ + + -	*SCG2 & AP1S3*
13	rs1407999	93841289	C	0.468	8.22E-05	0.032	2.68E-01	0.017	5.77E-01	-0.013	2.73E-04	0.025	1.20E-01	+ + + -	*GPC6**
11	rs550897	113947313	G	0.450	5.95E-05	0.033	1.32E-01	0.023	1.42E-01	-0.035	2.74E-04	0.025	2.25E-02	+ + + -	*FAM55D**
10	rs11245122	128415036	G	0.179	4.44E-06	0.048	2.14E-01	-0.025	3.50E-01	0.027	2.92E-04	0.032	1.11E-02	+ + - +	*C10orf90 & DOCK1*
2	rs11679585	40544972	A	0.184	2.06E-05	-0.044	6.94E-01	-0.007	5.82E-01	0.017	3.16E-04	-0.032	1.30E-01	- - - +	*SLC8A1 & LOC729984*

CHR	BP	BNP	minor allele	MAF	NELSON		LUSI group III		LUSI group IV		Meta-analysis			Direction of effect	Closest gene(s)
					p	B	p	B	p	B	p	B	Q		
12	rs2284424	13880137	A	0.304	7.49E-05	0.035	3.12E-01	0.017	4.20E-01	-0.021	3.74E-04	0.027	3.94E-02	+ + + -	*GRIN2B**
3	rs2811518	129496335	G	0.132	4.97E-05	-0.049	8.84E-01	-0.003	9.49E-01	-0.002	4.11E-04	-0.036	2.49E-01	- - - -	*EEFSEC**
3	rs17246389	168672301	C	0.248	7.60E-05	0.037	6.20E-01	-0.009	1.58E-01	0.037	4.29E-04	0.028	1.28E-01	+ + - +	*SERPINI2**
13	rs574662	22650396	C	0.309	1.33E-05	-0.038	8.83E-01	0.002	7.22E-01	0.009	4.56E-04	-0.026	1.38E-02	- - + +	*LOC100130029 & SGCG*
14	rs4905002	92420129	A	0.338	3.73E-05	0.036	9.03E-01	-0.002	8.50E-01	0.005	4.83E-04	0.025	6.57E-02	+ + - +	*GOLGA5 & CHGA*
17	rs2362396	75469049	G	0.281	1.01E-04	0.035	2.95E-01	-0.018	2.94E-02	0.059	5.34E-04	0.027	2.85E-02	+ + - +	*CBX4 &TBC1D16*
5	rs919336	31806568	A	0.430	4.91E-05	-0.034	7.54E-01	0.005	5.89E-01	-0.013	5.70E-04	-0.024	1.71E-01	- - + -	*C5orf22 & PDZD2*
13	rs534298	112791769	A	0.214	9.34E-05	-0.039	7.73E-01	-0.005	9.44E-01	0.002	8.14E-04	-0.027	1.13E-01	- - 0 0	*MCF2L**
8	rs1396976	131791979	A	0.185	5.56E-05	0.043	7.60E-01	0.006	4.54E-01	-0.022	8.86E-04	0.030	5.40E-02	+ + + -	*ASAP1 & ADCY8*
9	rs10429583	75056788	G	0.324	9.33E-05	-0.034	7.30E-01	-0.006	4.84E-01	0.017	1.26E-03	-0.024	1.40E-01	- - - +	*ANXA1 &LOC138971*
9	rs7047287	118828951	A	0.243	1.87E-05	0.040	6.11E-01	-0.009	5.98E-01	-0.015	1.29E-03	0.026	3.67E-02	+ + - -	*ASTN2**
10	rs509948	85538632	A	0.491	5.76E-05	-0.033	4.20E-01	0.013	5.92E-01	-0.012	1.35E-03	-0.023	7.47E-02	- - + -	*LOC728050 & GHITM*
5	rs4705990	132431332	G	0.271	6.62E-05	0.037	4.57E-01	-0.013	6.86E-01	0.011	1.64E-03	0.025	8.02E-02	+ + - +	*HSPA4**
13	rs554393	101633794	G	0.312	2.51E-05	0.037	6.49E-01	-0.008	2.24E-01	-0.030	2.58E-03	0.023	9.10E-03	+ + - -	*FGF14**
20	rs4815352	2497008	G	0.469	5.61E-05	0.033	1.17E-01	-0.025	4.22E-01	0.019	2.74E-03	0.021	1.15E-02	+ + - +	*TMC2**
3	rs10154906	135980476	C	0.212	5.22E-05	-0.040	3.86E-01	0.016	5.17E-01	0.018	4.54E-03	-0.024	1.64E-02	- - + +	*KY & EPHB1*
1	rs4845364	152408532	G	0.48	4.86E-05	-0.034	2.02E-01	0.020	9.42E-01	0.002	4.61E-03	-0.020	1.64E-02	- - + +	*TPM3**
4	rs679959	138143227	A	0.321	5.37E-05	0.035	3.81E-01	-0.015	2.81E-01	-0.029	5.77E-03	0.021	4.20E-03	+ + - -	*LOC646316 & LOC729307*
12	rs7303397	20385638	G	0.364	3.97E-05	0.035	2.27E-01	-0.021	2.64E-01	-0.027	7.14E-03	0.020	4.00E-03	+ + - -	*LOC644976 & LOC100131677*
7	rs10232434	88935516	C	0.267	3.49E-05	-0.038	4.76E-02	0.035	5.05E-01	0.018	1.43E-02	-0.019	1.20E-03	- - + +	*ZNF804B & DPY19L2P4*
4	rs2089540	101918405	G	0.278	3.03E-05	0.038	1.27E-01	-0.027	1.09E-01	-0.042	1.71E-02	0.018	6.00E-04	+ + - -	*EMCN & LOC728771*

*Group I: NELSON Groningen population; Group II: NELSON Utrecht population; Group III: LUSI-Toshiba population; Group IV: LUSI-Siemens population; MAF = minor allele frequency in group I and II; Q = p-value for heterogeneity; Direction of effect per cohort: each sign reflects one cohort, direction of effect is presented by: + = (B > 0), – = (B < 0), and 0 = no effect; *Corresponding SNP is located in an intron in this gene.*

Table 4. Association analysis on airway wall thickening of 47 SNPs (identified in NELSON), followed by replication in LUSI group III-males and group IV-males, and meta-analysis in NELSON and LUSI-males.

CHR	SNP	BP	NELSON		LUSI (group III) males		LUSI (group IV) males		Meta-analysis			Direction of effect	Close(st) gene(s)
			p	B	p	B	p	B	p	B	Q		
12	rs1391708	55591847	1.30E-05	0.06	1.13E-02	0.074	5.27E-01	0.031	4.85E-07	0.06	8.84E-01	+ + + +	*LOC100128944*
7	rs10251504	77527824	3.36E-06	0.038	1.26E-01	0.029	5.14E-01	0.02	1.01E-06	0.036	4.05E-01	+ + + +	*MAGI2**
2	rs734556	224269573	6.23E-07	0.042	5.17E-01	0.013	6.67E-01	0.013	1.65E-06	0.036	4.72E-01	+ + + +	*SCG2 & AP1S3*
17	rs4796712	37240656	1.04E-05	0.057	6.52E-02	0.058	8.58E-01	0.011	2.34E-06	0.055	8.89E-01	+ + + +	*NT5C3B**
15	rs11070836	49223689	5.56E-05	0.033	3.84E-01	0.017	5.28E-03	0.083	3.46E-06	0.034	3.13E-01	+ + + +	*TNFAIP8L3*
3	rs1382167	21611547	4.82E-06	0.051	9.55E-01	0.002	2.12E-01	0.055	7.79E-06	0.045	4.76E-01	+ + + +	*ZNF385D**
20	rs1291101	34927500	4.18E-05	-0.038	3.84E-01	-0.019	5.38E-02	-0.063	8.01E-06	-0.037	1.50E-01	- - - -	*C20orf117 & C20orf118*
10	rs10794108	128413863	8.60E-08	0.051	4.43E-01	-0.017	9.81E-01	-0.001	8.64E-06	0.037	1.74E-02	+ + - -	*C10orf90 & DOCK1*
12	rs220583	13852010	1.85E-05	0.04	7.37E-02	0.036	5.05E-01	-0.022	1.45E-05	0.035	3.74E-02	+ + + -	*GRIN2B**
10	rs7078439	128409974	2.27E-07	0.046	5.79E-01	-0.011	7.52E-01	-0.01	1.62E-05	0.034	2.94E-02	+ + - -	*C10orf90 & DOCK1*
13	rs2065550	102902377	1.24E-05	0.056	8.87E-01	0.004	1.58E-01	0.084	2.09E-05	0.049	3.82E-01	+ + + +	*LOC728183 & DAOA*
10	rs11259285	14782881	7.31E-05	0.033	1.38E-01	0.03	7.08E-01	0.011	2.80E-05	0.031	3.21E-01	+ + + +	*FAM107B**
12	rs2284424	13880137	7.49E-05	0.035	1.13E-01	0.032	8.42E-01	0.006	2.84E-05	0.033	2.25E-01	+ + + +	*GRIN2B**
2	rs10172774	224249717	3.02E-05	0.036	3.51E-01	0.019	7.33E-01	0.01	3.23E-05	0.032	7.55E-01	+ + + +	*SCG2 &AP1S3*
2	rs7576475	80581631	3.28E-05	-0.037	5.77E-01	0.012	5.12E-02	-0.07	7.31E-05	-0.032	1.09E-01	- - + -	*CTN2**
13	rs574662	22650396	1.33E-05	-0.038	9.62E-01	-0.001	8.03E-01	-0.008	7.47E-05	-0.031	6.70E-02	- - - -	*LOC100130029 & SGCG*
14	rs2029614	98617020	3.09E-05	0.056	4.87E-01	0.02	9.08E-01	-0.006	8.20E-05	0.047	2.54E-01	+ + + -	*RPL3P4 & BCL11B*
10	rs11245122	128415036	4.44E-06	0.048	2.55E-01	-0.028	3.70E-01	0.033	9.78E-05	0.036	3.30E-02	+ + - +	*C10orf90 & DOCK1*
3	rs925440	191325669	6.53E-05	-0.04	1.85E-01	-0.028	5.62E-01	0.02	9.95E-05	-0.034	2.72E-01	- - - +	*LEPREL1*
1	rs12039255	203875633	1.90E-05	0.073	2.00E-01	0.044	1.73E-01	-0.077	1.03E-04	0.057	8.04E-02	+ + + -	*ELK4 & SLC45A3*
14	rs1959775	47679625	3.78E-06	0.037	4.75E-01	-0.014	9.57E-01	-0.002	1.06E-04	0.028	3.16E-02	+ + - -	*RPS15AP3 & RPL18P1*
10	rs509948	85538632	5.76E-05	-0.033	7.68E-01	0.006	2.13E-01	-0.036	1.58E-04	-0.028	3.05E-01	- - + -	*LOC728050 & GHITM*
3	rs2811518	129496335	4.97E-05	-0.049	8.47E-01	0.006	5.52E-01	-0.027	2.11E-04	-0.04	3.57E-01	- - + -	*EEFSEC**
8	rs1396976	131791979	5.56E-05	0.043	4.83E-01	0.017	4.97E-01	-0.027	2.19E-04	0.035	1.49E-01	+ + + -	*ASAP1 & ADCY8*
17	rs2362396	75469049	1.01E-04	0.035	7.09E-01	-0.008	1.33E-01	0.055	2.22E-04	0.03	2.68E-01	+ + - +	*CBX4 &TBC1D16*
10	rs10794113	128425326	2.39E-06	0.044	3.16E-01	-0.021	6.79E-01	-0.013	2.29E-04	0.03	1.82E-02	+ + - -	*C10orf90 & DOCK1*
12	rs10774156	3534965	6.90E-05	0.037	9.82E-01	0	6.44E-01	0.016	2.47E-04	0.03	4.38E-01	+ + + +	*PRMT8**
2	rs10176854	224284646	2.78E-05	0.038	7.84E-01	0.006	4.15E-01	-0.027	2.75E-04	0.029	1.44E-01	+ + + -	*AP1S3 & AP1S3*

CHR	SNP	BP	NELSON		LUSI (group III) males		LUSI (group IV) males		Meta-analysis			Direction of effect	Close(st) gene(s)
			p	B	p	B	p	B	p	B	Q		
13	rs1407999	93841289	8.22E-05	0.032	6.32E-01	0.009	5.52E-01	-0.018	3.92E-04	0.026	1.35E-01	+ + + -	*GPC6**
2	rs11679585	40544972	2.06E-05	-0.044	5.08E-01	0.015	7.93E-01	-0.011	4.08E-04	-0.033	1.03E-01	- - + -	*SLC8A1 & LOC729984*
3	rs17246389	168672301	7.60E-05	0.037	4.06E-01	-0.018	2.05E-01	0.042	4.25E-04	0.03	1.25E-01	+ + - +	*SERPINI2**
13	rs554393	101633794	2.51E-05	0.037	7.43E-01	0.007	1.19E-01	-0.048	4.82E-04	0.028	2.53E-02	+ + + -	*FGF14**
9	rs7047287	118828951	1.87E-05	0.04	2.02E-01	-0.026	6.66E-01	0.017	7.44E-04	0.028	3.26E-02	+ + - +	*ASTN2**
13	rs534298	112791769	9.34E-05	-0.039	9.86E-01	0	7.00E-01	0.014	8.08E-04	-0.029	1.11E-01	- - - +	*MCF2L**
5	rs919336	31806568	4.91E-05	-0.034	3.54E-01	0.017	9.21E-01	-0.003	1.08E-03	-0.024	8.28E-02	- - + -	*C5orf22 & PDZD2*
3	rs10154906	135980476	5.22E-05	-0.04	1.75E-01	0.032	5.99E-01	-0.021	1.21E-03	-0.029	3.20E-02	- - + -	*KY & EPHB1*
10	rs11018027	133229112	8.03E-05	0.053	1.57E-01	-0.042	1.76E-01	0.069	1.21E-03	0.039	2.57E-02	+ + - +	*TCERG1L & FLJ46300*
14	rs4905002	92420129	3.73E-05	0.036	4.24E-01	-0.016	5.70E-01	-0.019	1.22E-03	0.025	2.00E-02	+ + - -	*GOLGA5 & CHGA*
4	rs679959	138143227	5.37E-05	0.035	3.77E-01	-0.018	3.42E-01	-0.036	2.04E-03	0.024	1.13E-02	+ + - -	*LOC646316 & LOC729307*
1	rs4845364	152408532	4.86E-05	-0.034	3.27E-01	0.018	6.37E-01	0.013	2.10E-03	-0.022	3.86E-02	- - + +	*TPM3**
5	rs4705990	132431332	6.62E-05	0.037	3.20E-01	-0.02	6.25E-01	-0.018	2.49E-03	0.025	4.55E-02	+ + - - -	*HSPA4**
11	rs550897	113947313	5.95E-05	0.033	6.34E-01	-0.009	1.19E-01	-0.046	2.86E-03	0.021	6.80E-03	+ + - -	*FAM55D**
20	rs4815352	2497008	5.61E-05	0.033	2.22E-03	-0.059	1.13E-01	0.047	4.94E-03	0.02	1.00E-04	+ + - +	*TMC2**
4	rs2089540	101918405	3.03E-05	0.038	5.13E-02	-0.04	4.41E-01	-0.028	5.62E-03	0.022	2.00E-03	+ + - -	*EMCN & LOC728771*
7	rs10232434	88935516	3.49E-05	-0.038	2.58E-02	0.048	2.75E-01	0.04	9.18E-03	-0.021	7.00E-04	- - + +	*ZNF804B & DPY19L2P4*
12	rs7303397	20385638	3.97E-05	0.035	3.48E-02	-0.044	1.50E-01	-0.044	1.12E-02	0.019	4.00E-04	+ + - -	*LOC644976 & LOC100131677*

Group I: NELSON Groningen population; Group II: NELSON Utrecht population; Group III: LUSI-Toshiba population;

*Group IV: LUSI-Siemens population; BP = Base pair; Q = p-value for heterogeneity; Direction of effect per cohort: each sign reflects one cohort, direction of effect is presented by: + = $(B > 0)$, – = $(B < 0)$, 0 = no effect; *Corresponding SNP is located in an intron in this gene.*

Table 5. Association analyses on airway wall thickening of 47 SNPs (identified in NELSON) followed by replication in LUSI groups III and IV including individuals with FEV_1/FVC < 80% and meta-analysis in NELSON and LUSI groups III (FEV_1/FVC < 80%) and IV (FEV_1/FVC < 80%).

			FEV_1/FVC < 80%									
CHR	SNP	BP	NELSON		LUSI group III		LUSI group IV		Meta-analysis		Direction of effect	Closest gene(s)
			p	B	p	B	p	B	p	B		
2	rs734556	224269573	6.23E-07	0.042	3.29E-02	0.050	4.02E-01	0.029	4.60E-08	0.043	+ + + +	*SCG2 & AP1S3*
10	rs7078439	128409974	2.27E-07	0.046	4.64E-01	0.018	8.57E-02	0.063	7.44E-08	0.044	+ + + +	*C10orf90 & DOCK1*
10	rs10794108	128413863	8.60E-08	0.051	5.74E-01	0.015	2.07E-01	0.048	7.53E-08	0.047	+ + + +	*C10orf90 & DOCK1*
17	rs4796712	37240656	1.04E-05	0.057	1.28E-01	0.058	3.21E-01	0.067	1.86E-06	0.057	+ + + +	*NT5C3B**
7	rs10251504	77527824	3.36E-06	0.038	2.75E-02	0.051	2.40E-01	-0.041	2.14E-06	0.036	+ + + -	*MAGI2**
15	rs11070836	49223689	5.56E-05	0.033	6.33E-01	0.011	6.13E-04	0.109	2.46E-06	0.036	+ + + +	*TNFAIP8L3 & CYP19A1*
10	rs10794113	128425326	2.39E-06	0.044	8.40E-01	0.005	3.48E-01	0.036	4.19E-06	0.039	+ + + +	*C10orf90 & DOCK1*
10	rs11245122	128415036	4.44E-06	0.048	6.29E-01	-0.015	3.65E-02	0.086	5.12E-06	0.044	+ + - +	*C10orf90 & DOCK1*
2	rs10172774	224249717	3.02E-05	0.036	5.29E-02	0.047	7.44E-01	0.011	5.70E-06	0.036	+ + + +	*SCG2 & AP1S3*
14	rs2029614	98617020	3.09E-05	0.056	8.10E-02	0.060	6.38E-01	0.029	6.16E-06	0.055	+ + + +	*RPL3P4 & BCL11B*
2	rs10176854	224284646	2.78E-05	0.038	1.11E-01	0.041	6.14E-01	0.020	7.14E-06	0.037	+ + + +	*SCG2 & AP1S3*
10	rs11259285	14782881	7.31E-05	0.033	3.03E-01	0.025	6.44E-02	0.058	9.97E-06	0.034	+ + + +	*FAM107B**
3	rs1382167	21611547	4.82E-06	0.051	8.00E-01	0.009	7.47E-01	-0.014	2.14E-05	0.044	+ + + -	*ZNF385D**
9	rs7047287	118828951	1.87E-05	0.040	7.88E-01	0.007	3.55E-01	0.036	2.49E-05	0.036	+ + + +	*ASTN2**
12	rs1391708	55591847	1.30E-05	0.060	1.51E-01	0.053	4.03E-01	-0.037	2.95E-05	0.051	+ + - -	*LOC100128944*
3	rs2811518	129496335	4.97E-05	-0.049	8.18E-01	-0.008	1.38E-01	-0.064	3.51E-05	-0.045	- - - -	*EEFSEC**
2	rs11679585	40544972	2.06E-05	-0.044	3.57E-01	-0.025	7.56E-01	0.013	3.60E-05	-0.039	- - - +	*SLC8A1 & LOC729984*
14	rs1959775	47679625	3.78E-06	0.037	5.91E-01	-0.013	9.62E-01	0.002	3.91E-05	0.031	+ + - +	*RPS15AP3 & RPL18P1*
13	rs2065550	102902377	1.24E-05	0.056	4.53E-01	-0.026	1.09E-01	0.090	3.93E-05	0.048	+ + - +	*LOC728183 & DAOA*
10	rs11018027	133229112	8.03E-05	0.053	6.25E-01	0.018	2.11E-01	0.066	5.00E-05	0.050	+ + + +	*TCERG1L & FLJ46300*
8	rs1396976	131791979	5.56E-05	0.043	2.08E-01	0.038	7.72E-01	-0.012	5.64E-05	0.039	+ + + -	*ASAP1 & ADCY8*
2	rs7576475	80581631	3.28E-05	-0.037	9.99E-01	0.000	4.91E-01	-0.027	6.36E-05	-0.033	- - 0 -	*CTN2**
13	rs534298	112791769	9.34E-05	-0.039	1.42E-01	-0.037	5.78E-01	0.021	8.43E-05	-0.035	- - - +	*MCF2L**

CHR	SNP	BP	FEV$_1$/FVC < 80% NELSON		LUSI group III		LUSI group IV		Meta-analysis		Direction of effect	Closest gene(s)
			p	B	p	B	p	B	p	B		
3	rs925440	191325669	6.53E-05	-0.040	2.12E-01	-0.031	4.13E-01	0.032	1.10E-04	-0.035	- - - +	*LEPREL1*
11	rs550897	113947313	5.95E-05	0.033	3.23E-01	0.022	5.23E-01	-0.022	1.10E-04	0.029	+ + + -	*FAM55D**
1	rs12039255	203875633	1.90E-05	0.073	2.96E-01	0.042	1.86E-01	-0.080	1.11E-04	0.059	+ + + -	*ELK4 & SLC45A3*
12	rs10774156	3534965	6.90E-05	0.037	7.90E-01	0.007	7.05E-01	0.014	1.25E-04	0.033	+ + + +	*PRMT8**
14	rs4905002	92420129	3.73E-05	0.036	7.60E-01	-0.008	5.78E-01	0.019	1.34E-04	0.030	+ + - +	*GOLGA5 & CHGA*
3	rs17246389	168672301	7.60E-05	0.037	8.38E-01	-0.005	2.38E-01	0.044	1.34E-04	0.033	+ + - +	*SERPINI2**
13	rs554393	101633794	2.51E-05	0.037	8.89E-01	-0.004	8.40E-01	-0.007	1.45E-04	0.031	+ + - -	*FGF14**
7	rs10232434	88935516	3.49E-05	-0.038	8.45E-01	0.005	8.30E-01	-0.008	1.51E-04	-0.032	- - + -	*ZNF804B & DPY19L2P4*
12	rs220583	13852010	1.85E-05	0.040	8.10E-01	0.006	3.73E-01	-0.033	1.57E-04	0.032	+ + + -	*GRIN2B**
12	rs2284424	13880137	7.49E-05	0.035	4.96E-01	0.016	7.21E-01	-0.013	1.63E-04	0.031	+ + + -	*GRIN2B**
5	rs919336	31806568	4.91E-05	-0.034	9.12E-01	0.002	9.17E-01	-0.003	2.24E-04	-0.028	- - + -	*C5orf22 & PDZD2*
20	rs1291101	34927500	4.18E-05	-0.038	8.14E-01	0.006	7.86E-01	0.010	2.97E-04	-0.031	- - + +	*C20orf117 & C20orf118*
3	rs10154906	135980476	5.22E-05	-0.040	7.72E-01	0.009	8.25E-01	0.008	3.40E-04	-0.033	- - + +	*KY & EPHB1*
17	rs2362396	75469049	1.01E-04	0.035	5.04E-01	-0.017	3.46E-01	0.034	3.72E-04	0.030	+ + - +	*CBX4 &TBC1D16*
20	rs4815352	2497008	5.61E-05	0.033	5.79E-01	-0.013	8.91E-01	0.004	3.75E-04	0.027	+ + - +	*TMC2**
1	rs4845364	152408532	4.86E-05	-0.034	5.47E-01	0.014	7.71E-01	0.009	5.73E-04	-0.026	- - + +	*TPM3**
13	rs1407999	93841289	8.22E-05	0.032	8.39E-01	-0.005	5.06E-01	-0.022	6.63E-04	0.026	+ + - -	*GPC6**
9	rs10429583	75056788	9.33E-05	-0.034	5.96E-01	0.013	8.43E-01	0.007	7.14E-04	-0.027	- - + +	*ANXA1 & LOC138971*
12	rs7303397	20385638	3.97E-05	0.035	2.90E-01	-0.028	3.08E-01	-0.037	1.03E-03	0.026	+ + - -	*LOC644976 & LOC100131677*
5	rs4705990	132431332	6.62E-05	0.037	1.52E-01	-0.037	7.86E-01	-0.011	1.67E-03	0.027	+ + - -	*HSPA4**
4	rs679959	138143227	5.37E-05	0.035	2.73E-01	-0.027	1.28E-01	-0.055	2.54E-03	0.024	+ + - -	*LOC646316 & LOC729307*
4	rs2089540	101918405	3.03E-05	0.038	7.48E-02	-0.044	1.39E-01	-0.052	4.37E-03	0.024	+ + - -	*EMCN & LOC728771*

*Group I: NELSON Groningen population; Group II: NELSON Utrecht population; Group III: LUSI-Toshiba population; Group IV: LUSI-Siemens population; BP = Base pair position; Direction of effect per cohort: each sign reflects one cohort, direction of effect is presented by: + = (B > 0.010), – = (B < -0.01) and 0 = no effect; *Corresponding SNP is located in an intron in this gene.*

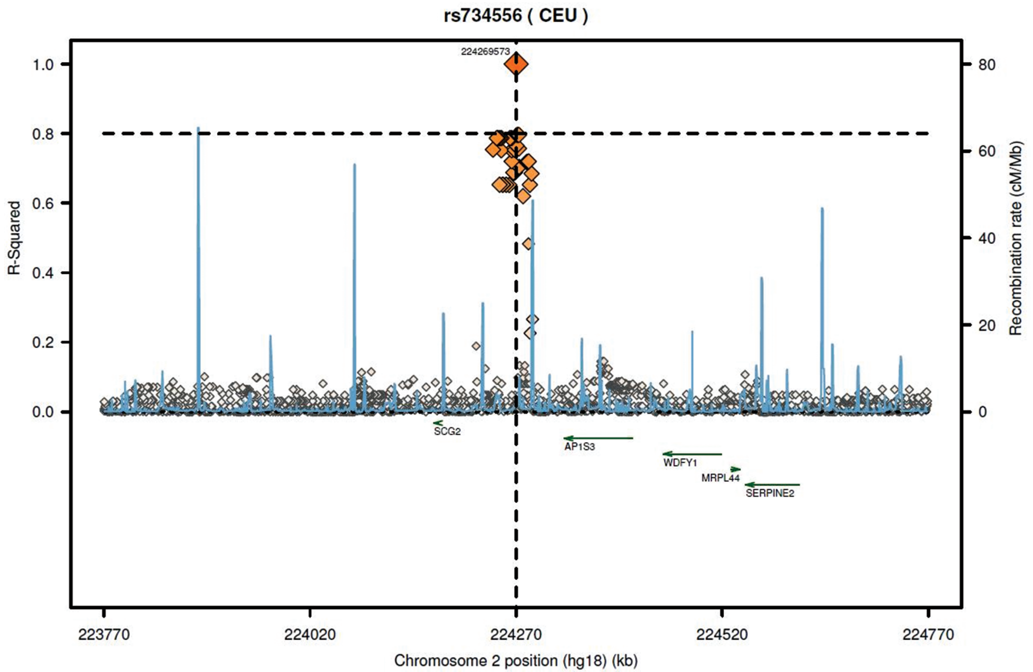


Figure 1. Position of rs734556 on chromosome 2.

References

1 Bolstad BM, Irizarry RA, Astrand M and Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics (Oxford, England)*. 2003;19(2):185-193

2 Hao K, Bosse Y, Nickle DC et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS genetics*. 2012;8(11):e1003029

3 Fehrmann RS, Jansen RC, Veldink JH et al. Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS genetics*. 2011;7(8):e1002197

4 van den Berge M, Steiling K, Timens W et al. Airway gene expression in COPD is dynamic with inhaled corticosteroid treatment and reflects biological pathways associated with disease activity. *Thorax*. 2014;69(1):14-23

Chapter 8

Summary, discussion and future perspectives

AE Dijkstra

Summary

This thesis encompasses studies on chronic mucus hypersecretion (CMH) and airway wall thickening (AWT). CMH and AWT are both features of chronic airway diseases and associated with impaired lung function. CMH is a troublesome symptom for many people; it is more prevalent at elderly ages and affects an individual's quality of life.
We studied risk factors for CMH in different populations with respect to demographic characteristics, environmental exposures and genetic influences. To determine AWT we first developed a method to optimally assess the thickness of the airway wall and used low dose CTs with new software, a novel approach for quantification of AWT. Besides studying demographic and clinical characteristics of AWT in the same population based cohort, we subsequently explored genetic risk factors by performing a genome wide analysis.

Chapter 1 presents an introduction on CMH and AWT. The underlying cause of CMH and AWT is chronic inflammation of the airway epithelium and underlying structures incited by environmental factors like exposure to cigarette smoke and occupational irritants. The consequence of this irritant exposure is the development of goblet cells (mucus producing cells) in the epithelium and mucus glands in the sub-epithelial layers as well as changes in the mucus constituents produced, contributing to expectoration of mucus in greater quantities and different compositions. Furthermore AWT is due to changes in the matrix composition and edema in both the epithelial and subepithelial layers of the airways. Several environmental factors may contribute to the development of both CMH and AWT. Since not every person who has been comparably exposed to these agents develops CMH or AWT, we investigated whether there are genetic differences between people with and without CMH, and people with abnormal and normal AWT that can explain this difference in susceptibility.

Chapter 2 reports results from a large cross-sectional general population based study, relating demographic characteristics, environmental tobacco smoke exposure (ETS), smoking habits and occupational exposures to CMH, stratified by COPD. Individuals with COPD had a higher prevalence of CMH compared to individuals without COPD, respectively 8.7% and 3.4%. A "job exposure matrix" was used In this study to link occupations and exposures. Exposure to gases & fumes was the most frequent occupational exposure in this population (40.1%). In individuals with and without COPD, the strongest predictor for CMH was a higher number of packyears smoked. In individuals with COPD occupational exposures were not significantly associated with

CMH. In contrast, next to smoking status, high occupational exposure to gases & fumes is an important risk factor of CMH in individuals without COPD. These findings suggest that the inducing mechanisms underlying CMH differ between individuals with and without COPD.

Chapter 3 describes our study on associations between CMH and genetic variants (SNPs) and corresponding genes. To this aim we performed a genome wide analysis of individuals with and without CMH participating in a large heavy smoking population based lung cancer study, including almost 500,000 SNPs. The most significant SNPs were analyzed in 11 other heavy smoking populations. A strong association with CMH, consistent across all cohorts, was observed with a SNP located in the *special AT-rich sequence-binding protein 1 locus (SATB1)* on chromosome 3, despite cohort differences in the definition of CMH and severity of airflow limitation. The odds ratio for this SNP suggests an additional risk of 17% per G allele to suffer from CMH in this population of ex- and current heavy smokers (n = 10,328, CMH prevalence is 26.2%). Additional functional studies confirmed this result. The risk allele was associated with higher mRNA expression of *SATB1* in lung tissue, presence of CMH in COPD patients was associated with increased *SATB1* mRNA expression in bronchial biopsies and SATB1 expression was induced during differentiation of primary human bronchial epithelial cells in culture. Taken together, our findings provide suggestive evidence that *SATB1* plays an important role in CMH.

Since not all individuals with COPD have CMH and, conversely, many individuals with CMH do not have COPD we investigated in **Chapter 4** whether there are specific genetic variants related to CMH in ex and current heavy smoking individuals with and without COPD. Genome wide association studies on CMH and subsequent replication yielded no genome wide significant results, neither in individuals with COPD nor in individuals without COPD. In addition, we found no significant overlap in SNPs associated with CMH in both groups. However, we found some genes with reasonably small p-values for association and replication with the same direction of effects of the SNPs in the identification and replication cohort(s). We explored these SNPs further to assess whether they were associated with the expression of a specific gene in lung tissue. In COPD, lower *GDNF* mRNA expression in bronchial biopsies was significantly associated with CMH, possibly by preventing expression by *GDNF-AS1,* our top gene. Furthermore, in individuals without COPD, a top-SNP in *MAML3* that nominally replicated in another cohort was an eQTL in lung tissue. Our results suggest genetic heterogeneity of CMH in individuals with and without COPD and indicate that it is worthwhile to repeat this study in much larger and homogeneously populated cohorts.

Chapter 5 describes the research we performed to investigate the relation between AWT and clinical and demographic characteristics of heavy smokers from a general population based cohort. To assess AWT we used advanced automated software that was able to select airways with a fixed internal diameter (3.5mm) in each lung lobe, and to calculate the corresponding AWT based on density differences. We assessed AWT and extent of emphysema of almost 500 heavy smoking individuals on low-dose CTs.
We concluded that post processing standardization of large numbers of airway wall measurements in all lung lobes is a feasible, reliable and useful method to assess AWT. We demonstrated that increased AWT is a more important factor for airflow limitation than emphysema in a smoking male population, independently from smoking behavior and respiratory symptoms. AWT at this lumen size was not significantly associated with CMH.
Since it is suggested that dimensions of airways with an internal diameter of 2.5 mm or more reflect the dimensions of small airways, and small airways are the major cause of airway obstruction in COPD, this method may provide insight into the changes therein [1, 2].

Chapter 6 describes the differences in airway wall measurements in larger airways (≥ 5mm) between male current and former heavy smokers, 50 with respiratory symptoms (cough, mucus, dyspnea and wheezing) and in 50 without these symptoms. AWT and the percentage of circumferential airway wall that could be measured were assessed in five selected bronchi, one in each pulmonary lobe, distributed over 3 internal luminal airway sizes. It was shown that male heavy smokers with chronic respiratory symptoms had significantly thicker airway walls than those without respiratory symptoms in airways with a luminal diameter from 5 to 10 mm, but not in larger airways. This result strengthen our idea to analyze different internal airway diameters all along the bronchial tree. Their relative contribution to symptoms and bronchial obstruction may provide different pathogenetic information in the various phenotypes of COPD.

Chapter 7 describes the genome wide association study we performed to search for SNPs/genes associated with AWT. To assess AWT we used recently developed software, the same as described in chapter 5, and low-dose CTs of male heavy smokers participating in the NELSON-study, from Utrecht and Groningen. We studied the association of mean AWT at 3.5 mm internal lumen size with > 500,000 SNPs, covering the whole genome. To replicate our top SNPs, i.e. SNPs showing the most significant association with AWT, we used a lung cancer screening study from Heidelberg, Germany, a cohort with the same inclusion criteria and availability of low-dose CTs. We identified three significant loci and replication of these 3 SNPs in cohorts with comparable lung function

reached genome wide significance. One of these SNPs was located close to *SERPINE2* and revealed to be an eQTL associated with higher SERPINE2 expression in lung tissue. *SERPINE2* was identified previously as a susceptibility gene for COPD and particularly emphysema, providing here the link with AWT. Furthermore, we found two nominally significant SNPs with effects in the same direction and additional functional studies showed that higher expression of one corresponding gene *(MAGI2)* was associated with lower numbers of bronchial inflammatory cells and the other gene (NT5C3B) with worse lung function in COPD patients. Moreover, a SNP in *NT5C3B* was an eQTL associated with lower *NT5C3B* expression in lung tissue. This study shows some of the genes associated with AWT are also associated with lung function, airway inflammation or emphysema, thus showing clearly the interrelationship between AWT and COPD.

Discussion and future perspectives

An important goal of medical research is to improve the health of human beings, taking into account the variation between humans that may result from differences in exposure to environmental factors and their genetic background. These health variations may lead to differences between individuals that are emphasized by their susceptibility to develop (airway) disease. This thesis is about investigating the complex interaction between environmental exposures and genetic variation that induce chronic mucus hypersecretion (CMH) and airway wall thickening (AWT).

CMH and risk factors

One of the prominent messages of this thesis is the overwhelming detrimental effect of smoking. This is shown by the difference in CMH prevalence between the study populations used in the current thesis. The prevalence of CMH is 28.5 % in the NELSON study, a lung cancer screening study including elderly individuals (male current and ex heavy smokers, mean age 60.3 ± 5.5 years, ≥ 20 packyears) primarily investigating the ability to detect early lung cancer as well as assessing the health gain of early detection of lung cancer [3]. In contrast, the prevalence of CMH was only 4.2% in the LifeLines cohort study, a general population based study (mean age 48.1 ± 10.9 years) intending to find an answer to the question why some people remain healthy into old age and others develop a chronic disease at relatively young age [4].

The main reasons for this difference can be attributed to the inclusion criteria: an elderly population in the NELSON study that smoked an extremely high number of packyears (mean packyears 40.7 ± 17.4) versus the general population of LifeLines (mean packyears 7.6 ± 10.8). From this comparison and our studies in CMH it becomes clear that smoking and age are important risk factors for CMH, and the longer you smoke, the higher the risk to develop CMH.

The obvious solution to prevent CMH development would be smoking cessation, or even better, to prevent starting smoking in youngsters and to prevent smoking and addiction to cigarettes completely. Despite the present knowledge and information about this hazard the prevalence of smoking hardly decreases (prevalence >15 yrs: males 27%, females 25%, 2012, RIVM). The current legislation by the Dutch government, i.e. a smoking ban in public areas and increased taxes on tobacco products, is difficult to enforce and in almost 30% of the Dutch pubs, smoking was still present as noticed in the newspaper in late November 2013 [5].

The approach of the medical world, namely stimulating awareness and smoking cessation counseling, fails to discourage smoking and to refrain young people from starting to smoke.

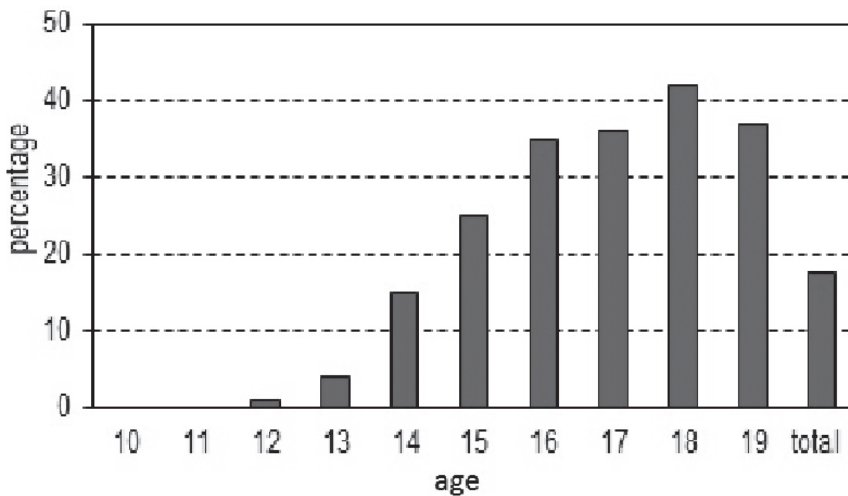


Figure 1. Percentage young people (10-19-yrs) that indicates in 2012 to have smoked during the last four weeks, by age. Due to small numbers in the sample, no distinction is made between boys and girls (RIVM).

A bright spot is the ban on selling cigarettes to young people under the age of 18 starting from January 2014 in the Netherlands. As the first symptoms of lung disease manifest at later age, an age that is far out of sight of adolescents, the risk of smoking is not taken seriously by these youngsters (Figure 1). Later, when addiction plays a role, smoking cessation is much more difficult. Thus preventing smoking at an early age is a health investment that pays off during the complete lifespan and for all diseases related to smoking, since not only respiratory disease is a smoking related disease. Also cardiovascular diseases, diabetes and a number of cancers, like head and neck and bladder cancers, are smoking-related diseases with a high mortality risk. In addition, smoking increases the risk for many diseases like Crohn's disease or leukemia and may adversely affect the course of a disease, e.g. multiple sclerosis [6]. This needs more attention in the society at large as well as more research on the best strategies to ban cigarette smoking worldwide.

However, banning smoking would not solve the whole problem, since people who never smoked and were not exposed to smoke by others, may develop CMH (1.4%, 58 out of 4,188 never smokers in the LifeLines cohort (asthmatics were excluded)). Other risk factors for CMH are inhaled agents individuals have been exposed to professionally, like gases & fumes (chapter 2). Exposure to gases & fumes is a.o. associated with jobs like heavy truck and lorry drivers, motor vehicle mechanics, welders and flame-cutters,

agricultural and industrial mechanics, plumbers and pipe fitters and painters. Through legislation and quality standards regarding working conditions people are may be less exposed in the near future to these hazards (in Western Europe). In addition, it is important that people are aware of the risks and that they are also responsible themselves.

There is growing evidence that COPD is not simply a disease of old age largely restricted to heavy smokers but finds its origin in childhood [7]. Recently it was shown that genes known to be associated with COPD are also associated with early childhood wheeze, which -by itself- is a risk factor for a worse lung function in child- and adulthood [8]. The question thus arises whether children with early symptoms of CMH like coughing and phlegm, are also more susceptible to develop CMH in adulthood. It would be of interest to investigate whether genes associated with CMH at adult age are also associated with coughing in childhood.

CMH has to be taken seriously. As already described in the introduction of this thesis, CMH was (and is) often not seen as an important and clinically relevant symptom: it simply is regarded to be "related to" smoking, without further consequences. The opposite is true: CMH is not an innocent symptom, and this can be deduced from the fact that the guidelines of the Global Initiative for Chronic Obstructive Lung Disease (GOLD Guidelines, 2014) frequently refer to the presence of CMH [9, 10]. Illustrative is the recommendation in these guidelines to use the Clinical COPD Questionnaire (CCQ) or the COPD Assessment Test (CAT), both questionnaires including respiratory symptoms, a.o. CMH, for assessing and monitoring COPD, based on the evidence that an FEV_1 is only a partial descriptor of disease status [11]. The CCQ was developed in Groningen and is a simple and short 10-items self-administered questionnaire to measure clinical control in COPD patients and a reliable and suitable tool [12].The CAT contains 8 items and has good measurement properties and is sensitive to the change in health status associated with exacerbations of COPD [13, 14].
Recently a longitudinal study by Putcha et al. investigating almost 6,000 smokers, results showed that the presence of mucus accompanied by cough at baseline, defined as usually having a cough and usually producing phlegm, was associated with accelerated lung function decline and with mortality [15]. A large international population based study (n > 6,000) described the relationship between chronic cough or phlegm and the Health Related Quality of Life score (HRQL) in adults (age 20-48 years) and revealed that presence of these symptoms was associated with a considerable impairment of the quality of life independently from other respiratory diseases like asthma, COPD or bronchial hyperresponsiveness [16]. Thus the presence of respiratory symptoms, which is easy to score, may partly explain a person's well-being [17].

Above corroborates my conclusion that further extensive research on the underlying mechanisms of and risk factors for CMH is relevant.

Diagnosis of CMH

Research on CMH is complicated by the fact that presence or absence of CMH is a subjective phenomenon, driven by perception of an individual and associated with cultural habits and acceptance. Furthermore, there is no objective instrument to determine the amount of produced mucus. Moreover, it is conceivable that some people actually cough up mucus and perceive it as a habit, while others silently swallow mucus without mentioning it in health questionnaires. Or even, as research by Chapman et al. showed, that women, consciously or unconsciously, are prone to underreporting of CMH (as CMH is not a ladylike symptom) and that cultural and socio-economic differences between countries affect the reporting frequencies of these symptoms [18, 19].

We performed several studies to determine demographic, environmental and genetic risk factors for CMH. An overview of results is given in the table 1.

As shown in this table, CMH is associated with demographic, environmental and genetic factors. When so many factors are involved, one would think that different CMH phenotypes are assumed to exist.

Table 1. Environmental and genetic factors associated with chronic mucus hypersecretion

Chapter		CMH		
		General population	COPD	Non COPD
2	Characteristics			Male gender Higher BMI
	Smoking	Packyears	Packyears ETS	Packyears Current smoking
	Occupational exposures	Gases & fumes (high) Aromatic solvents (low)		Gases & fumes (high) Mineral dust (low/high) Chlorinated solvents (high) Heavy metals (high)
3 + 4	GWAs – SNP	rs6577641	rs10794108	rs4863687
	GWAs – close(st) gene	*SATB1*	*GDNF-AS1*	*MAML3*
	mRNA expression in lung tissue	eQTL		eQTL
	Gene expression in COPD associated with CMH	*SATB1*	*GDNF*	

CMH = chronic mucus hypersecretion; ETS = environmental tobacco smoke

8

In chapter **2** we investigated risk factors for CMH in the general population, divided in individuals with and without COPD. We observed that in both groups a higher number of packyears smoked is an important risk factor for CMH. Besides this there are differences between individuals with and without COPD, the effect of gases & fumes exposure being the most remarkable. Exposure to gases & fumes is a risk factor for CMH in individuals without but not in individuals with COPD, as displayed in Figure 2. This supports our above mentioned hypothesis; there are subgroups of individuals with CMH, i.e. different CMH-phenotypes.

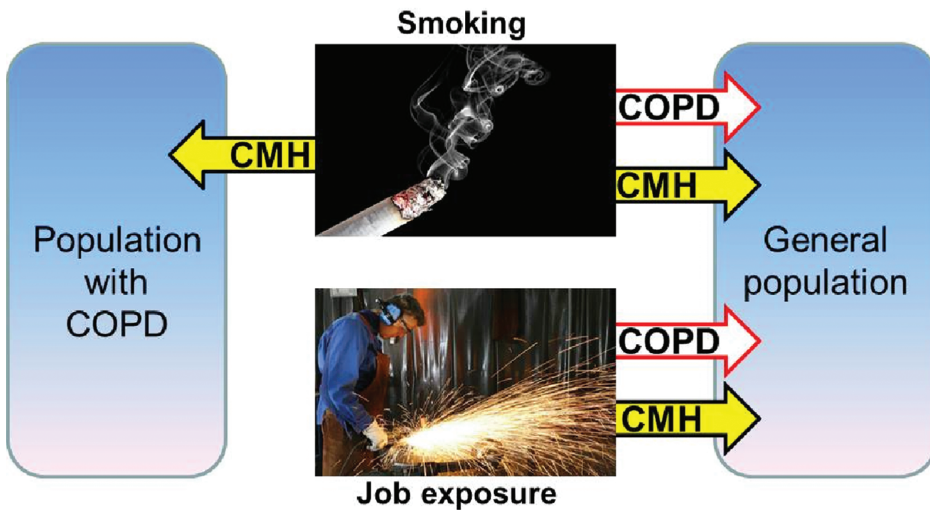


Figure 2. Smoking and occupational exposures are risk factors for COPD and for CMH in the general population. In COPD only smoking is a risk factor for CMH and occupational exposures are not.

There are several possible explanations for this observation. It is tempting to conclude that occupational exposures (gases & fumes) do not penetrate the smaller airways whereas smoke does, and only the larger airways are involved in individuals with CMH without COPD. However, recent research in a general population based study showed that FEF_{25-75}, the parameter commonly used to determine small airways obstruction, was significantly associated with exposures to gases & fumes. Even after exclusion of individuals with obstruction of the large airways (i.e. those with $FEV_1/FVC \leq 70\%$, $FEV_1 \leq 80\%$), this association was still present [20]. This suggests that gases & fumes also influence the smaller airways. Hence it might be that small airways are stimulated to produce less or different types of surfactants for instance, contributing to the tenacity of CMH. This requires further study.

Another possible explanation is that a gradual process is taking place in the airways. In the first phase there is a situation in which CMH prevails, then by continued exposure to cigarette smoke or occupational exposure to gases & fumes the inflammatory process expands and the individuals who are susceptible to the effects of these exposures develop COPD (i.e. inflammation and remodelling in the smaller and larger airways). By these changes, the airway wall is no longer susceptible for exposure to gases & fumes, or might even be protected from it.
Since CMH is also a feature of COPD and co-exists to COPD, the above is suggestive for the idea that CMH in individuals with COPD is different from CMH in individuals without COPD and that CMH in heavy smokers is different from CMH in individuals exposed to gases & fumes.
Findings in the literature combined with our suggestive findings for differential genetic variants and environments involved in CMH in individuals with and without COPD, show that it is worthwhile to study more precisely where mucus (over)production is located in the airways. As CMH is associated with an increased number of gobletcells, remodelling and inflammation of the airways, it is obvious that CMH is accompanied by airway wall thickening. Further research should find an answer to whether there is a relation between exposure to gases & fumes and AWT.
Since AWT measurements are based on density differences, it is a problem that in regions where surrounding tissue is present or the airway wall is less easy to distinct (the more peripheral airways), the technology to determine AWT is insufficient. Development of a method to measure small airway dimensions preferably on low dose CTs will be very meaningful to investigate the location of mucus production along the full airways.

We have seen that there are CMH differences regarding genetic and environmental risk factors and regarding absence or presence of airway obstruction. In addition, there may be differences regarding the location of mucus-production, the composition of produced mucus and the composition of the airway epithelium where it is produced. Are there differences between CMH associated with smoke exposure and occupational exposure, between CMH in individuals with or without COPD regarding composition, tenacity, viscosity and produced volume of mucus, as well as the type and level of inflammation of the epithelium and damage or loss of ciliary function? (Figure 3) Comparison of mucus composition and investigating the function of airway epithelial tissue in vitro may contribute to understanding CMH phenotypes.

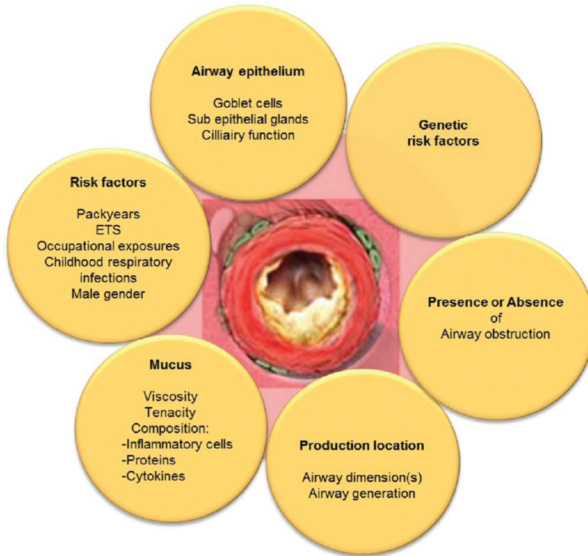


Figure 3. Features of and risk factors for CMH requiring further research.

Risk factors for AWT

AWT is the result of two ongoing processes: inflammation and tissue repair, both contributing to remodeling. As shown in the Table 2 below AWT is a phenotype that is associated with both genetic and environmental factors.

Table 2. Environmental and genetic factors associated with airway wall thickening

Chapter	AWT					
	Smoking		Packyears			
6	Lung function		FEV1,% predicted			
	Emphysema		Perc15			
7	GWAs	SNP	rs4796712	rs10251504	rs734556	rs10794108 & rs7078439
		in/close to gene	*NT5C3B*	*MAGI2*	*SERPINE2*	*C10orf90 & DOCK1*
	mRNA expression in lung tissue		eQTL		eQTL	
	Gene expression in COPD associated with		airflow obstruction	inflammatory markers		

AWT = airway wall thickness

We demonstrated that AWT at an internal lumen diameter of 3.5 mm ($AWT_{3.5}$) was associated with a lower lung function (FEV_1 % predicted); the thicker the airway wall, the more obstruction. We performed a cross sectional study and it is not clear how AWT evolves over time. This needs further study. Another important finding is that $AWT_{3.5}$ explained only 31.1% of the variance in FEV_1%predicted in the multivariate analysis (adjusted for emphysema, smoking behavior, lung volume and respiratory symptoms). The current hypothesis is that the inflammatory and remodelling process in the small airways, causing obstruction and obliteration of the terminal bronchioles, is the main cause of developing COPD [2]. The parameter representing small airways obstruction commonly used is FEV_{25-75}, and it would be of interest to study the association of AWT to this parameter. Possibly AWT is a better representative of the severity of airway disease than the commonly used FEV_1% predicted. Alternatively more specific techniques that can detect small airway obstruction would be of interest, like a Multiple Nitrogen Breath washout test, or Impulse Spirometry (IOS) [21, 22].

Identified SNPs and gene-expression

Some of the SNPs we found to be associated with CMH or AWT in the GWA-studies (Table 1 & 2) were expression quantitative trait loci; eQTLs. This means that these SNPs regulate expression levels of proteins for a gene. Furthermore we found *GDNF-AS1*, an antisense RNA transcribed to prevent translation of a complementary mRNA by base pairing to it and blocking translation. In this way *GDNF-AS1* prevents expression of *GDNF*.

We showed that some SNPs in genes that are relevant to CMH are associated with gene expression in lung tissue. It seems of importance, given the effects of smoking on DNA-methylation, to assess whether other factors besides the identified genes contribute to CMH. Thus methylation and other epigenetic phenomena, like micro-RNAs have to be studied further. To investigate gene expression levels of genes associated with CMH or AWT, we used bronchial biopsies and lung tissue. This may not specifically capture the cell type for which a gene is important. Thus, e.g. cultured epithelial cells may be an interesting cellular source since epithelial cells are the first barrier in the airways.

A cross-sectional study comparing gene-expression levels between individuals with CMH caused by smoking, by occupational exposures in the absence or presence of COPD compared to individuals without not exposed individuals without COPD and without CMH (Figure 4), could show gene expression differences, providing insight into interactions between genes and exposures.

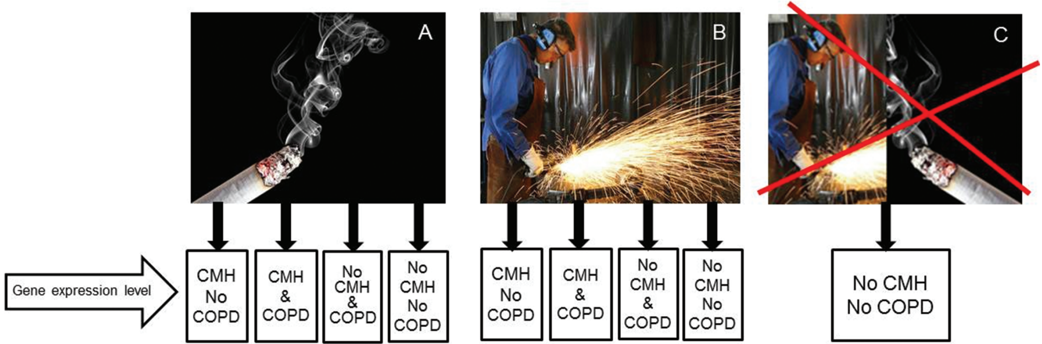


Figure 4. Research design: comparison of gene-expression levels in smoking cases with and without COPD or CMH, occupational exposed cases with and without COPD or CMH and never smoking, never occupational exposed controls with COPD nor CMH (Figure 3), could show gene expression differences.

Understanding the genetic mechanisms that underlie airway diseases like CMH and AWT will help to determine whether there are identifiable groups that respond differently to environmental exposures like smoking and occupational exposures. This thesis shows that environmental risk factors and genetic variants play a role in both diseases and that they affect each other. The underlying pathological processes are still largely unexplored.

References

1 Nakano Y, Wong JC, de Jong PA et al. The prediction of small airway dimensions using computed tomography. *American Journal of Respiratory and Critical Care Medicine : An Official Journal of the American Thoracic Society, Medical Section of the American Lung Association*. 2005;171(1073-449; 2):142-146

2 Hogg JC, Chu F, Utokaparch S et al. The nature of small-airway obstruction in chronic obstructive pulmonary disease. *The New England journal of medicine*. 2004;350(1533-4406; 0028-4793; 26):2645-2653

3 van Klaveren RJ, Oudkerk M, Prokop M et al. Management of lung nodules detected by volume CT scanning. *The New England journal of medicine*. 2009;361(23):2221-2229

4 Stolk RP, Rosmalen JG, Postma DS et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *European journal of epidemiology*. 2008;23(1):67-74

5 Volkskrant redactie. Rookverbod in horeca nog niet overal nageleefd. *Volkskrant*. January 16, 2014;(smoking cessation):1

6 Rijks Instituut voor Gezondheid en Milieu, ministerie voor Volksgezondheid, Welzijn en Sport. http://www.nationaalkompas.nl/gezondheid-en-ziekte/ziekten-en-aandoeningen/ademhalingswegen/copd/cijfers-copd-prevalentie-incidentie-en-sterfte-uit-de-vtv-2010/. 2010

7 Stocks J and Sonnappa S. Early life influences on the development of chronic obstructive pulmonary disease. *Therapeutic advances in respiratory disease*. 2013;7(3):161-173

8 Kerkhof M, Boezen HM, Granell R et al. Transient early wheeze and lung function in early childhood associated with chronic obstructive pulmonary disease genes. *The Journal of allergy and clinical immunology*. 2014;133(1):68-76.e4

9 Vestbo J. *Global Strategy for the Diagnosis, Management and Prevention of COPD*, Global . . 2014;2014

10 Vestbo J and De cramer M. *Global Strategy for the Diagnosis, Management and Prevention of COPD*, Global . . 2014;2014

11 Jones PW, Harding G, Berry P, Wiklund I, Chen WH and Kline Leidy N. Development and first validation of the COPD Assessment Test. *The European respiratory journal*. 2009;34(3):648-654

12 van der Molen T, Willemse BW, Schokker S, ten Hacken NH, Postma DS and Juniper EF. Development, validity and responsiveness of the Clinical COPD Questionnaire. *Health and quality of life outcomes*. 2003;1:13

13 Jones PW, Brusselle G, Dal Negro RW et al. Properties of the COPD assessment test in a cross-sectional European study. *The European respiratory journal*. 2011;38(1):29-35

14 Agusti A, Soler JJ, Molina J et al. Is the CAT questionnaire sensitive to changes in health status in patients with severe COPD exacerbations? *Copd*. 2012;9(5):492-498

15 Putcha N, Drummond MB, Connett JE et al. Chronic Productive Cough is Associated with Death in Smokers with Early COPD. *Copd*. 2013

16 Voll-Aanerud M, Eagan TM, Plana E et al. Respiratory symptoms in adults are related to impaired quality of life, regardless of asthma and COPD: results from the European community respiratory health survey. *Health and quality of life outcomes*. 2010;8:107-7525-8-107

17 Voll-Aanerud M, Eagan TM, Wentzel-Larsen T, Gulsvik A and Bakke PS. Respiratory symptoms, COPD severity, and health related quality of life in a general population sample. *Respiratory medicine*. 2008;102(3):399-406

18 Chapman KR, Tashkin DP and Pye DJ. Gender bias in the diagnosis of COPD. *Chest*. 2001;119(6):1691-1695

19 Stahl E, Postma DS, Juniper EF, Svensson K, Mear I and Lofdahl CG. Health-related quality of life in asthma studies. Can we combine data from different countries? *Pulmonary pharmacology & therapeutics*. 2003;16(1):53-59

20 de Jong K, Boezen HM, Kromhout H et al. Occupational exposure to vapors, gases, dusts, and fumes is associated with small airways obstruction. *American journal of respiratory and critical care medicine*. 2014;189(4):487-490

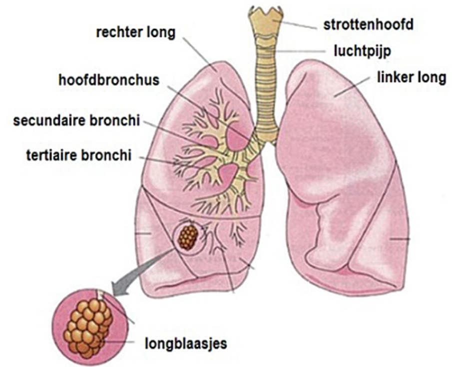
21 Verbanck S, Paiva M, Schuermans D, Hanon S, Vincken W and Van Muylem A. Relationships between the lung clearance index and conductive and acinar ventilation heterogeneity. *Journal of applied physiology (Bethesda, Md.: 1985)*. 2012;112(5):782-790

22 Boudewijn IM, Telenga ED, van der Wiel E et al. Less small airway dysfunction in asymptomatic bronchial hyperresponsiveness than in asthma. *Allergy*. 2013;68(11):1419-1426

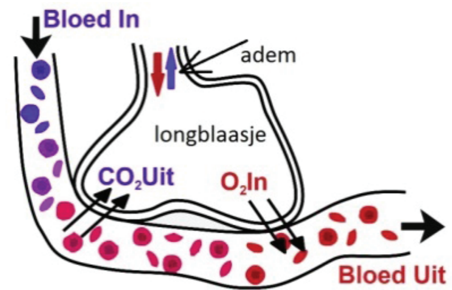
Chapter 9

Nederlandse samenvatting

(Dutch summary)



Figuur 1.
De bouw van een menselijke long

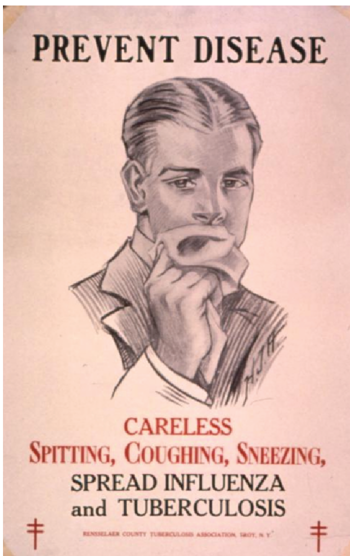


Figuur 2.
Uitwisseling van zuurstof en koolzuur in een longblaasje

Longen zorgen er voor dat we kunnen ademen en dat ons lichaam zuurstof krijgt. De lucht die we inademen stroomt via de mond door de luchtpijp die zich splitst in twee bronchiën, en vervolgens in een systeem van steeds meer, en steeds fijner vertakte pijpjes (bronchi), tot ze uiteindelijk uitmonden in de kleine longblaasjes (Figuur 1). Deze blaasjes worden omringd door een netwerk van kleine bloedvaatjes. De wand van de longblaasjes en van de bloedvaatjes is zo dun dat er gasuitwisseling kan plaats vinden: zuurstof (O2) aanwezig in ingeademde lucht wordt afgegeven aan het bloed en koolzuur (CO2) uit het bloed wordt afgegeven aan de lucht die uitgeademd wordt (Figuur 2).
In de 10.000-20.000 liter lucht die we dagelijks inademen bevinden zich kleine deeltjes, bacteriën en giftige stoffen waaraan onze luchtwegen constant worden blootgesteld. De luchtwegen worden beschermd tegen deze gevaren door een dunne laag slijm waarmee het onderliggende weefsel, het luchtweg epitheel, wordt afgedekt. Gezond slijm heeft een uitgebalanceerde samenstelling en bestaat voor een groot deel uit water aangevuld met stoffen die voor het plakkerige effect en de afweer tegen vreemde stoffen zorgen. Wanneer luchtwegen geïrriteerd worden ontstaat direct extra slijmproductie. Een binnengedrongen vreemd deeltje wordt onmiddellijk "ingepakt" met slijm en vervolgens richting de mond getransporteerd door trilhaartjes die een gecoördineerde beweging richting de uitgang (de mond) maken. Dit transport wordt ondersteund door te hoesten.

Wanneer irritatie van de luchtwegen aanhoudt kan de slijmproductie ontregeld raken. Er wordt bijvoorbeeld te veel slijm geproduceerd of de slijm samenstelling verandert of het wordt niet goed afgevoerd en blijft achter in de luchtwegen. In plaats van de oorspronkelijk beschermende functie zorgt het slijm nu voor problemen; het onderliggende weefsel wordt niet meer beschermd en raakt ontstoken wat gepaard gaat met verdikking van de luchtwegwand en benauwdheid (minder lucht).

Wanneer overproductie van slijm aanwezig is gedurende tenminste 3 maanden per jaar (gedurende de afgelopen 2 jaar) spreekt men van chronische slijm overproductie (ook wel chronische bronchitis genoemd).



In het begin van de 20ste eeuw werd chronische slijm overproductie vooral gezien als symptoom van tuberculose. Een ziekte waaraan destijds veel personen stierven (10.000 in Nederland in 1900). Door de ontdekking van medicijnen tegen tuberculose en verbetering van de sociale omstandigheden (betere huizen en hygiëne) verdween tuberculose en daarmee verdween ook de aandacht voor chronische slijm overproductie.
Na de 2de wereldoorlog nam het aantal personen wat rookte sterk toe en daarmee stak ook chronische slijm overproductie weer de kop op, wat in eerste instantie als een onschuldig bijverschijnsel van roken werd gezien.
Inmiddels is door onderzoek aangetoond dat chronische slijm overproductie niet zo onschuldig is en wel degelijk een gevaar vormt voor de gezondheid. Personen die er last van hebben, hebben een slechtere kwaliteit van leven; ze hebben het vaak benauwd of zijn kortademig, hebben frequenter en langduriger luchtwegontstekingen, worden vaker in een ziekenhuis opgenomen en gaan eerder dood dan personen zonder chronische slijm overproductie.
Het grootste risico om chronische slijm overproductie te ontwikkelen lopen mannen, oudere personen en personen die blootgesteld worden aan irriterende stoffen (o.a. rook) in hun omgeving. Het effect van "zelf" roken is het meest onderzocht maar ook meeroken (blootgesteld worden aan rook door anderen) met name ook tijdens de jeugd, en zelfs het hebben van een rokende moeder voor de geboorte, vergroot het risico op chronisch slijmproductie op volwassen leeftijd. Daarnaast wordt chronische slijm overproductie veel gezien bij personen die werken in de landbouw en in de textiel-, hout-, chemische- en voedselindustrie.

Een andere longziekte die (meestal) wordt veroorzaakt door roken is chronische obstructieve longziekte (COPD). COPD is een complexe ziekte waarvan toenemende benauwdheid het belangrijkste kenmerk is. Bij COPD vernauwen de kleinere luchtwegen door chronische ontsteking en ontstaan er problemen met in- en (vooral) uitademen.

Risicofactoren voor chronische slijm overproductie voor personen met en zonder COPD

Omdat niet iedereen met chronische slijm overproductie ook COPD heeft en niet iedereen met COPD ook chronische slijm overproductie, rijst de vraag of risicofactoren voor chronische slijm overproductie dezelfde zijn voor personen met COPD als voor personen zonder COPD. Om dit te onderzoeken gebruikten we de gegevens van meer dan 10.000 personen afkomstig uit de Noord Nederlandse LifeLines populatie. Om de mate van blootstelling aan sigarettenrook vast te stellen gebruikten we het antwoord op de vraag of men ooit gerookt had of nog rookte, en zo ja, hoeveel sigaretten men dan rookt(e) gedurende een bepaalde tijd. De eenheid die hiervoor gebruikt wordt is“pakjaar” (packyear in het Engels): 1 pakjaar is gelijk aan het dagelijks roken van 20 sigaretten gedurende 1 jaar. Dagelijks 20 sigaretten roken gedurende 20 jaar is dan 20 pakjaren, net als 40 jaar lang 10 sigaretten per dag. Om vast te stellen of personen ook blootgesteld werden aan irriterende stoffen vanwege hun werk, werd in een tabel opgezocht welke blootstelling bij een bepaald beroep hoort. Om te beoordelen of iemand COPD had werd een longfunctietest gedaan; de hoeveel lucht die iemand in 1 seconde kan uitademen (FEV_1) en de totale hoeveelheid lucht die kan worden uitgeademd wordt gemeten (FVC). Is het quotiënt van deze waarden (FEV_1/FVC) kleiner dan 0.70 dan heeft iemand COPD.

Risicofactoren voor chronische slijmproductie	Geen COPD	Wel COPD
Manlijk geslacht	+	-
BMI	+	-
Sigarettenrook		
Rook van anderen	-	+
Ex roken	-	-
Huidig roken	+	-
Pakjaren	+	+
Beroepsmatige blootstelling		
Biologische stoffen	-	-
Mineralen	+	-
Gassen & dampen	+	-
Pesticiden	-	-
Aromatische stoffen	+	-
Chloor	+	-
Zware metalen	+	-

Uit ons onderzoek blijkt dat, zowel voor personen met als zonder COPD, het aantal pakjaren dat men rook(te) het grootste risico vormt op chronische slijm overproductie; hoe hoger het aantal pakjaren hoe groter het risico. Echter, beroepsmatige blootstelling

aan gassen en dampen (gases & fumes) en in mindere mate aan mineralen, aromatische stoffen, chloor en zware metalen blijkt wel een risico te zijn voor personen zonder maar niet voor personen met COPD (Tabel 1). Van blootstelling aan gassen & dampen is sprake bij beroepen als lasser, loodgieter, schilder en vrachtwagenchauffeur.
Risicofactoren voor chronische slijm overproductie voor personen met en zonder COPD verschillen dus. Vragen die hier uit voorkomen zijn: is de samenstelling van het slijm en de plaats waar het slijm in de longen geproduceerd wordt verschillend tussen personen met en zonder COPD en zijn de betrokken genen ook verschillend. Dit moet verder onderzocht worden.

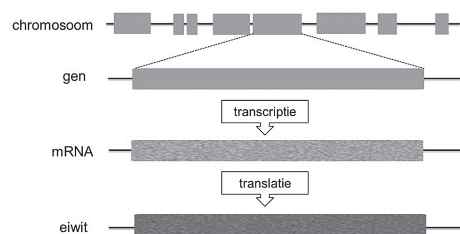
Genetische verschillen tussen rokers

Wanneer rokers worden vergeleken die ongeveer even oud zijn en evenveel gerookt hebben, blijken er grote verschillen te bestaan: er zijn rokers die nergens last van hebben, anderen hebben alleen last van chronische slijm overproductie, weer anderen alleen van COPD, maar er zijn ook rokers die last hebben van zowel chronische slijm overproductie als van COPD.
Ondanks dezelfde "rook" geschiedenis hebben ze dus niet allemaal dezelfde problemen. Daarom wordt er gedacht dat erfelijke aanleg voor het ontwikkelen van deze ziekten een rol speelt. Aanwijzingen hiervoor werden gevonden door aan te tonen dat in bepaalde families chronische slijm overproductie vaker voorkomt dan in andere families, terwijl de leef- en rook gewoonten van de onderzochte families vergelijkbaar waren. Hetzelfde geldt voor COPD.

Genetica

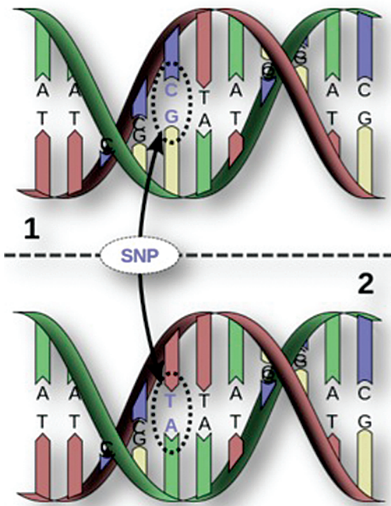
Elke lichaamscel bevat chromosomen, DNA moleculen, waarin zich dezelfde erfelijke informatie bevindt. Een afgebakend stuk DNA van een chromosoom heet een gen en heeft een vaste plaats op een chromosoom. Genexpressie is het proces waarbij informatie in een gen "tot uiting komt" en bestaat uit twee belangrijke stappen:

1) Transcriptie: het overschrijven van DNA in mRNA, en
2) Translatie; het vertalen van mRNA in eiwit (Figuur 3).



Figuur 3. Genexpressie is het proces waarbij informatie in een gen "tot expressie komt" doordat het gen wordt overschrijven van gen naar mRNA en vertalen van mRNA in eiwit.

DNA is opgebouwd uit 4 verschillende basen die paren vormen: adenine (A) met thymine (T) en guanine (G) met cytosine (C). De volgorde van basen in een gen bevat de code voor een eiwit. Eiwitten vervullen allerlei functies binnen en buiten de cel zoals afweer, bouwsteen en enzym. Menselijk DNA bevat ca. 1200.000.000 (1200 miljoen) basenparen waarvan slechts 1% verschilt tussen personen. Deze 1% bepaalt de variatie tussen personen: zoals het uiterlijk, en voor welke ziekte(n) iemand gevoelig is. Een basenpaar wat in >1% van de bevolking verschilt heet een polymorfisme of “Single Nucleotide Polymorphism”, kortweg SNP (snip) (Figuur 4). Door de volgorde van de basenparen van personen met en zonder een bepaalde ziekte te vergelijken kunnen aanwijzingen gevonden worden voor de betrokkenheid van genen bij het ontstaan van deze ziekte. Dit type onderzoek heet een genoom brede associatie studie (GWAs); alle erfelijke informatie (het hele genoom) van gezonde en zieke personen wordt hierbij op heel veel plaatsen vergeleken.

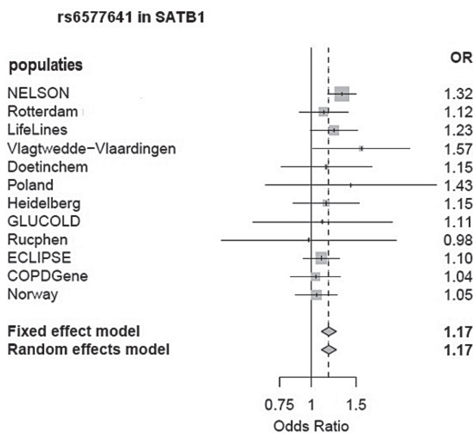


Figuur 4. Een basenpaar wat in >1% van de bevolking verschilt heet een polymorfisme of “Single Nucleotide Polymorphism” (SNP).

In dit proefschrift staan drie genoom brede associatie studies beschreven welke zijn uitgevoerd in de populatie van de NELSON-studie; een studie die in eerste instantie is opgezet om te onderzoeken of het opsporen van longkanker in een vroeg stadium op een driedimensionale foto (CT-scan), en het daardoor eerder kunnen starten met behandelen, (positieve) gevolgen heeft voor het verloop van de ziekte. De NELSON-studie bevat gegevens van ca. 3000 rokers en ex-rokers personen (voornamelijk mannen) die veel (minstens 20 pakjaren) hebben gerookt en daardoor een verhoogd risico op longkanker hebben maar ook een verhoogd risico op chronische slijm overproductie en/of COPD. Van deze personen werd op meer dan 600.000 plaatsen op het DNA vastgesteld welk basenpaar aanwezig was.

In de eerste studie werd het DNA van personen met en zonder chronische slijm overproductie vergeleken. De SNPs die sterk verbonden bleken met chronische slijm overproductie in de NELSON-studie werden daarna getest in 11 andere populaties. Een SNP, rs6577641, op chromosoom 3 in het *special AT-rich sequence-binding protein-1* gen (*SATB1*) liet in elke populatie het zelfde effect zien; wanneer het basenpaar hier G-C was in plaats van T-A was het risico om chronische slijm overproductie te ontwikkelen gemiddeld 17% hoger (Figuur 5). Uit de literatuur is bekend dat expressie van *SATB1* ook werd

aangetoond in luchtwegepitheel en dat *SATB1* een transcriptie factor is die de expressie van veel genen beïnvloedt. Het is voorstelbaar dat dit gen een rol speelt bij de vorming van slijm producerende cellen.



Figuur 5. Een zogenaamde "forest plot" laat de associatie zien van chronische slijm overproductie met rs6577641 in SATB1 op chromosoom 3 in elke populatie die is onderzocht is.

In de tweede studie voerden we twee genoom brede associatie studies uit om te kunnen vergelijken of de genen die betrokken zijn bij chronische slijm overproductie verschilden tussen mensen met en zonder COPD.

De bevindingen in NELSON personen met COPD werden gecontroleerd in 4 andere populaties. Een SNP (rs10461985) op chromosoom 5 in het *glial cell line-derived neurotrophic factor antisense RNA 1* gen (*GDNF-AS1)* bleek betrokken te zijn bij chronische slijm overproductie bij personen met COPD. In de literatuur vonden we dat dit gen de expressie van een ander gen (*glial cell line-derived neurotrophic factor, GDNF*) kon verminderen/blokkeren (Figuur 6).

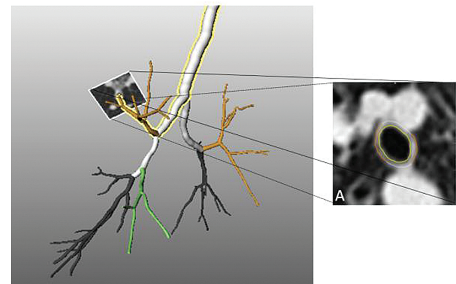
Figuur 6. GDNF-AS1 beïnvloedt de transcriptie van GDNF

Dit controleerden we in longweefsel van COPD patiënten: het bleek dat mRNA-expressie van *GDNF* in longweefsel inderdaad lager was bij personen met chronische slijm overproductie dan bij personen zonder chronische slijm overproductie. Het is aannemelijk dat dit wordt veroorzaakt door de aanwezigheid van de SNP in het *GDNF-AS1* gen. De productie van *GDNF* vermindert waardoor meer slijmproductie ontstaat.

Bij personen zonder COPD werden 4 SNPs gevonden die betrokken waren bij chronische slijm overproductie waaronder een SNP (rs4863687) op chromosoom in het *mastermind-like 3* gen (*MAML3*). Een aanvullend analyse in longweefsel van bijna 1100 personen liet zien dat de mRNA expressie van *MAML3* sterk verhoogd was bij personen met deze SNP vergeleken met personen zonder deze SNP. Uit de literatuur is bekend dat *MAML3* samenwerkt met β-catenin and NF-κB , beiden genen waarvan bekend is dat ze bij ontsteking van de longen betrokken zijn. Een rol voor *MAML3* bij ontstekingsprocessen in de longen lijkt aannemelijk maar welke rol het precies speelt, met name bij chronische slijm overproductie, moet uit vervolgonderzoek blijken.
Uit vergelijking van SNPs die sterk betrokken waren bij chronische slijm overproductie bij personen met COPD en SNPs sterk betrokken bij personen zonder COPD kwam slecht 1 SNP naar voren. Analyse van deze SNP in andere populaties kon betrokkenheid van deze SNP bij chronische slijm overproductie niet bevestigen.
De resultaten van deze studie suggereren sterk dat er genetische verschillen bestaan tussen personen met en zonder COPD met betrekking tot chronische slijm overproductie. Daarom is het de moeite waard maken om uitgebreider onderzoek te doen.
Echter, in het algemeen geldt dat hoe groter de groep personen is waarin iets onderzocht wordt, des te betrouwbaarder de resultaten zijn. Omdat de NELSON populatie voor dit onderzoek werd verdeeld in 2 groepen (personen met COPD en personen zonder COPD) leveren de resultaten van dit onderzoek een minder sterk bewijs.

Luchtweg wanddikte

Vroeger kon de dikte van de luchtweg wand alleen gemeten worden wanneer een (deel van een) long verwijderd was uit het lichaam. Tegenwoordig kan met behulp van een driedimensionale longfoto (CT-scan) en een computerprogramma een dwarsdoorsnede van een luchtweg worden gemaakt en de wanddikte worden gemeten (Figuur 7). Een luchtwegwand geeft een andere grijstint op een CT-scan dan bot, lucht of andere weefsel. Uit deze kleurverschillen is de wanddikte af te leiden.



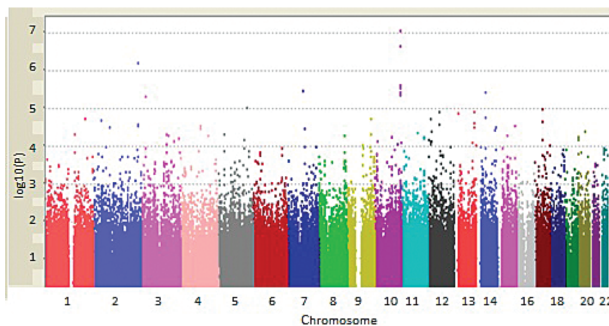
Figuur 7.
Dwarsdoorsnede van een luchtweg op een CT gemaakt m.b.v. software

De vraag die wij wilden beantwoorden is hoe de dikte van de luchtwegwand zich verhoudt tot longfunctie (mate van benauwdheid), rookgewoonten, respiratoire symptomen (hoesten, chronische slijm overproductie) en emfyseem. Emfyseem is een kenmerk van

COPD en wordt (meestal) ook veroorzaakt door roken. Bij personen met emfyseem is de elasticiteit van de longen verminderd waardoor de inhoud van de longen is vergroot en het uitademen bemoeilijkt wordt en benauwdheid ontstaat. De hoeveelheid emfyseem is ook meetbaar op een CT-scan.

Voor dit onderzoek gebruikten we weer de NELSON populatie; alle deelnemers rookten erg veel of hadden veel gerookt in het verleden. De CT-scans die van de NELSON-studie populatie zijn gemaakt zijn "low-dose CT-scans", d.w.z. dat de CT-scans met een lage dosis röntgenstraling zijn gemaakt wat minder stralingsbelasting geeft dan de gebruikelijke hoge dosis. We stelden in een vooronderzoek vast dat de luchtwegwanddikte gemeten op de lage dosis CT-scans vergelijkbaar is met de luchtwegwanddikte op hoge dosis CT-scans, en dat de wanddikte optimaal meetbaar is in luchtwegen met een interne diameter van 3,5 mm.
Onderzoek in een groep van 500 personen liet zien dat een dikkere luchtwegwand sterk geassocieerd is met meer pakjaren en met een slechtere longfunctie, maar niet met de aanwezigheid van chronische slijm overproductie of hoesten. Bovendien bleek dat de dikte van luchtwegwand meer van invloed was op de longfunctie dan de mate van emfyseem.

Om te onderzoeken welke genen betrokken zijn bij de dikte van de luchtwegwand voerden we een genoom brede associatie studie uit in meer dan 2600 personen uit de NELSON studie. Van deze personen werd de luchtweg wanddikte gemeten op een lage dosis CT-scan.



Figuur 8. Manhattan plot: op de x-as staan de chromosomen; elk puntje is een SNP; een SNP die (sterk) geassocieerd is met luchtwegwand verdikking staat ver verwijderd van de x-as.

Figuur 8 laat een "een Manhattan plot" zien: op de x-as verbeeldt elke kleur een chromosoom en is elk puntje een geteste SNP (totaal 520.00 SNPs). Des te verder een SNP verwijderd is van de x-as, des meer invloed deze SNP heeft op de dikte van de luchtweg wand. Bij chromosoom 2 en 10 zijn SNPs te zien die sterk geassocieerd zijn met luchtweg wanddikte. We herhaalden onze bevindingen in een vergelijkbare Duitse populatie zware rokers (> 20 pakjaren). DNA-varianten in het *membrane associated guanylate kinase WW and PDZ domain containing 2* (*MAGI2*) gen op chromosoom 7, in het 5′-nucleotidase cytosolic IIIB (*NT5C3B*) gen op chromosoom 17, dichtbij het *serpin peptidase inhibitor, clade E member 2* (*SERPINE2*) gen op chromosoom 2 en tussen het *Chromosome 10 open reading frame 90* (*C10orf90*) en het *Dedicated Of Cytokinesis* gen (*DOCK1*) op chromosoom 10 waren geassocieerd met luchtwegwanddikte. Aanvullend onderzoek liet zien dat *MAGI2* geassocieerd is met luchtwegwand ontsteking en uit de literatuur blijkt dat *SERPINE2* geassocieerd is met emfyseem en de SNP op chromosoom 10 eerder in verband is gebracht met benauwdheid, allemaal kenmerken van COPD.

Een belangrijk doel van medisch onderzoek is het verbeteren van de gezondheid van mensen, rekening houdend met de verschillen tussen mensen als gevolg van omgevings- en genetische factoren. Deze factoren, en de samenwerking tussen deze factoren, kunnen er toe leiden dat mensen meer of minder gevoelig zijn voor het ontwikkelen van (luchtweg) ziektes zoals chronische slijm overproductie en luchtweg wand verdikking. Dit proefschrift beschrijft onderzoek naar deze factoren maar roept ook weer vragen op die mogelijk in vervolgonderzoek kunnen worden beantwoord zoals:

- Hoe werken de genen die in verband zijn gebracht met chronische slijm overproductie en lucht
- Is slijm bij mensen me sen zonder COPD gelet op samenstelling
- Is het slijm veroorzaak le plaats in de luchtwegen als slijm v otstelling aan roken & dampen?

Buiten kijf staat dat roken ls risicofactor.

Dankwoord

Life is what happens to you while you're busy making other plans
John Lennon

Dankwoord

Hoe je leven verloopt is, net als het wel of niet krijgen van een (long)ziekte, afhankelijk van erfelijke eigenschappen en de omgeving waar je, vanaf het prilste begin van je bestaan, aan bloot staat. Je omgeving beïnvloedt de bewuste en onbewuste afwegingen die je maakt, wat zich uit in gedrag en keuzes. Mensen - zowel in je directe omgeving als verder in de wereld - aan wiens gedrag je je spiegelt, brengen je keuzes en gedrag in beweging. Het geheel is een dynamisch proces; je levenspad.

Wat heeft dat nou te maken met een dankwoord in een proefschrift? In mijn beleving alles!
Zie het schrijven van een proefschrift ook als een pad waar je langs loopt. Wanneer je de juiste dingen mee neemt of ontvangt voor onderweg en je mensen tegenkomt die je de weg wijzen, bijsturen en aanmoedigen, dan helpt dat om je doel te halen. Kies je een verkeerde afslag of stuit je op een afgesloten weg dan wordt het lastiger, en ben je nog afhankelijker van anderen.

Het vertrekpunt van mijn pad ligt bij mijn ouders, broer en zus, die me m'n bagage voor onderweg mee gaven. Lieve mama en papa - die er helaas al lang niet meer is - jullie waren de basis, heel veel dank hiervoor! Koos en Geke, nu we "groot" zijn weet ik hoe belangrijk we voor elkaar zijn. Zussen kunnen ook hartsvriendinnen zijn. Dank!

Ook de mensen in het buurtje waar ik opgroeide, op de lagere school en op de middelbare school drukten hun stempel op me. Soms kom ik hen tegen op begrafenissen, delen we herinneringen en maken we plannen om een reünie te houden, maar dat komt er (nog) niet van. Alleen Linda zie ik nog geregeld en dan pakken we de draad zo weer op en dat is altijd weer fijn. Bedankt!

Aan de jaren als analist bij 'laboratorium kindergeneeskunde' bewaar ik goede herinneringen, omdat we daar als collega's fijn samenwerkten. Daar heb ik m'n hartsvriendinnen Marga en Sjaan leren kennen (op alfabet). Altijd kan ik bij jullie terecht. Marga, geweldig dat je m'n paranimf wilt zijn. Sjaan, ons leven loopt zo parallel en we kunnen zo veel delen. Dank!

De vader van mijn kinderen ontmoette ik ook in deze periode. We hadden goede jaren maar helaas redden we het toch niet samen. Het mooiste in mijn leven is wel uit deze relatie voortgekomen: Jaap, Roos en Bram. Met vallen en opstaan zijn jullie goede, zelfstandige en weerbare mensen geworden waar ik verschrikkelijk veel van houdt en heel trots op ben. Lieve schatten, bedankt!

Na veel overwegingen nam ik een afslag en ging gezondheidswetenschappen studeren. Een keuze waar ik geen moment spijt van heb gehad. Er werd een honger gestild.

Dat dit ingeslagen pad zou eindigen in een promotie was toen nog heel, heel ver weg. Mijn promotietraject verliep klassiek: een rustige start waarna de druk langzaam maar gestaag op liep. Gedurende dit traject verzuchtte ik vaak "komt het ooit zover dat ik een dankwoord mag schrijven?" Ja dus, en daarvoor wil ik verschillende mensen bedanken.

Allereerst mijn (co)promotoren. Dirkje, altijd gedreven en enthousiasmerend, snel en doelgericht. Mens, wat heb ik veel van je geleerd! Harry, steeds scherp en nauwkeurig formulerend. Als je tijd had, dan schoof je alles aan de kant, dat heb ik altijd in je bewondert. Marike, je woorden "maar Akkelies, dat weet je toch wel" hebben me, na eerst vertwijfeling, steeds weer diep doen graven. Judith, drempelloos toegankelijk en altijd precies en geduldig. Ik had je vaak nodig en kon altijd bij je terecht. Bedankt!

Additionally I would like to thank professors Vestbo, Brusselle and Kerstjens for the effort of judging this thesis; I feel honored by your interest.

Dit proefschrift zou er niet zijn geweest zonder deelnemers aan de NELSON en de LifeLines studie en de grote inzet van medewerkers die hebben bijgedragen aan de dataverzameling.

Sietske, Trudy, Heleen, Stephanie en eerder Gonny en Renee, ook jullie waren onmisbaar en ik wil jullie bedanken voor alles wat jullie voor me opzochten, regelden of planden.

GRIAC-ers bedankt voor alle leerzame bijeenkomsten, presentaties, input en discussies.

Sinds een jaar ben ik weer werkzaam bij de afdeling Laboratoriumgeneeskunde. Ik wil mijn collega's daar bedanken voor hun medeleven en voor de vrijheid die ik kreeg om dit proefschrift af te ronden.

En last but not least, mijn directe collega's van longziekten: Maartje, Jorine, Susan, Erica, Ilse, Fransien, Anda, Grietje, Wytske, Eef, Karin en Ruth, met jullie heb ik veel gedeeld. Tranen van het lachen maar ook van frustratie en woede. Geregeld voelde ik me de kinderachtigste van de club. Tips en trucs met betrekking tot statistiek en computerprogramma's kreeg ik van jullie. De grabbelton voor dieptepunten, taart of koek bij zelfs de kleinste vreugdevolle gelegenheid en een kroket op vrijdag olieden ons onderzoekersbestaan. Gouden tijden! Bedankt.

Ik ga weer verder. Het is niet mogelijk om alle mensen die mijn levenspad kruisten en me richting gaven te noemen en te bedanken. Het risico dat ik iemand te kort doe door het toch te proberen is te groot, daarom: lieve mensen, jullie weten het wel, zonder jullie was het niets geworden. Bedankt!

Nu ga ik
* een cursus sterrenkunde doen
* breien op de bank
* het Pieterpad lopen
* een moestuin aanleggen
* vrijwilliger worden bij Amnesty International
* vogels tellen op de dijk
* elke dag tennissen
* vrijdags altijd wijn drinken met Geke

Annelies