Impact of gene-environment interactions within inflammatory and oxidative stress pathways on the development of chronic obstructive lung disease (COPD)

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät der
Universität Basel

von

Ivan Curjuric

aus Zürich (ZH) und Kroatien

Basel, 2012

creative commons

Namensnennung-Keine kommerzielle Nutzung-Keine Bearbeitung 2.5 Schweiz

---

Sie dürfen:

das Werk vervielfältigen, verbreiten und öffentlich zugänglich machen

Zu den folgenden Bedingungen:

Namensnennung. Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen (wodurch aber nicht der Eindruck entstehen darf, Sie oder die Nutzung des Werkes durch Sie würden entlohnt).

Keine kommerzielle Nutzung. Dieses Werk darf nicht für kommerzielle Zwecke verwendet werden.

Keine Bearbeitung. Dieses Werk darf nicht bearbeitet oder in anderer Weise verändert werden.

- Im Falle einer Verbreitung müssen Sie anderen die Lizenzbedingungen, unter welche dieses Werk fällt, mitteilen. Am Einfachsten ist es, einen Link auf diese Seite einzubinden.
- Jede der vorgenannten Bedingungen kann aufgehoben werden, sofern Sie die Einwilligung des Rechteinhabers dazu erhalten.
- Diese Lizenz lässt die Urheberpersönlichkeitsrechte unberührt.

Quelle: http://creativecommons.org/licenses/by-nc-nd/2.5/ch/ Datum: 3.4.2009

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Nicole M Probst-Hensch	Dissertationsleitung
Prof. Marcel Tanner	Fakultätsvertreter
Prof. Peter Vollenweider, Universität Lausanne	Korreferent

Basel, den 15. November 2011

Prof. Dr. Martin Spiess

Dekan,
Philosophisch-naturwissenschaftliche
Fakultät, Universität Basel

Table of Contents

Acknowledgements

I owe my gratitude to many persons for successfully conducting and finishing this thesis work.

First of all, I would like to thank my thesis supervisor Prof. Nicole M. Probst-Hensch, who has constantly supported me throughout these years and has given key inputs whenever it was necessary to take the next step. Nicole, thank you for the rewarding collaboration during my thesis work. I very much enjoyed it, and particularly appreciated the excellent balance you keep between providing us the necessary freedom for developing new ideas, and setting the necessary limits or pushing things forward when needed. This has substantially helped to develop new skills and made my thesis work a rewarding experience. Thank you also for keeping an open ear and eye to the things that happen outside the working environment.

A large thank you also goes to Medea Imboden together with Nicole, for introducing me to the do's and don'ts of studying genes in the context of chronic, complex disease, for their tireless inputs towards my project specific analyses, and the scientific guidance during my thesis work. Their long standing highly professional engagement laid the foundation for my thesis, and continuously fuelled my motivation to pursue the determinants of chronic disease during the work. I would further like to thank Martin Adam and Gian Andri Thun for the ongoing mutual exchange of tips and tricks making an epidemiologist's everyday life easier, and their cheerful company which helped to overcome sometimes disappointing analysis results.

These people all belong to the same team whose diverse but well –tempered composition make up a very enjoyable and stimulating working environment. I am very happy and grateful to have the opportunity to be a part of it.

I also thank the members of my thesis committee: professors Hans-Peter Beck (SwissTPH and University of Basel), Wolfgang Berger (University of Zürich), Charlotte Braun-Fährländer (SwissTPH and University of Basel), Dirkje S Postma (University of Groningen, The Netherlands), Thierry Rochat (University Hospital and University of Geneva), and Erich W. Russi (University Hospital and University of Zürich). They supported my work by assessing the initial thesis year and providing input to project-specific work. At the same time, I would also like to thank Prof. Peter Vollenweider (University Hospital and University of Lausanne) and Prof. Marcel Tanner (SwissTPH and University of Basel) who stepped in as co-referee and faculty representative during the final phase of my thesis, when the merger of our previous institutions required changes in the committee composition.

Particular thanks also go to Prof. Christian Schindler who, although not being part of my thesis committee, gave me a solid foundation in applied statistical analysis from which I will continue to benefit in my professional life. Remembering the time when I was in the midst of clinical training and thinking back to the circumstances making me switch from clinics to epidemiological research, I also want to thank Prof. Ursula Ackermann-Liebrich, former director of the Institute of Social and Preventive Medicine Basel for wakening my interest in epidemiology and Public Health. Finally, I'm also grateful to all scientific members of the SAPALDIA study who gave their continuous inputs to my work, and to the participants who provided the data for the study.

My thesis work was supported by the ProDoc programme of the Swiss National Science Foundation (SNF) in the framework of the “SSPH+ PhD program Public Health” from the Swiss School of Public Health Plus (SSPH+). I would thus like to thank the SNF for supporting my thesis work and Prof. Charlotte Braun-Fahrländer for the student support and educational programme provided by the SSPH+, as well as her initial guidance in administrative matters.

But my professional life would not have been half as enriching without my family. First of all I would like to thank my beloved wife: Nives, thank you for your support during these years of my studies, for keeping me on the grounds of normal everyday life, and of course for managing everything at home in times of my absence during courses, meeting and conferences. Thank you for going through these sometimes difficult years together. One of the most intense courses I attended in the last years and still keep attending time-management taught by my wonderful daughter Cvijeta and son Marin. By rightfully seeking my attention and company every day, sometimes also challenging me, they opened up a new dimension in my life and made it a richer and much deeper experience. Their smile often freshened my mind, and freed it up for new ideas. Finally, I’m also very grateful to my parents and parents-in-law for helping out during critical bottlenecks in our schedules, as well as for providing a safe haven to rest and refill the batteries throughout the year.

I want to conclude by expressing my hope that the parts of my professional life to follow will be as diverse and interesting as my thesis experience was.

Zusammenfassung

Einleitung

Chronisch obstruktive Lungenerkrankung (COPD) ist für einen bedeutenden Anteil der chronischen Erkrankungen, Gesamtsterblichkeit und Gesundheitskosten verantwortlich und ist deswegen von grosser Bedeutung für die öffentliche Gesundheit. Die Prävalenz-Schätzwerte internationaler Studien driften jedoch aufgrund unterschiedlicher Krankheits-Definitionen weit auseinander. Trotz ihrer grossen Bedeutung wird die Erkrankung in der Regel erst in späteren Stadien diagnostiziert, in denen keine effektive Therapie verfügbar ist.

Die COPD Diagnose wird dabei vornehmlich auf die Lungenfunktionsmessung abgestellt, unter Berücksichtigung weiterer Faktoren wie zum Bespiel hoher Tabakkonsum. Um die klinische Früherkennung zu fördern sowie vergleichbare Daten in der Forschung zu ermöglichen schlug die „global initiative for obstructive lung diseases“ (GOLD) eine einfache Definition vor. Diese beinhaltet, dass bei COPD das Verhältnis des Volumens, das bei der Spirometrie in der ersten Sekunde ausgeatmet wird, über das gesamte Ausatmungsvolumen (FEV_1/FVC) kleiner als 0.7 ist. Die Spirometrie-Werte müssen dabei nach pharmakologischer Atemwegs-Dilatation gemessen werden, um COPD von undiagnostiziertem Asthma differenzieren zu können. Die Definition wurde zunächst weitläufig angewendet, einschliesslich epidemiologischer Studien, kam aber in der Folge unter starke Kritik, da sie eine Überdiagnose bei älteren Probanden begünstigt (der FEV_1/FVC Wert von 0.7 kann durch die Lungenalterung zum Teil bereits mit 45 Jahren erreicht werden). Alternativ kommt vermehrt das Kriterium FEV_1/FVC < lower limit of normal (LLN) zum Einsatz. Das LLN ist als fünftes Perzentil der FEV_1/FVC Verteilung bei gesunden Nichtrauchern definiert. Wenn möglich sollten COPD Studien heute beide Definitionen einschliessen.

Die der COPD Erkrankung zugrunde liegenden Krankheitsprozesse sind immer noch unklar. Rauchen wurde lange Zeit als entscheidender Risikofaktor angesehen, aber neuere Studien deuten auch auf Effekte durch Berufsexpositionen, Verwendung von biologischen Materialien (wie Kohle) zum Heizen und Kochen, Passivrauchen sowie möglicherweise Luftverschmutzung hin. Zwei biologische Reaktionswege wurden als wahrscheinliche Krankheitsprozesse definiert. Erstens könnte ein Ungleichgewicht zwischen Proteasen und Antiproteasen im Körper zur Schädigung von Lungengewebe und den für COPD typischen Veränderungen führen. Zweitens könnte ein Ungleichgewicht zwischen Oxidantien und Antioxidantien zu oxidativem Stress führen, unter welchem freie Radikale (Oxidantien) mit normalen Zell-Strukturen wie Proteinen, Zellwänden und DNA chemische Reaktionen eingehen und Zell-Schädigungen auslösen. Dies löst Entzündungsreaktionen aus, welche oxidativen Stress weiter begünstigen. Kandidaten-Gen Studien konnten Gene aus den beiden Systemen in Verbindung mit COPD bringen, aber oftmals wurden diese Assoziationen in Folgestudien nicht repliziert. Grund dafür könnte die Nichtbeachtung von Gen-Umwelt-Interaktionen sein.

Die meisten der zitierten Umwelt-Risikofaktoren induzieren oxidativen Stress im Zielgewebe, doch das Ausmass der Schädigung hängt von der individuellen Empfindlichkeit ab. Diese wird durch Varianten in den zugrundeliegenden Genen und Reaktionswegen mitbeeinflusst. Die Identifizierung von Empfindlichkeits-Faktoren muss durch die Untersuchung von Gen-Umwelt-Interaktionen erfolgen.

Es wurde nachgewiesen, dass die Schwebestaub-Komponente von Luftverschmutzung oxidativen Stress induziert. Epidemiologische Studien haben bedeutende und konsistente Beziehungen zwischen Schwebestaub-Exposition und Mortalität, Hospitalisierungen aufgrund kardiopulmonaler Beschwerden, respiratorischer Symptome sowie beschleunigter Lungenalterung gezeigt. Daher ist Luftverschmutzung einer der wichtigsten, heutigen Umwelt-Risikofaktoren und muss als Risikofaktor für COPD untersucht werden.

Das Ziel dieser Dissertation war es, die Inzidenz von COPD in der Schweiz zu bestimmen anhand verfügbarer bevölkerungsbezogener Daten. In diesem Kontext soll die Nützlichkeit epidemiologischer Daten abgeschätzt werden. Weiter soll die Rolle von Luftverschmutzung in der Entstehung von COPD untersucht werden, mit Fokus auf die mögliche Interaktion mit „Oxidative Stress" Genen. Schliesslich soll der geschätzte Beitrag von Luftverschmutzung zur COPD-Entstehung auf Populationsebene mit jenem von Rauchen, dem wichtigsten bekannten Risikofaktor, verglichen werden in Bezug auf involvierte Gene und Reaktionswege.

Methoden

Die Arbeiten der vorliegenden Dissertation basierten auf den Daten der Schweizer Studie zur Luftverschmutzung und Lungenerkrankungen bei Erwachsenen (SAPALDIA). Dies ist eine bevölkerungsbezogene Kohortenstudie, welche im Jahr 1991 gestartet wurde mit dem Ziel, die Gesundheitseffekte von Langzeit-Expositionen gegenüber Luftverschmutzung auf Lungen- und Herzerkrankungen sowie Allergien zu untersuchen. Die erste Folgeuntersuchung wurde 2002 durchgeführt, eine zweite wurde 2010/11 abgeschlossen. Die jetzigen Arbeiten beruhen auf den Daten der Basis- und ersten Folgeuntersuchung, an denen 9651 respektive 8047 Personen teilnahmen. Teilnehmer beantworteten einen detaillierten Gesundheits-Fragebogen mit Fragen zu Rauch- und anderen Lebensgewohnheiten, beruflichen Expositionen und vorbestehenden Erkrankungen. Die Lungenfunktion wurde mit denselben Testgeräten bei beiden Erhebungen und in standardisierter Weise getestet. Die Messungen wurden strikten Qualitätskontrollen unterworfen. Dabei fand keine Atemwegs-Dilatation statt. Schätzer der individuellen Luftverschmutzungs-Exposition waren für die Schwebestaubfraktion mit medianem Durchmesser unter 10µm (PM_{10}) verfügbar, parametrisiert als Expositions-Veränderung oder kumulative Belastung während des Follow-ups. Die Schätzer beruhten auf einem Gauss'schen Dispersionsmodel nationaler Emissionsdaten für die Jahre 1990 und 2000, mit Interpolation der Werte für die Zwischenjahre anhand historischer Trends bei Luftmessstationen. Bei der Folgeuntersuchung 2002 wurden auch Blutproben abgenommen, wodurch DNA-Proben von über 6000 Personen für Untersuchungen von Kandidaten-Genen verfügbar waren. Durch die Zusammenarbeit im internationalen GABRIEL Konsortium wurden genomweite Daten für 1457 Personen gewonnen, welche alle Asthmatiker und eine Zufallsstichprobe von Nichtasthmatikern umfassen.

Die Art der verfügbaren Daten implizierte folgende Entscheidungen betreffend des Analysedesigns: Das Fehlen von Spirometrie-Messungen nach Atemwegs-Dilatation machte die Untersuchung einer modifizierten GOLD COPD-Definition nötig. GOLD-Kriterien wurde dabei auf prä-dilatatorische Werte angewendet. Alternativ wurde die longitudinale Abnahme der Lungenfunktion als Proxy-Mass für COPD Entwicklung untersucht. Zweitens konnten lediglich die Schätzer für PM_{10}-Exposition als valides Mass der individuellen Belastung verwendet werden, da die PM_{10}-Fraktion der Luftverschmutzung sich räumlich homogener verteilt als andere Komponenten. Drittens musste die Untersuchung von involvierten Reaktionswegen auf die relativ kleine Gruppe von Nicht-Asthmatikern mit verfügbaren genomweiten Daten beschränkt werden.

Resultate

In der ersten Arbeit beobachteten wir mit der modifizierten GOLD-Definition eine Inzidenz von 14.2 Fällen/1000 Personenjahre (PJ). Dies liegt am oberen Ende von publizierten Schätzwerten aus vergleichbaren bevölkerungsbasierten Studien, und konnte nur partiell mit unterschiedlicher Altersverteilung, Tabakexposition und Dauer des Follow-up erklärt werden. Während ein positiver Zusammenhang zwischen Inzidenz und Alter sowie Rauchen konsistent in der Literatur beschrieben ist, beobachteten wir einen solchen auch für chronische Bronchitis. Die LLN-Definition lieferte erwartungsgemäss eine kleinere Inzidenz von 7.2 Fällen/1000 PJ. 20.9% der obstruktiven Fälle bei der Basisuntersuchung zeigten normale Spirometriewerte in der Folge. Eine Progression milder Atemwegs-Obstruktion zu moderaten bis schweren Stadien während des Follow-up war assoziiert mit höheren Raten von Atemnot und Arztbesuchen, während persistent milde Obstruktion nicht damit assoziiert war. Fälle von milder Obstruktion, die nicht persistierten, waren assoziiert mit mehr Arztbesuchen, zeigten im Durchschnitt aber normale FEV_1 und FVC Werte. Dies könnte durch nicht-diagnostiziertes Asthma bedingt sein. Die Schlussfolgerung dieser Arbeit war, dass wiederholte Lungenfunktionsmessungen ohne Atemwegs-Dilatation künftige gesundheitliche Beeinträchtigungen vorhersagen können, jedoch nicht gut zwischen Asthma und COPD differenzieren. Der Einbezug weiterer klinischer Parameter könnte helfen, Fälle von milder Obstruktion mit hohem Progressionsrisiko zu charakterisieren.

In der zweiten Arbeit untersuchten wir, ob Varianten in den Genen *Hämoxygenase-1* (*HMOX-1*) und der *Glutathione S-Transferase* (*GST*) Superfamilie die Wirkung einer Luftverschmutzungs-Reduktion auf die Lungenfunktionsabnahme modifizieren. Diese Gene gehören zur ersten Abwehrlinie des Körpers gegen oxidativen Stress. Die Analyse war streng an eine zuvor publizierte Arbeit angelehnt, welche zeigte, dass eine Verbesserung der Luftqualität mit einer Verlangsamung der natürlichen Lungenalterung assoziiert war, besonders in den kleinen Atemwegen, welche durch den Parameter FEF_{25-75} repräsentiert wurden (dieser ist durch die Flussgeschwindigkeit im mittleren Teil der Ausatmung definiert). Wir beobachteten, dass Mutationen in *GSTP1* und *HMOX-1* die Wirkung einer reduzierten PM_{10}-Belastung auf die Lungenalterung signifikant modifizierten, mit größten Veränderungen bei FEF_{25-75}. Der Nutzen einer verbesserten Luftqualität verteilte sich daher nicht zu gleichen Teilen auf die Bevölkerung, sondern Personen mit unterschiedlicher Fähigkeit, oxidativen Stress zu verarbeiten, profitierten davon in unterschiedlichem Masse. Dies ist potentiell relevant für die Grenzwertsetzung. Es war jedoch schwierig die Resultate auf biologischer Ebene zu interpretieren, da bei den interagierenden genetischen Varianten die funktionellen Auswirkungen bezüglich Abbau von Luftschadstoffen nicht bekannt sind.

In der letzten Arbeit wurde die Interaktion zwischen PM_{10}- oder Tabak-Belastung mit Genen und Reaktionswegen, die für oxidativen Stress relevant sind, verglichen. Die Analyse fußte auf 878 Nicht-Asthmatikern mit genomweiten Daten. 152 Gene, 14 Reaktionswege und 12679 Mutationen wurden durch eine Pathway-Analyse gemäss der ARTP-Methode untersucht. Nach Korrektur für multiples Testen fanden wir, dass die Gene *CRISP2* signifikant, und *SNCA* marginal mit kumulativer PM_{10}-Belastung auf die Veränderung von FEV_1/FVC interagierten. Eine vergleichende Analyse auf Mutations-Ebene brachte neben der *SNCA*-Interaktion eine weitere Mutation im Gen *PARK2* hervor, die mit PM_{10} die longitudinale Veränderung von FEV_1 beeinflusste. Der Vergleich von nominal mit PM_{10}- oder Tabak-Exposition interagierenden Genen (P-Wert für Interaktion <0.05) zeigte, dass die Überlappung der Interaktionsmuster zwischen den Expositionen sehr gering war. Bei Fokussierung auf die am stärksten interagierende Mutation innerhalb eines nominal signifikanten Gens zeigte sich, dass die

Tabak assoziierten Interaktions-Effekte größer ausfielen als jene bei PM_{10}. Der Prozentsatz der erklärten Variabilität in der Lungenfunktionsveränderung war jedoch vergleichbar zwischen den Expositionen, und reichte bis zu 28.5%. Dies ist vermutlich eine Überschätzung aufgrund des so genannten „winner's curse" Effekts, der auftritt, wenn Effektschätzer bei kleinen Stichproben nach Stärke gefiltert werden. Weiter wurde diese Schätzung relativiert durch die fehlende Replikation der beobachteten Interaktionen in *CRISP2* und *PARK2* in der restlichen SAPALDIA Population. Schlussfolgernd waren statistisch signifikante Interaktionen auf Mutations- oder Gen-Ebene nicht nachweisbar, die Resultate deuten jedoch darauf hin, dass unterschiedliche Gene die Effekte von Luftverschmutzung und Rauchen auf die Lungenfunktionsabnahme vermitteln.

Diskussion und Schlussfolgerungen

Unsere Resultate legen nahe, dass die COPD Inzidenz in der Schweiz am oberen Ende von vergleichbaren internationalen Schätzern liegt, unter Vorbehalt der Überschätzung durch die fehlende Atemwegs-Dilatation und damit verbunden nicht diagnostiziertes Asthma. Dieses schlägt sich jedoch auch in einer erhöhten Belastung des Gesundheitswesens nieder. Der Einbezug weiterer klinischer Parameter könnte eine bessere Identifizierung von Personen mit milder Obstruktion und hohem Progressionsrisiko ermöglichen. Für eine klare Differenzierung zwischen COPD und Asthma und ihrer Auswirkungen auf das Gesundheitssystem ist die Atemwegs-Dilatation vor Lungenfunktionsmessung unumgänglich. Die beobachteten Interaktionen zwischen PM_{10}-Belastung und „Oxidative Stress" Genen legen nahe, das die Luftverschmutzung über die Modulation der Lungenfunktionsabnahme zur COPD-Entstehung beiträgt. Der Nutzen einer verbesserten Luftqualität wird umgekehrt nicht gleich hoch in der ganzen Bevölkerung sein, sondern hängt von der individuellen Empfindlichkeit gegenüber oxidativem Stress ab. Die Resultate der Pathway-Analysen deuten darauf hin, dass unterschiedliche Gene und Reaktionswege durch PM_{10} und Tabak-Exposition aktiviert werden, möglicherweise aufgrund der unterschiedlichen Intensität des verursachten, oxidativen Stresses. Potentiell könnte der Beitrag von Luftverschmutzung in der COPD Entstehung auf Populationsebene jenem von Tabak vergleichbar sein, der Expositions-spezifische Anteil erklärter Outcome Variabilität muss jedoch von weiteren Studien untersucht werden.

Diese Resultate sind von Public Health Relevanz, da sie die Bedeutung von oxidativem Stress im natürlichen Alterungsprozess der Lunge, und damit möglicherweise auch bei der Entstehung von COPD, hervorheben. Zukünftige methodische Weiterentwicklungen werden die Identifikation von Schlüssel-Enzymen und –Reaktionswegen und damit auch die Entwicklung neuer Präventions- und Behandlungs-Strategien ermöglichen. Zum jetzigen Zeitpunkt wäre die Empfehlung Überlegungen wert, in Zeiten höherer Luftverschmutzung Antioxidantien an empfindliche Bevölkerungsgruppen wie Kinder und ältere Leute mit Vorerkrankungen zu verabreichen. Dies könnte relativ einfach durch eine gemüse- und früchtereiche Ernährung oder Nahrungszusätze erfolgen. Auch sollte die Rolle der Luftverschmutzung in der COPD-Entstehung größere Beachtung in Public Health und Forschung finden.

Die Analysen profitierten von verschiedenen Charakteristiken der zugrunde liegenden Daten. Das bevölkerungs-basierte Design ermöglichte die Untersuchung von COPD Frühstadien wie beschleunigter Lungenalterung. Weitere Stärken waren die Verfügbarkeit von validierten Schätzern für die individuelle Luftschadstoffbelastung, standardisierte Lungenfunktionsmessungen und DNA Proben von hoher Qualität, sowie die Verwendung moderner Analysemethoden zur Untersuchung von Reaktionswegen. Die wichtigsten Limitationen beinhalteten die fehlende Atemwegs-Dilatation bei der Lungenfunktionsmessung, die Beschränkung auf zwei Lungenfunktionsmessungen (dadurch könnten Mess-

fehler zur Unterschätzung von Luftschadstoff-Effekten führen), und die limitierte Stichprobengröße für genomweite Analysen.

Diese Einschränkungen könnten durch den Aufbau einer großen, nationalen Kohorte mit detaillierten Daten zu Krankheitscharakteristiken, Umweltexpositionen und genomweiter genetischer Variabilität überwunden werden. Eine solche Unternehmung würde das Engagement aller Schlüsselstellen im Schweizer Gesundheitsbereich bedingen. Aus wissenschaftlicher Sicht wäre dies äußerst wertvoll, da neben dem Beitrag von Luftverschmutzung zur COPD-Entstehung Determinanten vieler weiterer chronischer Erkrankungen untersucht werden könnten.

Summary

Introduction

Chronic obstructive lung disease (COPD) is of major Public Health importance in terms of its global impact on morbidity, mortality and health care costs, although international estimates of its prevalence and burden differ widely due to the use of different disease definitions. Despite this large impact, the disease is often not diagnosed but in an advanced stage, where no effective therapy exists to date.

Disease diagnosis is thereby based predominantly on lung function measurement, while taking account of additional risk factors such as smoking. To facilitate early detection and comparability across studies, the global initiative for obstructive lung disease (GOLD) has proposed an easily applicable disease definition based on the ratio of the forced expiratory volume in the 1st second of exhalation over the totally exhaled volume (FEV_1/FVC) smaller than 0.7, measured after pharmacological airway dilatation. The latter is required to distinguish COPD from hidden asthmatic disease. The definition was first widely adopted including epidemiological studies, but was subsequently criticized for causing over-diagnosis in older ages. Owing to the natural lung function decline, the critical threshold could be reached by healthy persons at age 45 years. The alternatively proposed FEV_1/FVC lower limit of normal, defined as the 5th percentile of the distribution in a healthy non-smoking population, has since substantially gained ground in the clinical as well as research setting. In consequence, the use of both definitions is warranted in studies on COPD disease burden today.

Though COPD represents a major public health problem, the etiological pathways upon which it arises are not yet clear. Tobacco smoking has traditionally been the most important risk factor, but emerging evidence from recent years points to the importance of occupational exposures, domestic biomass burning for cooking and heating, environmental tobacco smoke exposure as well as ambient and traffic related air pollution as important determinants. Two major pathways have thereby been proposed as etiological frameworks for COPD. First, an imbalance of endogenous proteases and antiproteases could lead to destruction and alterations of lung tissue typical of COPD. Second, according to the oxidant/antioxidant imbalance hypothesis an overload of oxidants compared to the antioxidant defenses of the body could lead to oxidative stress. Oxidants are free radicals that react and interfere with normal cell structures like cell walls, proteins and the DNA, and cause inflammatory reactions with further oxidative stress. Genes belonging to these systems have successfully been related to COPD or lung function decline in candidate gene studies, underlining their importance. However, replication of these findings has proven difficult, even with the availability of genome wide studies. Failure to account for gene environment interaction could underlie this difficulty.

Most of the environmental exposures suggested as COPD risk factors above are known to induce oxidative stress due to their content in free radicals or by stimulation of endogenous production. Further, the adverse effects of exposure on the tissue are not uniform, but depend on individual susceptibility to oxidative stress, determined by the functional capacity in underlying enzymes and

hence also variation in the respective genes and pathways. Identifying individual susceptibility factors thus requires the study of gene-environment interaction.

For ambient air pollution, and especially its particulate matter component which is known to induce oxidative stress, substantial adverse health effects have been described including mortality, hospitalizations for cardiopulmonary morbidity, respiratory symptoms and decreased lung function growth. These make ambient air pollution one of the major environmental threats today, and an important candidate risk factor in the context of COPD.

The aim of the thesis was thus to estimate the burden of COPD in Switzerland based on available population based data, while assessing the suitability of epidemiological data for this endeavor. Further, the role of ambient air pollution in the development of COPD is to be investigated via its possible interaction with variants in primary oxidative stress candidate genes. Finally, the estimated impact of air pollution exposure at the population level is to be compared to that of the traditional COPD risk factor, tobacco smoking, focusing on involved genes and pathways which could mediate the effects.

Methods

The work related to this thesis was based on data from the Swiss Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA). This is a population based cohort study initiated in 1991 to study the health effects of long term exposure to ambient air pollution on lung, allergic and heart diseases. A first follow-up has been conducted in 2002, and a second one has just been completed in 2010/11. The current work drew upon data from the baseline and first follow-up assessment in which 9651 and 8047 persons participated, respectively. Participants underwent a detailed health questionnaire including questions on smoking and other life style habits, occupational exposures, and preexisting disease. Lung function testing was applied in both examinations without dilatation of the airways but according to a standardized protocol and strict quality control including use of the same devices at both time points. Individual air pollution exposure estimates were available for particulate matter of median diameter less than 10µm (PM_{10}) in terms of change and cumulative exposure over 11 years of follow-up. These were based on a Gaussian dispersion model using Swiss emission data from years 1990 and 2000 and interpolation using historical trends of fixed air pollution monitoring stations. Blood samples were drawn in 2002 and DNA samples of over 6000 persons were available for investigation of candidate genes. In the framework of the large international GABRIEL consortium, genome wide data was obtained on 1457 persons, including all asthmatics and a random sample of non-asthmatic participants.

The nature of the available study data had the following implications on investigating the thesis questions: As no lung function measurements after airway dilatation were available, a modified GOLD definition of COPD was used (based on the pre-bronchodilation values), or alternatively change in lung function decline was studied as proxy measure for COPD development. Due to the more homogeneous distribution of particulate matter pollution compared to other components, only PM_{10} exposure could be used as valid measure of individual air pollution exposure. Third, investigation of pathways involved in disease causation had to be limited to the relatively small group of non-asthmatic persons with available genome wide data.

Results

Using a modified GOLD-definition of COPD, we found an incidence of 14.2 cases/1000 person years (PY) in the first article. This was at the higher end of estimates from comparable population-based cohorts and could only partly be explained by differences in age, smoking distribution or length of follow-up. While positive associations of incidence with age and smoking were consistently described in the previous literature, we found chronic bronchitis also to be a significant determinant. Using FEV_1/FVC < lower limit of normal (LLN) to define obstruction yielded expectedly lower incidence estimates of 7.2 cases/1000PY. 20.9% of obstructive cases at baseline did not persist, most of them presenting mild baseline obstruction. Progression from mild to moderate or severe stages or persistence in moderate to severe stages of obstruction during 11 years of follow-up was associated with more frequent health service use and dyspnea at follow-up, in contrast to persisting in mild obstruction. Non-persistence of obstruction was marginally associated with higher health service use, despite broad adjustment for asthma, and on average, this category had normal lung function values but a mismatch between FEV_1 and FVC. This was possibly due to cases of hidden asthma. We concluded that pre-bronchodilation spirometry as frequently used in epidemiological studies has prognostic value in predicting future adverse health events and health service use, though misclassification with hidden asthma might be substantial. Additional clinical characteristics could be useful to identify participants at risk of progression.

In the second paper, we investigated whether variants in genes *heme-oxygenase 1* (*HMOX-1*) and the *glutathione S-tranferase* (*GST*) gene superfamily, genes which belong to the body's first line defense against oxidative stress, modify the effect of a reduction in PM_{10} exposure during follow-up on lung function decline. The employed analysis model was strongly based on previously published work, which had shown that a reduction in air pollution exposure was associated with attenuation in the natural, age-related lung function decline, particularly in the smaller airways as measured by FEF_{25-75} (the velocity of the airflow in the mid-portion of exhalation). We observed that variants in *GSTP1* and *HMOX-1* significantly modified the effect of a reduced PM_{10} exposure, with the strongest interaction effects again observed for decline in FEF_{25-75}. The benefits of a reduction in air pollution exposure are thus not equally distributed across the population, but individuals with a differing endogenous capacity to cope with oxidative stress profit from it to a different extent. This finding potentially has policy implications, but limited knowledge about how the interacting variants alter the biological processing of air pollutants warrant caution in interpretation.

Finally, in the last paper we compared the interaction between ambient PM_{10} or tobacco smoke exposure and genes and pathways involved in oxidative stress on lung function decline in a subset of 878 non-asthmatic adults with available genome-wide data. The study comprised 152 genes, 14 pathways, and 12679 single nucleotide polymorphisms (SNPs). A pathway analysis was applied using the published ARTP-method. After multiple testing correction, we found that genes *CRISP2* significantly, and *SNCA* marginally interacted with PM_{10} on the decline in FEV_1/FVC. In comparison, a pure SNP-level analysis yielded one additional interaction: a SNP in gene *PARK2* significantly interacted with PM_{10} on FEV_1-decline. When looking at the interaction pattern across nominally significant genes ($p_{interaction}<0.05$), we observed that different genes and pathways were interacting with PM_{10} and tobacco smoke exposure. Focusing on the strongest SNP in nominally significant genes, tobacco smoke exposure presented larger interaction effects on lung function decline than PM_{10}. However the percent explained variability in lung function decline was similar for both exposures, ranging up to 28.5%. This is probably an overestimation due to the so-called "winners

curse" phenomenon. The estimates were further questioned by non-replication of the interactions with *CRISP2* and *PARK2* in the remainder of the SAPALDIA population. This was possibly related to small sample size. In conclusion, we were not able to detect consistent, significant interactions using either SNP-level or gene-level analysis after strict correction for multiple testing. However, by evaluating nominally significant interactions, we observed suggestive evidence that different genes and pathways are involved in mediating the effects of PM_{10} and tobacco smoke exposure on lung function decline.

Discussion and Conclusion

Our results imply that the incidence of COPD in Switzerland is at the higher end compared to data from other countries, with the reservation that estimates are based on pre-bronchodilation spirometry. They might thus be overestimated by hidden asthma (which also imposed burden on health service use). Besides spirometry, clinical criteria might be needed to discern who among mildly obstructed persons is at risk of progression. Separating the respective contributions of COPD and asthma to the overall disease burden warrants bronchodilation. The observed interactions between ambient PM_{10} exposure and oxidative stress defense genes suggest that air pollution contributes to COPD development by modulating lung function decline. In turn, the expected benefits from cleaner air are not going to be equally distributed among all members of society, but are determined by individual genetic make-up defining the capacity to cope with oxidative stress. Finally, according to the results of our pathway interaction analyses, genes and molecular pathways activated by ambient air pollution potentially differ from those induced by tobacco smoke. This could be related to different levels of oxidative stress. Finally, the impact of air pollution at the population level might potentially be comparable to that of tobacco smoke, but further research needs to address the exposure-specific percentages of explained variability in lung function decline.

These findings are important for Public Health and prevention, as they underline the important role of oxidative stress in shaping the natural decline in lung function, and hence possibly the risk of COPD. Future methodological improvements might lead to the identification of key enzymes and pathways in COPD causation with subsequent development of new preventive measures and therapies. But for the time being, thought is warranted about whether to recommend antioxidant administration via diets enriched with fruits, vegetables or supplements to susceptible population groups such as children or the elderly during periods of high pollution. Further, the role of air pollution in COPD causation deserves greater attention in Public Health, policy and epidemiological research. Our studies benefitted from several strengths including a population based design that allowed studying pre-clinical COPD stages like accelerated lung function decline, validated individual level air pollution exposure estimates, highly standardized lung function measurements, high quality DNA from a large part of the population, and application of modern analysis techniques. The most important limitations include the lack of post-bronchodilation spirometry, availability of only two lung function measurements (implying possible underestimation of air pollution effects by non-differential measurement error), and the reduced sample size for genome wide analyses.

These limitations could be overcome with a new, large national cohort with detailed phenotype, exposure and genome wide variation data. Such an endeavor requires commitment from all stakeholders in the Swiss Public Health field, but has high scientific merit as it would not only allow defining the fraction of COPD attributable to air pollution, but also studying determinants of other chronic diseases.

Abbreviations

ATS	American Thoracic Society
COPD	Chronic obstructive lung disease
DNA	Deoxyribonucleic acid
ERS	European Respiratory Society
FEF_{25-75}	Velocity of the airflow in the middle portion of a forced expiratory maneuver. Proxy measure for the patency of small airways in the lung.
FEV_1	Volume blown out in the first second of a forced expiratory maneuver
FVC	Volume blown out totally in a forced expiratory maneuver
GOLD	Global Initiative for Obstructive Lung Disease
GWAS	Genome-wide association study
PM_{10}	Particulate matter air pollution with a median diameter less than 10µm
$PM_{2.5}$	Particulate matter air pollution with a median diameter less than 2.5µm
SNP	Single nucleotide polymorphism, meaning a single base pair mutation.
UFP	Ultrafine particles, corresponding to the particulate matter fraction with median diameter less than 100nm.

1. Introduction

1.1. Chronic Obstructive Lung Disease (COPD)

Chronic obstructive lung disease (COPD) is the most widespread chronic airway disease besides asthma in industrialized countries today. The disease is associated with abnormally elevated levels of inflammation in lung tissue leading to a progressive loss of lung function and consequently increasing dyspnoea, reduced exercise capacity and quality of life[1]. The Global initiative for Obstructive Lung Disease (GOLD) and American Thoracic Society (ATS)/ European Respiratory Society (ERS) guidelines define the disease in the following way: "A disease state characterized by airflow limitation that is not fully reversible. The airflow limitation is usually both progressive and associated with an abnormal inflammatory response of the lungs to noxious particles or gases" [2,3].

1.1.1. Epidemiology and Public Health Burden

COPD is substantially and progressively contributing to the global burden of disease[4,5]. While ranking as 12th cause of disability and 6th cause of mortality worldwide in 1990, according to estimates from the Global Burden of Disease Study, it will likely be the 5th cause of disability and 3rd cause of mortality by 2020[6]. Consequently, direct and indirect economic costs arising from treatment and loss of productivity are enormous[4,7]. In 1998, the US National Heart, Lung and Blood Institute (NHLBI) and the World Health Organisation (WHO) formed the **GOLD initiative** to raise awareness for the disease and establish a consensus on diagnosis and treatment[1].

In light of the disease's Public Health importance, several studies were performed to provide data on COPD prevalence and incidence rates in the general population. While their results varied considerably depending on the chosen disease definition, most of them used diagnostic criteria recommended by the GOLD initiative, where a ratio of the forced expiratory volume in the 1st second over forced vital capacity (FEV_1/FVC) <0.7 in post-bronchodilation spirometry defines presence of airflow obstruction. FVC thereby represents the total volume of air that can be blown out in a forced expiratory manoeuvre. According to GOLD criteria, values of FEV_1 are then used for further severity classification: $FEV_1 \geq 80\%$ defines mild GOLD stage I, < 80% moderate stage II, <50% severe stage III and <30% very severe stage IV disease.

Prevalence and incidence estimates cited in the following paragraphs are most often based on the GOLD-definition, although many of the epidemiological studies did not apply bronchodilation required to distinguish fixed airway obstruction as occurring in COPD from a reversible form mostly present in asthma. Further details on disease definitions are given in section 1.1.6.

1.1.2. Prevalence

The overall prevalence of COPD was estimated to be 9.2% (95% CI: 7.7-11.0%) in a meta-analysis by Halbert and colleagues[8] based on data from international population-based studies published

between 1990 and 2004, with estimates ranging from 5.5% (3.3-9.0%) to 21.8% (4.7–61.4) depending on the chosen spirometric criteria to define disease. Importantly, the review noticed that most of the prevalence estimates from epidemiological studies were based on pre-bronchodilation spirometry, although most guidelines require post-bronchodilation spirometry for the COPD diagnosis. To overcome these limitations, the international Burden of Obstructive Lung Disease (BOLD) study was designed with a strong focus on achieving comparability, and applied standardized questionnaires and post-bronchodilation spirometries to assess the prevalence of COPD in 12 countries based on representative population-samples[9]. The study focused on prevalence estimates for moderate to severe stages of disease in persons aged 40 years and over, because COPD starts manifesting in this age range and the mild forms are subject to misclassification with the naturally occurring age-related lung function decline[9] (see section 1.1.5). Despite careful consideration of differing sampling strategies and applying uniform methodology across countries, the authors found a wide heterogeneity of prevalence estimates for moderately severe COPD, ranging from 5.9% in Germany to 19.1% in South Africa, which could not be explained by differences in age or smoking distribution. The overall prevalence was estimated to be 10.1%. A similar picture emerged in the PLATINO study, a population-based survey of metropolitan areas in five Latin-American countries, where estimates for mild COPD ranged from 7.8% in Mexico to 19.6% in Uruguay, and did not change substantially after adjustment for sex, body mass-index, ethnic origin, smoking exposure, domestic biomass and coal pollution as well as exposures in the workplace[10]. In a recent systematic review on available European data, Atsou et al. showed that spirometry based prevalence rates varied across countries from 2.1-26.4% overall, and still from 4.5% to 26.1% in the upper age-range of 40 years and more[11]. Thereby both, pre- and postbronchodilation spirometry was used by the original studies. The authors related part of the variability to differing study population characteristics (general population samples, patient or occupational cohorts) and different levels of tobacco smoke exposure, but different spirometric disease definitions also played an important role[11].

In summary, the comparability of many earlier prevalence studies is hampered by differing study characteristics and disease definitions[8,11]. In contrast, the BOLD and PLATINO studies benefited of standardized disease definitions and methodology yielding highly comparable estimates across countries, but the results imply a wide range of site-specific prevalence estimates up to 26%, which cannot be explained by differences in age structure or distribution of smoking exposure. Further, population based prevalence studies yielded substantially higher prevalence estimates than those based on physicians diagnoses[11], which underlines the public health importance of the disease and suggests considerable under-diagnosis.

1.1.3. Incidence

Compared to available prevalence estimates, COPD incidence data is sparse and was only estimated by a few large, prospective cohort studies comprising different populations in terms of age range, follow-up time and study setting. Annual COPD incidence was estimated to 3 cases/1000 person-years in young adults aged 20-44 years[12], and 16 cases/1000 person-years in an older population aged 46-77 years [13]. A study in Copenhagen, Denmark observed a 20% cumulative incidence over 25 years of follow-up in subjects over 19 years of age[14]. In contrast, only 6% of a Norwegian random

population sample aged 18-74 years developed COPD after nine years[15]. Thus incidence rates varied widely depending on age distribution and from country to country.

1.1.4. Burden of COPD in Switzerland

In Switzerland, no population-based prevalence study was carried out until recently, although published lung function data from the first examination of the Swiss Study on Air Pollution and Lung Disease in Adults (SAPALDIA) in 1991[16,17] had been used in the meta-analysis by Halbert and colleagues[8]. Bridevaux et al focused on data from the SAPALDIA follow-up survey in 2002 due to the favourable age distribution (median age 50 years, range 30-73), and estimated the prevalence of moderate airflow obstruction to 7% according to the GOLD definition, and 5.1% according to a ratio of FEV_1/FVC smaller than its lower limit of normal (LLN)[18]. The LLN is calculated as the 5th percentile of the normal distribution in healthy never-smoking adults. These estimates are at the lower end of comparable studies across Europe including the European centres of the BOLD study[9]. Population-based COPD incidence data was not available before publication of the work related to this thesis (first thesis paper).

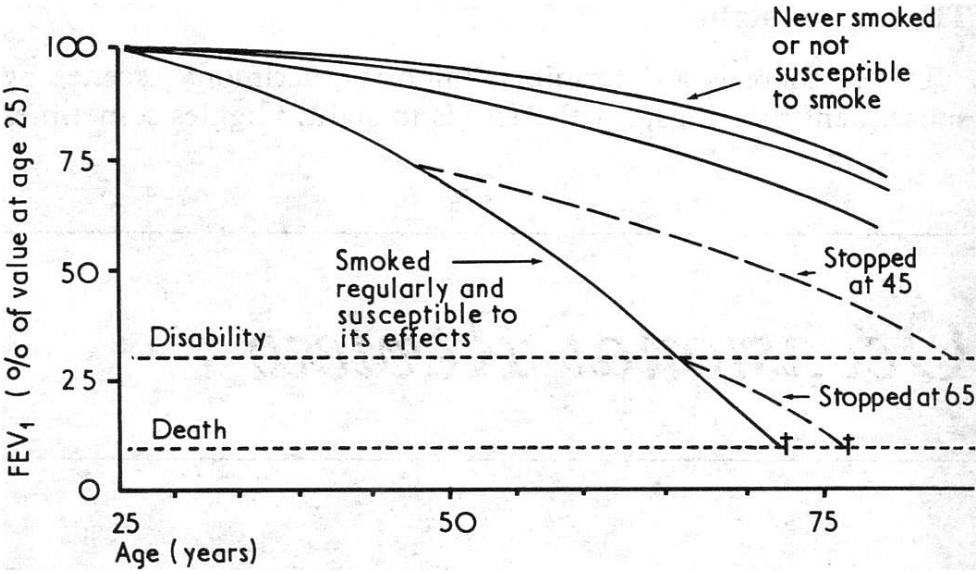
1.1.5. Symptoms and clinical presentation

COPD is characterized by ongoing loss of breathing capacity due to an underlying, progressive narrowing of the airways and loss of gas exchanging lung surface[1]. Severe stages of disease are frequently associated with respiratory symptoms such as chronic cough and phlegm production, as well as shortness of breath while walking. These symptoms are accentuated while the disease progresses over several years, and ultimately patients suffer from incapacitating dyspnea, reduced exercise tolerance and also impaired quality of life[19]. COPD is often accompanied by other chronic conditions such as loss of muscle mass, weight loss (decline in BMI), cardiovascular disease, disturbances of lipid and glucose regulation, osteoporosis and depression, which further add to the disease burden and also economic costs[20,21]. The pathogenetic link between COPD and these comorbidities is currently not clear and actively researched[21].

A cardinal feature of the disease is the accelerated loss in FEV_1 compared to the normal, aging-related lung function decline, which starts after lung growth has reached its peak by the age of 25 years. The accelerated decline is induced and maintained by exposure to noxious inhalatory particles like tobacco smoke and others. This characteristic of the disease has long ago been described by Fletcher and Peto in one of the first publications on COPD-disease (**figure 1**)[22]. Importantly, the authors noted already at that time that not all individuals are susceptible to the effects of smoking.

On the tissue level the disease presents mainly in the form of two processes of differing intensity, one being characterized by progressive loss of tissue walls between the lung alveoli, i.e. in the gas exchanging part of the lung, which results in lung emphysema, the other presenting as inflammation and tissue remodeling in the larger, air transporting airways, resulting in fibrotic changes and fixed narrowing of the bronchi. Airway obstruction is also enhanced by the emphysematous processes via

Figure 1 Course of decline in FEV_1 according to exposure to tobacco smoke from [22]



loss of elastic lung fibers whose traction normally helps to maintain open airways. Progressive obstruction leads to air trapping in the lungs, which is clinically often observed in chest radiographs in the form of hyperinflation and a flattened diaphragm, while pronounced emphysema can lead to large confluent air spaces presenting as bullae on radiography.

1.1.6. Diagnosis

One of the large problems to solve to tackle COPD-related morbidity is timely diagnosis. According to epidemiological data, more than half of the patients remain undiagnosed and untreated, even when their disease has already progressed to a moderate severity stage[9,23,24]. Hence at the time of diagnosis, often a large part of the patient's lung function has already been lost. The problem is that in early stages of the disease the symptoms are rather unspecific and the degree of impairment in everyday life is minor.

In contrast, there is consensus among experts that the disease needs to be diagnosed in an early stage by timely detection of the typical spirometric changes, if a chance of applying effective preventive or therapeutic measures shall be kept1. The diagnosis of COPD is predominantly based on spirometry and different spirometry-based COPD-definitions exist. Appendix 1 gives a detailed overview on pulmonary function testing.

1.1.6.1. Spirometric criteria to define COPD

Different spirometric criteria have been proposed to define COPD[25], and it is well known that the different definitions produce a wide range of prevalence estimates[26]. The most frequently used

definitions are those proposed by the GOLD initiative and the European Respiratory Society (ERS)/American Thoracic Society (ATS) based on a fixed cutoff of 0.7 for the FEV_1/FVC ratio to define airway obstruction[3,27], although the validity of this cutoff is much contested today (see section 1.1.6.2).

GOLD COPD definition:

COPD is present if the following criteria are met in post-bronchodilation spirometry:

- Mild GOLD stage I: FEV_1/FVC < 0.7 and FEV_1 >=80% of the predicted value
- Moderate stage II: " and 80%> FEV_1 >=50% "
- Severe stage III: " and 50%> FEV_1 >=30% "
- Very severe stage IV: " and 30%> FEV_1 "

FEV_1 values are thereby compared to expected values, based on prediction equations derived from a healthy, non-smoking adult population of the same ethnicity. Many population-specific prediction equations have been published to date[28], and most of them take into account the proband's sex, age, and possibly height. In the European context the most frequently used prediction equations are those published by Quanjer and colleagues in 1993 for the European Community of Coal and Steel (ECCS equations)[29,30]:

Females:

FEV_1:	3.95 * height – 0.025 * age -2.60	RSD=0.38
FVC:	4.43 * height – 0.026 * age -2.89	RSD=0.43
FEV_1/FVC: (in %)	89.1 – 0.19 * age	RSD=6.51

Males:

FEV_1:	4.30 * height – 0.029 * age -2.49	RSD=0.51
FVC:	5.76 * height – 0.026 * age -4.34	RSD=0.61
FEV_1/FVC: (in %)	87.21 – 0.18 * age	RSD=7.17

Thereby volumes are in litres, height is in metres, and age in years. For adults aged <25 years, 25 should be substituted for age. RSD means residual standard deviations of the prediction equations (i.e. the standard deviation of the residual difference between predicted and measured lung function values after taking account of height and age).

For studies on the North-American continent, the most widely applied equations are those published by the NHANES III study[31].

Comparative studies have shown that the ECCS equations often produce lower predicted lung function values than equations calculated for specific geographic populations[32-35], or those derived from large population based studies such as SAPALDIA[36] or NHANES III[32]. The probable reason for the frequent underprediction is that the ECCS equations were derived from lung function measurements and equations in different populations including smokers over three decades (1954-1980)[30,32]. They thus more represent average prediction values across populations, and might not optimally fit single ones. In spite, or perhaps also because of this shortcoming, they are still used in comparative analyses of population based studies in order to use a common set of prediction equations that approximately fits the involved studies to the same extent.

ERS/ATS COPD definition:

The spirometric criteria proposed by the ATS/ERS are practically identical to those of the GOLD initiative, with the exception that a ratio of 0.7 is already classified as pathologic[3]. Further, besides stressing that COPD is a preventable and treatable disease, the ATS/ERS consensus statement[37] emphasizes that these criteria should be applied to high risk groups such as persons with respiratory symptoms or heavy smokers.

1.1.6.2. Fixed FEV_1/FVC cutoff versus lower limit of normal

There is currently much debate about the validity of using a fixed cutoff of $FEV_1/FVC<0.7$ to define airway obstruction[38-40].

This fixed cutoff was initially proposed by an expert panel to have diagnostic criteria which are easy to remember and readily applicable in the clinical setting with the hope to facilitate early detection of airway obstruction[38]. Moreover, use of a fixed cutoff for the first time facilitated comparability of prevalence estimates across epidemiological studies.

The inherent problem of using a fixed ratio cutoff is however that due to the natural lung function decline, which starts from the age of 25 years, this threshold can normally be reached around the age of 45 years, even by non-smokers[25,40,41]. It was estimated that by using the fixed FEV_1/FVC ratio cutoff of 0.7 between 35-60% of healthy men aged 70 years or older would be falsely classified as diseased[25,42,43]. On the other side of the age spectrum, younger adults with a ratio above but close to the 0.7 threshold would still be classified as healthy, although their expected normal value would be considerably higher[25]. Accordingly, between 7-30% of young adults aged <50 years with irreversible airways obstruction might be missed by focusing on a fixed threshold of 0.7[43,44]. Overall, the fixed ratio cutoff proposed by the GOLD initiative and ATS/ERS thus leads to considerable over-diagnosis in older age, while at the same time young adults are under-diagnosed.

This misclassification issue has brought up the proposal of using the lower limit of normal (LLN) of the FEV_1/FVC ratio to define the threshold of obstruction. The LLN would thereby be represented by the 5^{th} percentile of the distribution of FEV_1/FVC observed in a healthy, non-smoking, asymptomatic adult population of the same ethnic origin as the tested proband. Based on physiology, the expected lung volumes and hence also the LLN will naturally be different for males and females, for different heights, and would also depend on age.

While study-specific LLN values have been published, including for the SAPALDIA study[36,45], the LLN is most often calculated by subtracting 1.645 standard deviations from the value predicted by one of the widely used standard equations in populations of European origin[29,31]. As an example, using the equations by Quanjer et al.[29], the LLN for FEV_1/FVC would be calculated as:

LLN FEV_1/FVC: 89.10 – 0.19 * age -1.645 * 6.51 (RSD) in females

87.21 – 0.18 * age -1.645 * 7.17 (RSD) in males

where RSD=residual standard deviation, as defined on page 19.

Obviously, also the LLN criterion leads to a small degree of misclassification, as 5% of the healthy, adult population would be classified as diseased. It is however considerably more restrictive than the fixed cutoff criterion, and resulting estimates for the prevalence of airway obstruction are about 40% lower compared to using a fixed ratio of 0.7[25,46,47], with the difference being particularly prominent in the oldest age-segments. In the BOLD study, Vollmer and colleagues have found that the variability of COPD prevalence estimates was considerably reduced when using LLN, although still significantly different between study sites, ranging from 7.1-8.6%[48].

The discussion on how to best define airway obstruction is still ongoing, and in clinical practice as well in research, both definitions can be observed.

In the clinical setting, the fixed 0.7 cutoff-criterion is thereby less prone to misclassification, as most of the diseased persons present with respiratory symptoms, frequently a smoking history, as well as in later disease stages. In such a population the prevalence of disease is likely to be high, and thus the fixed cutoff criterion will often perform well, except in young adults aged 50 years or less[43]. This might be also a reason why the GOLD initiative still propagates it in its position statements[1]. Likewise, ATS/ERS emphasized that the definition shall be used in persons with a positive history for these risk factors in its COPD position paper from 2004[3]. However, in light of the consequences that a false-positive test-result and hence over-diagnosis in elderly persons entails (lifelong treatment and possible anxiety), the ATS/ERS shortly afterwards changed its recommendations to using the LLN for the interpretation of spirometries[28], and an urgent change of the major COPD guidelines is currently advocated[39,40].

In the research setting, and particularly in population-based epidemiological studies, the issue of misclassification when using the fixed cutoff criterion is even more pronounced (although without direct negative consequences for participating individuals), as most study participants are in good respiratory health, and the pre-test probability of disease is small. The stricter LLN criterion thus helps to minimize the rate of misclassification, from which many epidemiological studies lacking post-bronchodilation spirometry might even more benefit, as they struggle with additional misclassification by presence of hidden asthma. However, only a few population-based epidemiological studies have so far investigated and shown the prognostic value of using the LLN criterion to predict future adverse health events[49-51], while there is more (although contested) evidence for the fixed cutoff in this respect[49,52-54]. The comparative predictive performance of the criteria thus remains to be determined.

To summarize, the lower limit of normal criterion has gained ground not only in research but also in the clinical setting during the last years, and it will probably become standard in the future[28].

Currently prediction equations for a wide spectrum of different ethnic populations are available (a good overview is given on the homepage of one of the leading respiratory epidemiologists, Prof. P Quanjer, under www.spirexpert.com/epidemiol7.htm, as accessed on July 15th 2011), but their readily application in clinical practice is hampered by the failure of many spirometer manufactures to make them available in their devices. At the same time, the GOLD and older ATS/ERS criteria including the fixed ratio cutoff are still widely used in epidemiological research for reasons of comparability with earlier studies. However, parallel or sensitivity analyses using the LLN criteria are clearly warranted in light of the limitations of using a fixed ratio criterion.

1.1.7. Treatment

Therapeutic options to slow the progression of COPD are still scarce. So far, the therapy consists primarily in removing the noxious inhalatory exposures, most often meaning smoking cessation, and symptomatic treatment with broncho-dilating drugs (which relieves the notion of respiratory distress), and possibly corticosteroids, given in combination with antibiotics in the case of infectious exacerbations (to suppress respiratory deterioration by inflammatory reactions). Recent results from the TORCH, UPLIFT and GLUCOLD trials reported promising effects on slowing the accelerated FEV_1 decline in moderate COPD stages by a long-term maintenance therapy with long-acting bronchodilators and inhaled steroids, given either alone or in combination[55-57]. While these findings challenge the negative results of previous studies with shorter follow-up, they are based on a posthoc analysis, and the case for clinically applying maintenance treatment in early COPD stages is thus still open[58,59]. The partly conflicting findings from these studies also stimulated the idea that only specific sub-phenotypes of COPD might respond to a certain maintenance treatment[60].

In light of the large mortality and morbidity burden the disease imposes worldwide[6], new insights into the molecular processes underlying COPD are urgently needed to identify new targets for prevention and therapy.

1.1.8. Risk Factors and Disease Aetiology

1.1.8.1. Environmental Risk Factors

Since it was described by Fletcher and Peto[22], COPD has traditionally been understood as a smoking-related disease, and there is wide consensus that smoking is still the most important risk factor for COPD[1,3]. But in the last years, it has increasingly been realized that many other inhalatory exposures contribute substantially to COPD development[61,62]. Current estimates state that about one third of prevalent COPD cases affect lifetime never smokers[61], and in line with this, 30% of all cases of prevalent airway obstruction in the SAPALDIA study were never-smokers[18]. Other potentially relevant exposures comprise respiratory infections in infancy, tuberculosis infection, exposure to household dust and fumes from cooking and heating, industrial and occupational dust and fume, as well as environmental tobacco smoke and ambient air pollution[61,62]. Among these, consistent associations

with COPD-risk were found for workplace exposures to dusts and chemicals[63-65]. There is also growing evidence for an important role of environmental tobacco smoke (ETS)[66], besides its known effects on respiratory symptoms and lung function[67,68]. Finally, evidence is also emerging that ambient air pollution might contribute to COPD development[69,70].

The effect of these alternative risk factors on COPD development is substantial due to their large exposure prevalence. Thereby indoor air pollution from burning of biomass fuel for heating and cooking constitutes a particularly important source of exposure in developing countries[61,71,72], whereas in industrialized countries environmental tobacco smoke, occupational exposures, and ambient air pollution from industry, heating and traffic predominate. The role of these latter factors is however increasing in many developing countries due to their transition to market economies and a westernized lifestyle. From a public health perspective, ambient air pollution is of particular interest due to its unavoidability and high exposure level in densely-populated urban centers caused by production facilities, heating sources and traffic. Moreover, no biological exposure threshold for harmful effects of ambient air pollution has been identified to date[73].

It is however important to note that by far not all exposed individuals develop disease. In the case of smoking for example, it has been estimated that only about 15% of smokers[22] eventually develop COPD. Data for the other exposures is not available. It is thus thought that the individual risk to develop disease depends on the underlying susceptibility to noxious exposures, which is in turn defined by the individual make-up of the genes coding for protective enzyme systems in the body.

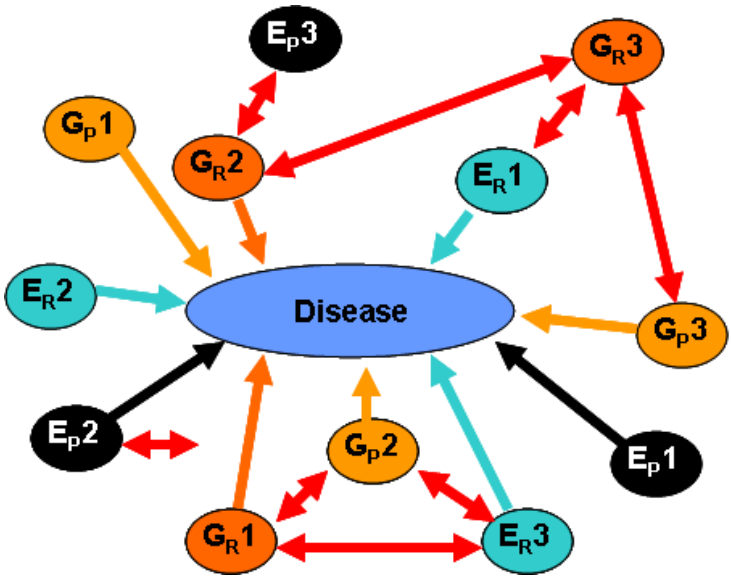
1.1.8.2. Individual susceptibility and the role of oxidative stress genes

COPD thus belongs to the group of complex diseases, in which single environmental exposures or genetic mutations are not sufficient to cause disease, but a complex network of interactions between environmental and genetic factors forms the basis for underlying susceptibility. Each genetic alteration thus only contributes a small part of the overall risk, and possibly only in interaction with an environmental exposure **(figure 2)**. Such a "polygenic disease model" represents the current understanding of chronic disease causation in genetic epidemiology[74].

Protease/anti-protease imbalance hypothesis

Genetic studies have so far made promising attempts to uncover the genetic factors related to COPD. The most long standing and widely recognized genetic determinant of COPD is α1-antitrypsine deficiency, which is caused by genetic mutations in the SERPINA1-gene[75,76]. In the most severe-form, this mutation leads to a 90% loss of function of the enzyme α1-antitrypsine[77], which normally inhibits neutrophil elastase present in lung tissue, and hence to degradation of elastic lung tissue and emphysema. This severe form is however rare, and accounts for only about 1-2% of COPD cases[77,78], which cluster in families. More importantly, the process only manifests with concomitant exposure to tobacco smoke, as the inflammatory reactions induced by tobacco smoke attract sufficient quantities of neutrophils into the tissue[79]. The interaction between mutations in the SERPINA1-gene and smoking represents the classic paradigm of gene-environment interactions, which is now known for several decades[75]. It led to the formulation of the "protease/anti-protease imbalance" hypothesis in

Figure 2 Polygenic disease model



G_P: genetic protective factor; **G_R**: genetic risk factor; **E_P**: environmental protective factor; **E_R**: environmental risk factor

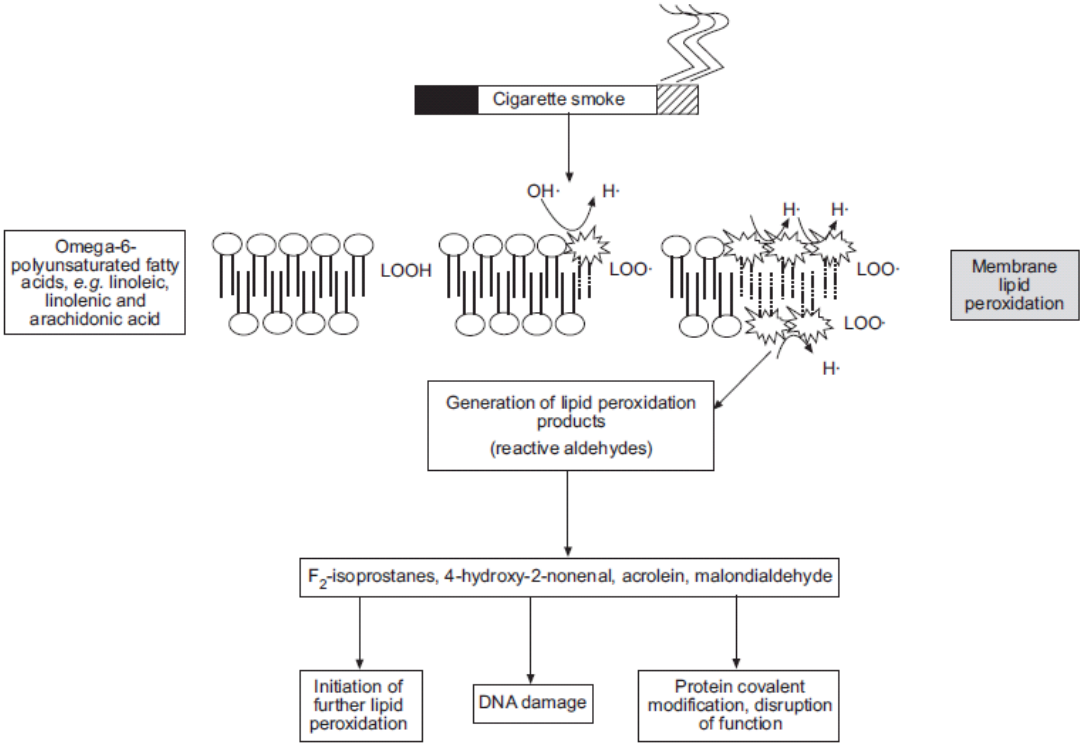
the causation of COPD[80], meaning that the imbalance between deficient α1-antitrypsine activity and excessive release of elastase by inflammation activated neutrophils causes emphysema in lung tissue. On the basis of this hypothesis candidate gene studies indentified further genes involved in lung tissue and extracellular matrix maintenance as risk factors[79].

Oxidant/anti-oxidant imbalance hypothesis

The other important enzyme systems potentially involved in the development of COPD are the anti-oxidant and inflammatory pathways in the body, which are highly activated upon exposure to tobacco smoke and other inhalants like air pollution that impose oxidative stress on the lungs[81]. These exposures either directly contain high amounts of free radicals (highly-reactive substances containing free, unpaired electrons) or compounds which trigger their extensive formation in the body via a cascade of inflammatory reactions and consecutive activation of macrophages and neutrophil leucocytes that normally protect the body against infection-related tissue damage[82]. On the cell level, exposure to free-radicals leads to peroxidation of outer cell walls, further propagation of reactions into the cell, and ultimately protein and DNA alterations **(figure 3)**, which damages the cell and further enhances inflammation[83]. The extent of these reactions will thereby depend on the individual capacity to handle the oxidative stress burden, and mutations affecting the activity of pivotal enzymes in these cascades are likely to influence the risk of COPD substantially[82]. Thus, analogously to the protease/anti-protease hypothesis the other large paradigm in COPD-development is an imbalance between the oxidative stress burden and the activity of protective anti-oxidant systems[84].

Though the two hypotheses are often separately referenced and investigated, they are biologically linked via inflammatory cascades and the secretion of pro-inflammatory substances from white blood cells triggered by oxidative stress (**figure 4**).

Figure 3 Cascade of peroxidation reactions in the cell upon exposure to free radicals, from [83]



Candidate gene studies

As a consequence of both, the oxidant/anti-oxidant and protease/anti-protease imbalance hypotheses, many genes involved in inflammation (like *tumor necrosis factor alpha* or *interleukins*), oxidative-stress (like *gluthathione s-transferases*, *microsomal epoxide hydroxylase*, and *heme-oxygenase 1*), and maintenance of the extracellular matrix (like *matrix metallo-proteinases*) have been investigated and associated with COPD or related traits in candidate gene studies[85-87]. However, only a small part of these findings could be consistently replicated in later studies and a recent meta-analysis across different studies and populations[87,88]. According to the meta-analysis, among the genes showing consistent associations with COPD were *transforming growth factor, beta 1 (TGFB1)*, *interleukin 1 receptor antagonist* (*IL1RN)*, and *tumor necrosis factor alpha TNF-α*[87,89].

Genome-wide association studies (GWAS)

Due to technological developments costs of genotyping have progressively declined during the last decade, and gene-chips covering several hundred thousand mutations dispersed over the whole genome have become available for epidemiological research. As a consequence, the search for genetic determinants of COPD was extended to the whole genome in recent years. Analysis was carried out in a hypothesis-free manner by association-testing of all available variants on the gene-chip with COPD in the purpose of identifying new, previously unsuspected genes and pathways.

Figure 4 The role of proteases, oxidants and inflammation in COPD causation, from [82]

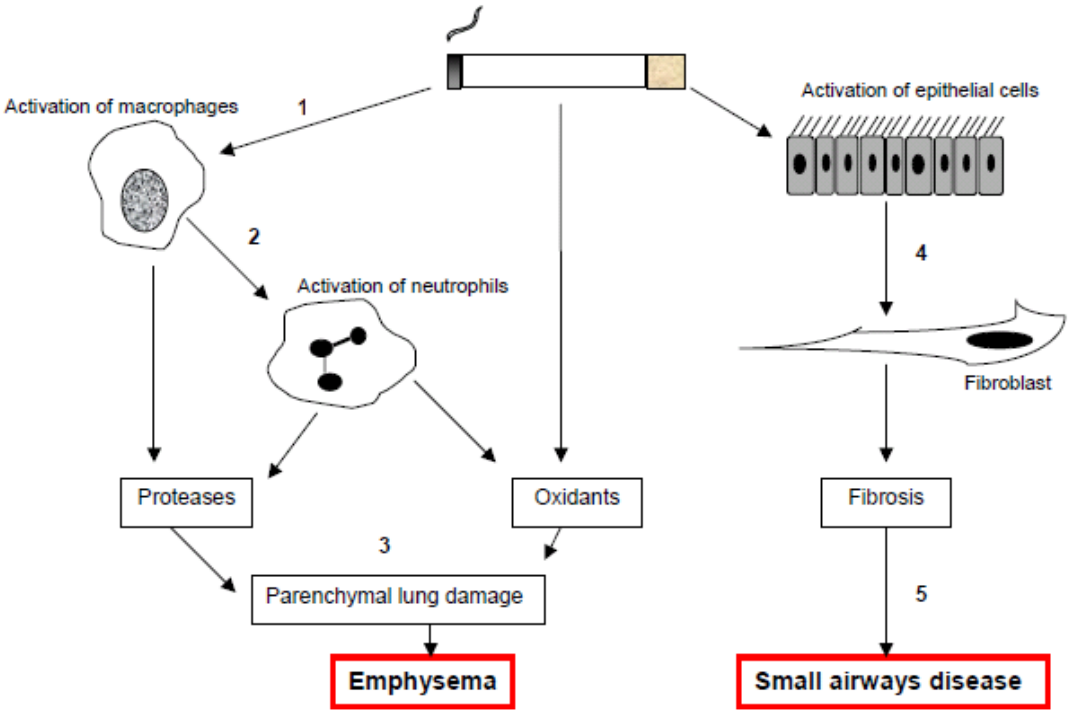


Figure 1
The pathogenesis of COPD. Cigarette smoke activates macrophages (1), leading to the direct release of proteases or neutrophil chemotracctants (2), together with the release of oxidants resulting in subsequent breakdown of connective tissue in the lung (3), causing emphysema. Epithelial cell stimulation promotes fibroblast activity (4), eventually leading to small airways disease (5).

Capitalizing on this agnostic nature, genome wide association studies (GWAS) have indeed suggested new genes and genomic loci affecting the risk of COPD. The set of identified genes/loci comprises *hedgehog interacting protein HHIP*, a locus near the *alpha-nicotinic acetylcholine receptor genes3 and 5 CHRNA3/5* or *iron responsive element binding protein 2 IREB2* (both located in the same linkage-disequilibrium block), and *family with sequence similarity 13 member A FAM13A* (which was also associated with hypoxia and lung function)[90,91]. Further, a meta-analysis of two large GWAS on cross-sectional lung function from the European SpiroMeta and North-American CHARGE consortia[92,93] identified loci at *FAM13A*, *serotonin 4 receptor subtype (HTR4)*, the *receptor for advanced glycosylation end products (AGER)* , *thrombospondin type-1 domain-containing protein (THSD4)* and a locus near *HHIP* to be associated with FEV_1/FVC, and *tensin 1 (TNS1)* as well as *glutathione S transferase C-terminal domain (GSTCD)* with FEV_1. A recent follow-up analysis could also confirm their association with COPD[94].

However, it is currently not clear by which mechanism these genetic variants influence the risk of COPD on the molecular level, as the identified variants for both lung function and COPD were hardly related to oxidative stress defense or the proteinase/anti-proteinase system. Further, a subsequent analysis focusing on previously defined candidate genes for COPD and lung function within the dataset of the SpiroMeta consortium failed to replicate previous candidate gene associations with the disease[95]. Moreover, the overlap between the most strongly associated SNPs from the cited genome-wide studies on COPD and lung function was small.

Gene-environment interaction

In summary, the findings from both, candidate gene studies and GWAS on COPD were not consistent, and non-replication of previously identified associations was frequently observed. A subsequent analysis of the COPD GWAS by Pillai and colleagues[96] implicated that the three genomic regions might differentially affect different sub-phenotypes of COPD: while all were to some degree associated with airway obstruction, the variant near *CHRNA3/5* was also associated with smoking intensity and emphysema, and *HHIP* with fat free mass and body mass index (BMI). This finding underlines the importance of a detailed phenotypic characterization in COPD studies, and might explain a part of the observed non-replication of study findings, as genetic heterogeneity across disease subphenotypes might level out gene effects if not properly accounted for. Alternatively, non-replication could also be due to lack of statistical power or genetic heterogeneity in study populations of differing ethnic origin. But these factors are not likely to operate in large scale GWAS, that regularly control for ethnic background of study participants in the analysis, where the issue of non-replication is nevertheless present.

Another important explanation of non-replication could be presence of gene-gene- and gene-environment interactions[97,98]. Interestingly, not many studies on possible interactions between genetic and environmental factors have been conducted in the search for genetic determinants of COPD[75,99], although the disease represents the classic paradigm of gene-environment interaction since the association of lung function, α1-antitrypsine deficiency and smoking has been described 50 years ago[75,77].

Variants in several candidate genes belonging to the first line defense against oxidative stress have been investigated regarding their interaction with air pollution on different respiratory outcomes in the past, including asthma, respiratory symptoms and lung function[100]. Often studied genetic variants included SNPs in *glutathione S-transferases class M1* (*GSTM1)*, *P1* (*GSTP1*)and *T1* (*GSTT1)*, *NAD(P)H dehydrogenase 1 (*NQO1*)*, *superoxide dismutase 2 (SOD2)*, *glutathione peroxidase 1 (*GPX1*)*, *heme oxygenase-1 HMOX-1*, *epoxide hydrolase 1 (EPHX1)* and *catalase (CAT).* A systematic review of studies published up to April 30th 2009 investigating interactions between genetic variants and ambient air pollution on symptoms, lung function and asthma found evidence for the presence of gene-environment interactions in oxidative stress systems, particularly in populations of children who are more susceptible[101]. However, the strength of evidence was hampered by potential issues regarding multiple testing correction, selective reporting and inconsistent directions of association in the respective studies[101]. This is in line with a concurrent review on gene air pollution interactions focusing on asthma outcomes that was largely based on the same underlying studies[102]. Similarly, interactions have been described for oxidative stress genes with tobacco smoke exposure on lung function, and for inflammatory genes regarding lung function decline in adults with airway obstruction[103-106].

In contrast, no genomewide study on COPD was currently identifiable in the literature that assessed gene-environment interaction, and only one genome-wide study on childhood asthma assessed interaction with farming exposure in the whole field of respiratory epidemiology to date[107].

This lack of gene-environment interaction studies likely reflects the higher sample size requirements for successful investigation of gene-environment interactions compared to gene-main effects studies. The issue is further aggravated when performing GWAS, that typically require several thousand cases and controls to cope with multiple testing correction after genome-wide analysis of genetic main effects already[108,109]. Potential higher order interactions between genes and environmental factors can further limit the power of classic analysis methods[108,109]. In recognition of these methodological difficulties, large scale collaborations to examine the mutual roles of genes, environmental factors and their interaction in COPD and other chronic respiratory diseases have been proposed[110]. But these need to be accomplished with a strong focus on data comparability to avoid spurious associations introduced by differential measurement error between studies. The assembly of large

datasets or meta-analysis of interactions across several studies thus requires comparable data quality also on the environmental side, which is not easy to achieve in the field of air pollution exposure measurement.

In summary, in view of the large Public Health burden and limited therapeutic options to treat COPD, it is important to pursue the investigation of gene-environment interactions, as uncovering the complex network of interacting genetic and environmental factors offers the possibility to develop new and urgently needed therapeutic and preventive measures.

1.2. Air Pollution

1.2.1. Air pollution as environmental threat to health

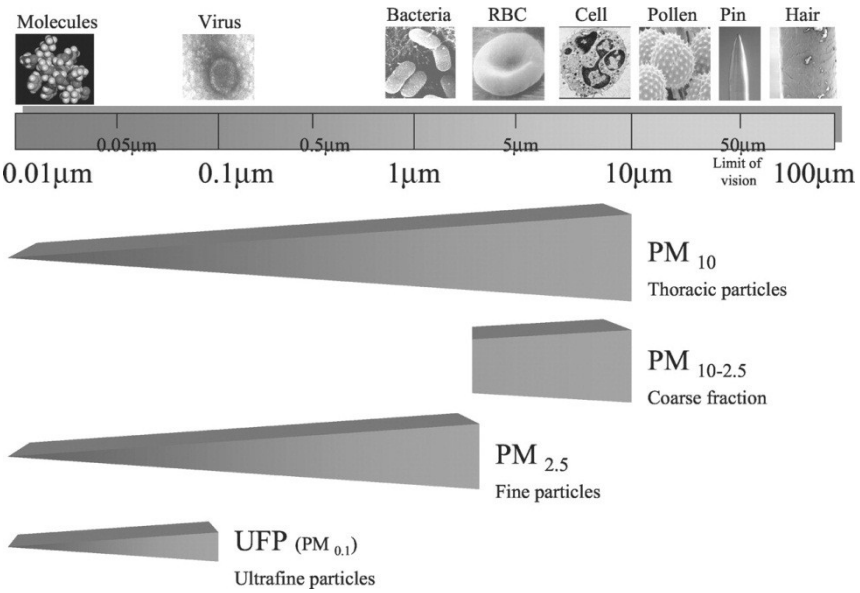
Since the "Great Smog" in London triggered a high number of emergency visits and excessive deaths in 1952, the health effects of ambient air pollution have received attention environmental health research. This excessive pollution episode occurred in December 1952 due to a stagnation of the weather conditions and concentration of pollutants like soot and sulfur dioxide (SO2) originating from heating and combustion sources[111].

By the 1980s, clean air legislation and technological changes have substantially lowered the amount and changed the composition of air pollution, with smoke and sulfur components loosing importance compared to ozone (O3), nitrogen dioxide (NO2) and particulate matter (PM). But the results of three large Unites States (U.S.) cohort studies, the Harvard Six Cities study[112], the American Cancer Society study[113] and the Seventh-day Adventist study[114], brought the health effects of air pollution again to the attention of policy makers, the public as well as the research community. They provided robust evidence that the relatively low particulate matter exposure present at that time was still associated with overall mortality. Thereby, exposure contrasts in the range observed within the studies potentially entailed differences in life expectancy of 1-2 years on average. These observations were intensely questioned from the side of industry and respective stakeholders, but withstood detailed scrutiny[115,116]. The reanalyses eventually even helped to further improve the methodology of studying air pollution effects[117]. As a consequence, polluted air was de novo perceived as one of the main environmental threats to human health, which considerably fuelled the research in the field.

1.2.2. Sources, composition and exposure assessment

In the context of air pollution studies, it is important to consider that air pollution is actually a complex mixture of exposures[111]: it consists of gaseous pollutants such as NO2, SO2 and O3, as well as liquid and solid state components constituting particulate matter exposure. This has important implications for research, as the gaseous fraction is mostly generated locally by combustion processes (e.g. from traffic or power plants) and ensuing chemical reactions (e.g. synthesis of O3 by reaction of NO2 with hydrocarbons under sunlight exposure), where it also exerts its health effects before being degraded. Particulate matter (PM) relevant to human health in the context of air pollution consists of the particle size fraction with median diameters of 10µm or less (PM_{10}). These particles remain suspended in the air after generation and are breathable, i.e. they enter and deposit in the airways and lungs. Different particle sizes have been defined according to the preferential place of deposition in the airways (**figure 5**): Particles with median diameter of 2.5-10µm ($PM_{2.5-10}$) deposit in the trachea and pulmonary bronchi, while the fine fraction (PM2.5, median diameter less than 2.5 µm) can also deposit in the lung alveoli[118]. Ultrafine particles (UFP, median diameter less than 100nm) can penetrate the alveolar wall and reach the blood circulation[119,120], and there are even studies showing their deposition in the brain via the nasopharynx and olfactory nerves[121].

Figure 5 Fractions of particulate matter air pollution, from [118]



Particles are generated in different ways: coarse particles (median diameters between 2.5-10 µm) are formed by mechanical abrasion and attrition of diverse surfaces in agriculture, on construction sites and during transport including asphalt, tires, and rubber wear off. The finer portion $PM_{2.5}$ (less than 2.5µm in diameter) originates mostly from incomplete combustion processes of motor vehicle engines, power generation, industrial activities, but can also arise from condensation of liquid drops and by chemical reactions of gaseous substances. The composition of particles varies depending on the source and can consist of a mixture of organic compounds, inorganic mineral dusts, heavy metals, acids or even biological specimens such as bacteria.

The concentration of gaseous and particulate matter pollutants can be measured using devices that either sample the air over longer time spans while retaining the pollutant for later laboratory analysis, or that directly measure their concentration in the air. This is achieved using chemo luminescence methods for gaseous, and gravimetric methods for particulate pollutants.

For the estimation of personal exposures from environmental measurements the physicochemical properties of gaseous and particulate matter pollution play a crucial role. The most important difference is that in contrast to gaseous pollutants which are volatile and undergo chemical reactions with other atmospheric components, particles are relatively inert, remain suspended and are further dispersed by air movements. They can be transported over several hundred kilometers by atmospheric flows. Pollution hotspot like large industrialized cities can thus induce health effects over wide geographic areas. From a methodological point of view, this makes it easier to obtain sufficiently accurate exposure estimates for particulate matter in epidemiological studies than for gaseous pollutants. Accordingly, most of the robust and consistent associations with adverse health effects have been described for particulate matter exposures (see next paragraph). Similarly, sophisticated methods to improve the accuracy of personal exposure estimates have mainly been developed for particulate matter pollution. A more detailed description of available measurement methods and currently employed models for exposure estimation is given in Appendix 2.

1.2.3. Air pollution health effects

Research on air pollution related health effects has concentrated on two main lines: short and long term exposure effects. Studies investigating short term effects have thereby usually looked at how variation in day-to-day air pollution values correlated with daily rates in the outcome of interest in so-called time-series analyses, whereas long term studies investigated air pollution exposure over longer periods using traditional study designs and statistical methods.

1.2.3.1. Short term effects

Regarding short time effects, many smaller panel studies and two large scale cohort studies from Europe and the U.S. have shown positive associations with all-cause, respiratory and cardiovascular mortality as well as hospital admissions for cardiovascular and pulmonary reasons[122,123]. An increase in PM_{10} exposure of 10µg/m^3 was associated with an increase in daily mortality of about 0.6% (95%-confidence interval (CI) 0.4-0.8%) and increases in hospitalization rates for asthma and COPD in elderly persons between 1.0-1.5%, and 0.5-1.1% for cardiovascular disease[111,124-127]. Similar associations have also been found in these studies for the gaseous pollutants NO2 and O3, but not after adjusting for particulate matter. This illustrates the frequently present difficulty to disentangle the causal substance in air pollution research due to their mutual correlation. Positive associations were also found with symptoms and exacerbations of asthma and COPD[123,128]. There is also evidence that short term increases in particulate matter air pollution lead to elevated arterial blood pressure, and can trigger myocardial infarction, ischemic stroke, and hospitalizations for cardiovascular causes[122,123]. For other, subclinical outcomes like alterations in heart rate variability, ischemic ST-segment depressions or myocardial repolarization disturbances results were variable and sometimes inconsistent[122].

1.2.3.2. Long term effects

Long term exposure to particulate matter air pollution (PM_{10} and $PM_{2.5}$) was also associated with all-cause, cardiovascular and pulmonary mortality[112,113,122,129], reduced lung function and symptoms of chronic bronchitis in cross sectional studies of adults including SAPALDIA[123,130,131]. Increase has also been related to worsening of lung function and respiratory symptoms in asthma and COPD patients[123,132,133]. Further, a reduction of ambient PM_{10} exposure was associated with an attenuation of the natural lung function decline in the SAPALDIA study[69]. First associations with asthma incidence in adults were found for traffic air pollution[133]. On the cardiovascular side, there is favorable though not conclusive evidence for associations of $PM_{2.5}$ exposure with exacerbation of congestive heart failure, possibly incidence of stroke and non-fatal myocardial infarction, as well as progression of atherosclerosis[122]. In children, exposure to air pollution has been related to lower lung function, respiratory symptoms, incident asthma, and also attenuated lung function growth[134-138]. The latter is particularly important, as lung growth reaches a plateau phase about the age of 25 years, after which the natural, slow decline in lung function sets in. High air pollution exposure in the growth phase thus

entails a lower start point for decline, and might constitute higher susceptibility to respiratory disease in later life.

1.2.3.3. Susceptibility

Based on the results from epidemiological studies, it appears that the parts of the population which are most susceptible to the deleterious health effects of air pollution comprise children and the elderly, persons with pre-existing cardiopulmonary diseases or diabetes, and low socio-economic status[122,123,139].

1.2.4. Biological mechanisms

Particulate matter and gaseous pollutants induce oxidative stress upon exposure in the lungs, with consecutive inflammatory reactions, which are potentially also propagated to the rest of the organism[111,122,140,141]. NO2 has also been described to impair the function of alveolar macrophages and epithelial cells, which could lower the protection against infections and might be a mechanism explaining the observed associations with exacerbations of respiratory diseases[111]. Oxidative stress, inflammation and in the long term also tissue remodeling is responsible for many of the symptoms and morbidity seen at the lung level. Possible mechanisms for the cardiovascular effects include spillover of inflammation from the lung to the systemic circulation[118,122,141] (with consecutive alteration of endothelial function, dysbalances in blood coagulation, and inflammatory reactions in vessel walls), triggering of pulmonary reflexes impacting on the autonomous nervous system (and hence heart rate variability), or direct effects via penetration of fine and ultrafine particles into the blood stream.

1.2.5. Public health significance

The large, consistent and growing epidemiological evidence that air pollution exposure is associated with important adverse health outcomes such as overall and cardiopulmonary mortality, morbidity, and health service utilization makes clear that air pollution is one of the major Public Health issues today. Two other aspects further stress the Public Health significance of air pollution, the first being that up to now, no clear biological exposure threshold has been identified for adverse health effects, especially regarding particulate matter exposure[73,111,118]. Second, the fact that air pollution exposure cannot be avoided by affected persons like harmful lifestyle habits, makes it a clear target for community or national level preventive and regulatory action[73]. Air pollution is thus a prominent environmental hazard and an important potential risk factor to study in the context of COPD.

2. Aims

The introductory part on COPD can be resumed as follows: COPD is a disease of large and global Public Health burden in terms of morbidity and mortality. The estimates of the disease's prevalence comprise a wide range, even after standardizing the case definition across different study populations and accounting for differences in their age structure and distribution of tobacco smoke exposure. Population-based data on incidence are very scarce, and completely missing in the case of Switzerland. The unexplained portion of the disease distribution suggests on one side that other important environmental risk factors have not been accounted for, and on the other the presence of individual susceptibility factors that determine the actual risk of disease conveyed by environmental exposures (this is also in line with the observation that by far not all smokers develop COPD). Two main pathways for COPD causation have been proposed comprising protease/anti-protease as well as oxidant/anti-oxidant imbalance. In light of the large oxidative-stress burden many inhalatory exposures impose on the lungs, studies were performed on the effect of variants in genes coding enzymes of oxidative-stress defense or its endogenous production on COPD risk. Their evidence was inconsistent and non-replication frequent. This is also the case for the results gathered by recent, hypothesis-free genome-wide association studies, which successfully identified new, unsuspected genetic risk factors for COPD. An explanation for the observed non-replication could be failure to account for gene-environment interactions.

Ambient air pollution is an important environmental risk factor and could contribute to COPD via oxidative stress reactions. Investigating the role of ambient air pollution in COPD development is of high Public Health interest because of the high exposure prevalence and reduced capacity of individuals to avoid exposure.

In light of the state of research, the thesis work aimed to find answers to the following questions:

I) What is the burden of COPD in Switzerland, what are its main determinants, and how does it impact on individual well-being and use of health service resources?
II) Does ambient air pollution contribute to the COPD burden after accounting for tobacco smoke exposure and age-and sex-related changes in respiratory health?
III) How does the impact of air pollution on COPD development compare with that of tobacco smoke on the population level?

These aims were to be accomplished by assessing the following research questions using data from the population-based SAPALDIA cohort study:

I a) What is the incidence of COPD in Switzerland: can adequate estimates of disease incidence be made on the basis of available epidemiological data?

I b) What are the main determinants of COPD-incidence in terms of age, sex, smoking, chronic bronchitis symptoms?

I c) How is the observed COPD-incidence related to use of health services and individual breathing capacity in our epidemiological study?

II a) Does long-term exposure to ambient air pollution affect the development of COPD and related lung function outcomes, after accounting for smoking and other identified risk factors?

II b) Does variation in genes constituting the body's first-line defense against oxidative stress modify the effects of air pollution on the risk of COPD and related outcomes?

II c) Is it possible to characterize a population subgroup with greater susceptibility to adverse effects of air pollution?

III a) Starting from oxidative stress related genes, can larger biological pathways be identified, variation in which affects the individual susceptibility to air pollution regarding development of COPD or related traits?

III b) Are the biological pathways mediating the effects of air pollution on individual risk of COPD or related traits the same as those activated by tobacco smoke?

III c) How large is the respective impact of pathways mediating ambient air pollution and tobacco smoke effects on risk of COPD and related outcomes in terms of the size of environmental exposure effects and predictive power at the population level?

3. Methods:

3.1. Description of the SAPALDIA study

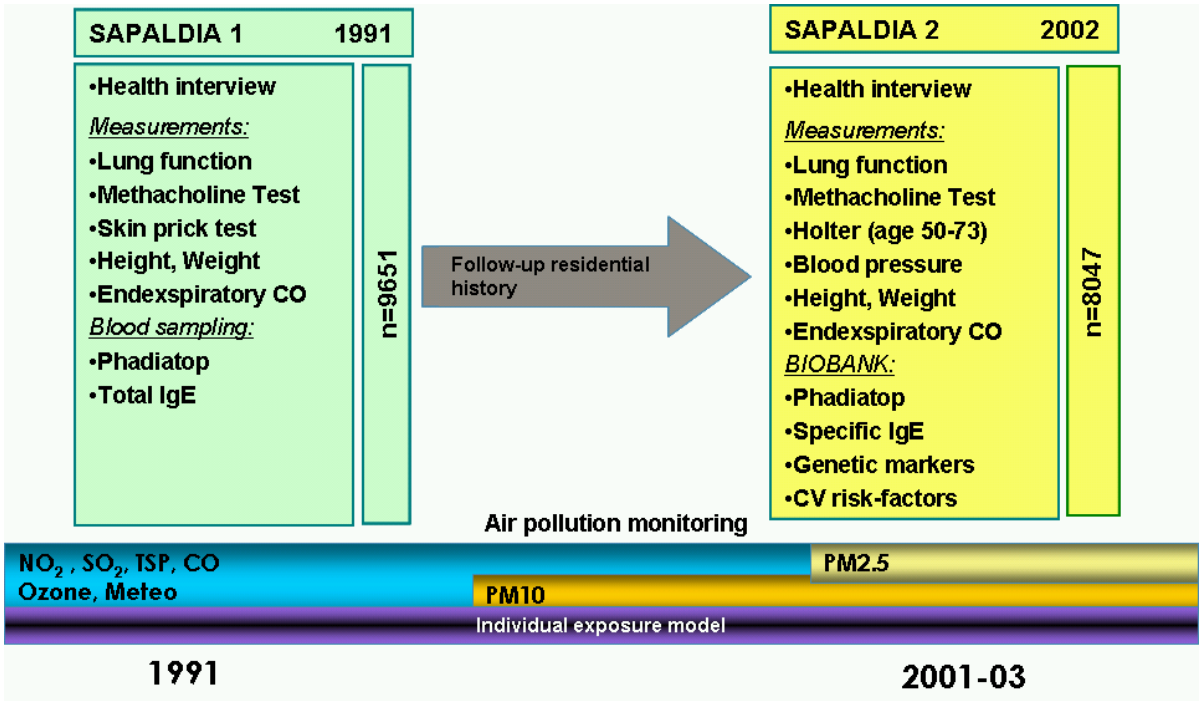
3.1.1. Study design

The SAPALDIA study is a cohort study established in 1991 to study the health effects of long-term exposure to ambient air pollution in a sample of Swiss adults from the general population[16,17]. The study drew upon eight communities with different grades of urbanization that represented the whole range of ambient air pollution exposure throughout Switzerland. Swiss resident adults aged 18-60 years were randomly sampled from the respective population registries and invited for participation. All persons participating in the baseline assessment in 1991 (SAPALDIA 1) were re-invited for follow-up examinations in 2001-2003 (SAPALDIA 2). Meanwhile, a second follow-up examination has been accomplished during 2010/11 (SAPALDIA 3), but the current thesis is based on the data from the first 2 assessments.

The SAPALDIA study was approved by the Swiss Academy of Medical Sciences and the regional ethics committees. All study participants gave written informed consent prior to study examinations.

An overview of the examinations applied in the SAPALDIA study is given in **figure 6**, and details on study procedures and protocols have been published in methodological papers accompanying the study assessments[16,17]. In the following paragraphs, only a description of the study examinations relevant to the thesis will be given.

Figure 6 Overview of measurements in the SAPALDIA study



3.1.2. Questionnaires

A computer-assisted interview was performed at both examinations. The interview comprised questions about presence of respiratory symptoms and allergic diseases, family history of respiratory disease, personal smoking history, exposure to environmental tobacco smoke at home or at the workplace, exposure to dust and fumes at work, medication use, living conditions, and socio-economic and demographic factors. In the follow-up assessment, additional questions on chronic illness including heart disease, dietary habits and physical fitness were included. Questions targeting time changing exposures and characteristics were asked identically at both examinations. An overview of the questionnaire used in SAPALDIA 2 is given in Appendix 3.

Definitions related to smoking

Based on the answers concerning smoking history, participants were defined as never smokers if they had smoked less than 20 packs of cigarettes or 360g of tobacco in their lifetime. Smokers reported current, active smoking at the time of interview, and ex-smokers were defined as those who had quit smoking at least one month before the examinations[142]. For current and ex-smokers, pack years were calculated as the number of cigarette packs consumed per day multiplied by the years of consumption. Thus a person with a cumulative tobacco smoke exposure of 2 pack years has for example smoked a pack of cigarettes a day over the previous 2 years, while another one having 20 pack years of exposure might have smoked 2 packs per day over 10 years.

3.1.3. Lung function measurements

Lung function testing was performed at both examinations according to the protocol developed in the ECRHS study[143]. Participants were in an upright sitting position and performed three to eight forced expiratory spirometry maneuvers complying with quality criteria as set out by the American Thoracic Society (ATS) in 1994[144]. Lung function testing was done without broncho-dilation due to resource but also time constraints, as methacholine tests to assess airway reactivity were foreseen in both examinations. Spirometry devices were checked regularly and calibrated on a daily basis. The same devices were used at the baseline and first follow-up examination eleven years later (Sensor Medics 2200 SP, Sensor Medics, Yorba Linda, USA), and comparability of spirometry measurements was assessed at each time point[145,146]. In the testing procedure, measurements of lung volumes FEV_1 and FVC were recorded, as well as the airflow velocity in the mid-portion of FVC between its 25^{th} and 75^{th} percentile (FEF_{25-75}).

3.1.4. Air pollution modeling

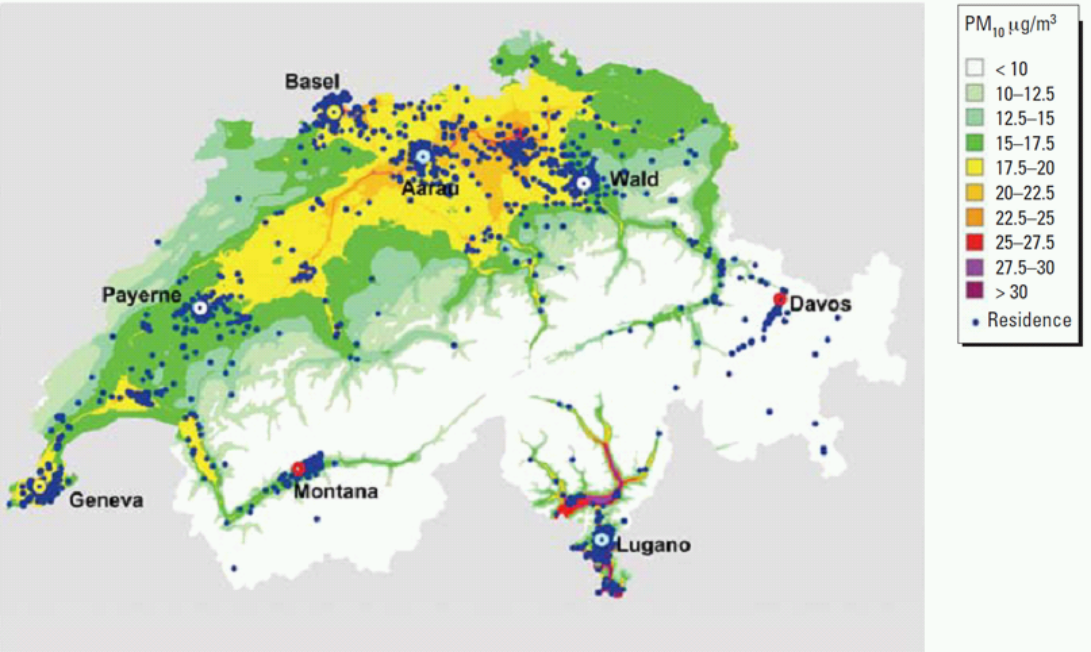
Air pollution exposure data from the beginning of the SAPALDIA study throughout the first follow-up period up to 2002 was based on central site air pollution measurements from monitors. Exposure resolution was thus limited, as only a small number of monitors were operating per study site, and all participants residing within a certain distance from the monitor were assigned the same, measured

exposure level. Accordingly, in each study site, gaseous (e.g. SO4, NO2, NOx) and particulate air pollution (total suspended particles (TSP) or PM_{10}) was measured at central monitoring sites maintained by national and cantonal air hygiene authorities. In two rural study areas additional measurements of the fine fraction of particulate matter pollution (i.e. $PM_{2.5}$) were carried out by the SAPALDIA study from 1999 onwards, as no official measurements were available. To better capture spatial exposure resolution, in- and outdoor NO2-concentrations at the home addresses were additionally measured by Palmes tubes for a selected subsample of study participants[147]. Further details of air pollution monitoring data are available in the methodological publications accompanying each assessment[16,17].

From the first follow-up assessment onwards, air pollution exposure assessment was much improved by the use of individual exposure estimates[148]. These were based on the PolluMap model version 2[149,150], a Gaussian dispersion model which predicted mean average annual exposures to PM_{10} and NO_2 for the years 1990 and 2000 with a 200m grid resolution throughout Switzerland **(figure 7)**. The estimates were based on an emission inventory comprising transport, industrial, commercial, and construction sources, household heating, and agricultural and forestry activities. The dispersion of pollutants across geographical areas was modeled as Gaussian distribution taking into account Swiss topography, the emission source type and height above ground, as well as meteorological data. A detailed comparison of the exposure estimates predicted by the model with those measured at air pollution monitoring sites showed that the model produced robust estimates for individual PM_{10} exposure even in highly exposed sites, that can be used for health assessments. The model probably profited from the fact that particulate matter is transported over long ranges spanning several kilometers and shows only limited variability on short distances[148]. On the opposite, the dispersion model insufficiently captured the high spatial variation typically found in traffic related pollutants like NO_2, and the respective model estimates were thus only comparable with measured ones in background sites. Details of the model evaluation have been published by Liu and colleagues[148].

Individual exposures were derived by mapping the geo-referenced residential address of each study participant to the corresponding model grid cell. Changes in residency were thereby accounted for. Additionally, for PM_{10}, individual annual average exposure estimates were derived for all years between 1990 and 2002 by using an algorithm that interpolated the model estimates from 1990 and 2000 based on historical trends from fixed monitoring stations[69]. Individual longitudinal exposure estimates were then estimated in terms of the change in PM_{10} exposure between baseline and follow-up examination, calculated as the difference between the mean annual exposures in 2002 and 1990. A detailed residence history of each participant also allowed estimating a cumulative PM_{10} exposure value by summing mean annual exposure estimates for each study year and respective grid cell of the participant's residency. This resulted in exposure estimates of $\mu g\ PM_{10}/m^3$ * years. A comparison of effect estimates on lung function[69] showed that the estimates for change and cumulative exposure to PM_{10} are highly correlated, and cannot always be disentangled in analytical studies.

Figure 7 Dispersion model estimates for home outdoor PM_{10} exposure in 2002, from [148], and adapted from [149]



Blue dots represent the residential addressed of study participants at SAPALDIA 2, clustered around the 8 study areas

3.1.5. Genetic data

Blood samples of 45ml were taken at the follow-up examination from all participants who had given consent for serum or genetic analyses[16]. Each blood sample was partitioned into 40 aliquots if possible, resulting in a blood bank of about 250000 blood aliquots. For safety reasons the blood bank was split into two identical parts, which were kept in two different locations (University Hospitals of Geneva and Zürich).

7ml of whole blood was collected for later DNA-extraction into EDTA-buffered tubes and stored at minus 80°C. In preparation for analysis, DNA was extracted manually using the PUREGENE purification kit (GENTRA Systems, Minneapolis, USA), and 1ml of frozen EDTA-blood yielded between 7-80µg DNA of high quality. DNA working solutions with concentrations of 10ng/µml were generated and kept at minus 20°C for long term storage.

Ensuing genotyping for health related analyses was guided by published associations in genetic epidemiology studies and molecular biology studies related to possible health effects of air pollution. Likewise, genotyping results are available for many polymorphisms in oxidative-stress genes on which associations with respiratory or cardiovascular diseases have been published[101,102,151]. Genotyping of variants in glutathione S-transferases and heme-oxygenase 1 genes, which both belong to the lungs first line defense systems against oxidative stress, is of most relevance for the work in this thesis.

In the framework of the GABRIEL study[152], a large collaborative effort funded by the 7th European Framework Program to uncover the genetic and environmental determinants of asthma, genome-wide genotyping of all self-declared asthmatics (n=663) and a random sample of never-asthmatics (n=997) was obtained in 2009. Genotyping was done on the Illumina Human 6010quad BeadChip comprising 610'000 single nucleotide polymorphisms (SNPs) dispersed throughout the genome.

After excluding samples with <97% genotyping success rate, non-European origin, cryptic relatedness or sex-inconsistencies, genome-wide data was available for 1457 participants, comprising 878 non-asthmatics and 579 asthmatics.

From the 610000 SNPs covered by the chip, 567589 autosomal SNPs passed strict quality control filters, including Hardy-Weinberg equilibrium p-value<10^{-4}, genotyping call rate <97%, and minor allele frequency (MAF) <5%. All SNPs were used for imputation to 2.5 Mio SNPs using MACH v 1.0 software[153]. The most comprehensive data on common genetic variation in humans stems from the HapMap project[154], which applied high resolution genotyping in four populations of different ethnic origin (North Americans of European origin, Japanese, Han Chinese, and Africans). The HapMap version 22 CEPH panel of Utah residents with ancestry from northern and western Europe[154] served as a reference frame for imputation, giving 2'168'681 SNPs with good imputation quality and minor allele frequencies >5%.

3.2.Implications for current research

The characteristics of the available data in the SAPALDIA study entail the following implications for investigating the research questions stated in section 2, page 33:

I) As only pre-bronchodilation spirometry results are available, no strict definition of COPD is possible (this requires broncho-dilation before lung function testing). As a consequence, the outcomes which are amenable to investigation are

a. a modified definition of COPD: the GOLD COPD disease definition was applied to pre-bronchodilation spirometry results. Thereby a vigorous control for known as well as hidden asthma in the analysis is necessary, as well as a sensitivity analysis using FEV_1/FVC < lower limit of normal to define airway obstruction.

b. lung function decline: decline is calculated by subtracting the baseline from the follow-up spirometry measurement. Accelerated lung function decline is a proxy measure and cardinal feature of COPD, with decline in the FEV_1/FVC ratio preceding airway obstruction, while an accelerated decline in FEV_1 determines severity in presence of obstruction. Further, decline in FEF_{25-75} can serve as proxy for processes in the small airways. Lung function declines were also studied in the case of interaction testing, as a continuous outcome measure offers higher statistical power than a binary outcome.

II) To assess the effects of air pollution on COPD/lung function decline and its potential interaction with oxidative stress related genes, only individual PM_{10} exposure estimates could be used, since the validity of NO2 estimates was not sufficient to warrant a use in health related analyses. Further, the installment of a stricter clean air policy during the 1990s led to a decline in air pollution levels throughout Switzerland between the baseline and follow-up examination. Our estimates of change in exposure between surveys were thus on average negative. Previous work in our study had shown that a larger decline in PM_{10} exposure during follow-up was associated with attenuation in the normal, age-related lung function decline[69].Thereby the effects of an exposure change could not be distinguished from those of cumulative exposure, as the two were highly correlated. These factors were considered in the analysis of a possible interaction between air pollution and oxidative stress genes on lung function decline.

III) To assess the role of biological pathways comprising oxidative-stress related genes, only 878 non-asthmatic adults with genome-wide data were available. Focusing on oxidative-stress relevant genes and pathways still entailed the investigation of over 12000 SNPs. Thus the small sample size compared to the large number of variables represented a problem of high data dimensionality with the main issue being low statistical power. This could not be addressed by classic statistical analysis methods, but required applying modern pathway analysis methods, which make use of pre-existing biological knowledge to integrate lower level association signals (from SNPs) onto the upper biological levels (genes and pathways). Such an approach results in fewer data dimensions, and higher statistical power.

4. Paper 1:
Assessing the burden of COPD in Switzerland, its determinants, and its impact on individual well-being and health service utilization. (aim I)

4.1. Longitudinal change of prebronchodilator spirometric obstruction and health outcomes: results from the SAPALDIA cohort.

(published in: Probst-Hensch NM, **Curjuric I**, Pierre-Olivier B et al. Longitudinal change of prebronchodilator spirometric obstruction and health outcomes: results from the SAPALDIA cohort. Thorax 2010; 65(2): 150-6. Equal first authorship with Prof. Probst-Hensch NM.

Article link: http://thorax.bmj.com/content/65/2/150.full?sid=a6f06257-f9fd-4100-b746-857946480fb7)

LONGITUDINAL CHANGE OF PRE-BRONCHODILATOR SPIROMETRIC OBSTRUCTION AND HEALTH OUTCOMES – RESULTS FROM THE SAPALDIA COHORT

Authors:

NICOLE M. PROBST-HENSCH[*,1] PHD, IVAN CURJURIC [*,1, 3] MD, BRIDEVAUX PIERRE-OLIVIER[2] MD, PHD, URSULA ACKERMANN-LIEBRICH[3] MD, ROBERT W. BETTSCHART[4] MD, OTTO BRÄNDLI[5] MD, MARTIN BRUTSCHE[6] MD PHD, LUC BURDET[7] MD, MARGARET W. GERBASE[2] MD PHD, BRUNO KNÖPFLI[8] MD, NINO KÜNZLI[3] MD, PHD, MARCO G. PONS[10] MD, CHRISTIAN SCHINDLER [3] PHD, JEAN-MARIE TSCHOPP[11] MD, THIERRY ROCHAT[2] MD, ERICH W. RUSSI[12] MD

**Probst/Curjuric: equal first author*

Affiliations: [1] Chronic Disease Epidemiology , Institute of Social and Preventive Medicine, University of Zürich, Switzerland; [2] Division of Pulmonary Medicine, University Hospitals of Geneva, Geneva, Switzerland; [3] Institute of Social and Preventive Medicine at Swiss Tropical Institute, Basel, Switzerland; [4] Lungenzentrum, Hirslanden Klinik Aarau, Switzerland; [5] Zürcher Höhenklinik Wald, Switzerland; [6] Division of Pulmonary Medicine, Cantonal Hospital of St. Gallen, St. Gallen, Switzerland; [7] Hôpital Intercantonal de la Broye, Payerne, Switzerland; [8] FMH Pädiatrie, Pneumologie und Sportmedizin, Davos, Switzerland; [10] Dipartimento di Medicina, Ospedale Regionale di Lugano, Lugano, Switzerland; [11] Centre Valaisan de Pneumologie et Département de Médecine Interne du Centre Hospitalier du Centre du Valais, Crans-Montana, Switzerland; [12] Division of Pulmonary Medicine, University Hospital of Zürich, Zürich, Switzerland

Correspondence to: Prof. Dr. Nicole Probst-Hensch,

Institute of Social and Preventive Medicine

Chronic Disease Epidemiology

University of Zürich

Sumatrastr. 30, CH-8091 Zürich, Switzerland

Nicole.Probst@ifspm.uzh.ch

Keywords: Epidemiologic Studies, Health Services, Incidence, Prognosis, Chronic Obstructive Pulmonary Disease

Abbreviations: COPD Chronic Obstructive Pulmonary Disease, FEV1 Forced expiratory volume in 1 second, FVC Forced vital capacity, GOLD Global Initiative for Chronic Obstructive Lung Disease

Abstract

Background: Understanding the prognostic meaning of early COPD stages in the general population is relevant for discussions about under-diagnosis. So far, COPD prevalence and incidence were often estimated using pre- instead of post-bronchodilation spirometry. In the SAPALDIA Cohort we investigated time course, clinical relevance and determinants of severity stages of obstruction using pre-bronchodilator spirometry.

Methods: Incident obstruction was defined as FEV1/FVC ratio ≥0.70 at baseline and <0.70 at follow-up, non-persistence inversely. Determinants were assessed in 5490 adults with spirometry and respiratory symptom data in 1991 and 2002 using Poisson regression controlling for self-declared asthma and wheezing. Change in obstruction severity (defined analogously to GOLD classification) over 11 years was related to shortness of breath and health service utilization for respiratory problems by logistic models.

Results: Incidence rate of obstruction was 14.2 cases/1'000 person years. 20.9% of obstructive cases (n=113/540) were non-persistent. Age, smoking, chronic bronchitis and non-current asthma were determinants of incidence. After adjustment for asthma, only progressive stage I or persistent stage II obstruction was associated with shortness of breath (OR 1.71 (0.83-3.54), OR 3.11 (1.50-6.42) respectively) and health service utilization for respiratory problems (OR 2.49 (1.02-6.10), OR 4.17 (1.91-9.13) respectively) at follow-up.

Conclusions: The observed non-persistence of obstruction suggests that pre-bronchodilation spirometry, as used in epidemiological studies, might misclassify COPD. Future epidemiological studies should consider both pre- and post-bronchodilation measurements and take specific clinical factors related to asthma and COPD into consideration for estimation of disease burden and prediction of health outcomes.

Introduction

Chronic obstructive pulmonary disease (COPD) is a major cause of morbidity and premature mortality worldwide.[1] At diagnosis, often more than half of the lung function has been lost and subsequent need for medical care is high.[2] This raises concerns about under-diagnosis, particularly regarding earlier disease stages[3, 4] which are expected to be more amenable to preventive action and improvement of quality of life. Timely diagnosis may also reduce health care costs.[5] For the clinical identification of early stages, the Global Initiative for Chronic Obstructive Lung Disease (GOLD) provided an international standard for diagnosis based on a forced expiratory volume in 1 second (FEV1) over forced vital capacity (FVC) ratio < 0.70 measured by post-bronchodilator spirometry.[4] Severity classification depends on FEV1, expressed as percentage of the predicted value: ≥80% mild GOLD stage I, < 80% moderate stage II, <50% severe stage III and <30% very severe stage IV disease.

Population based epidemiological studies are fundamental to understand the time course and prognostic meaning of COPD GOLD stages in the general population. In recent years, a modified GOLD definition omitting bronchodilation has been widely adopted by these studies.[1] The ease of use and straightforwardness of the FEV1/FVC cutoff facilitates standardization and comparability of observations,[6] and overcomes the shortcomings of previous inconsistent case definitions producing a wide range of prevalence and incidence estimates, and complicating evaluation of health care needs.[7] Although pre-bronchodilation measurements may overestimate COPD prevalence by up to 50%,[8, 9] and might be unreliable when assessing COPD determinants because of reversible airflow obstruction, it is not known whether they perform worse than post-bronchodilator measurements for predicting future health outcomes.[1] So far, GOLD stages II and more have consistently been associated with mortality and reduced quality of life in epidemiological studies using pre-bronchodilation spirometry.[10-13] The picture is less straightforward for stage I, which is most relevant for discussions about underdiagnosis. It has been associated with increased mortality in population studies,[10, 12, 14] but partially respiratory symptoms might be responsible for that.[12, 14] Similarly, in the SAPALDIA cohort we could recently show that stage I predicted rapid decline in FEV1, a cardinal feature of COPD,[4] lower quality of life and increased health care utilization for respiratory problems 11 years later, but only in the presence of respiratory symptoms at baseline.[15]

In this current study based on pre-bronchodilation spirometry data from the SAPALDIA cohort, we investigated the time course and clinical relevance of severity of spirometric obstruction according to modified GOLD criteria while controlling for the effects of overt and undiagnosed asthma.

MATERIALS AND METHODS

Study population

The SAPALDIA cohort[16-18] consists of a random sample of 18-62 year old adults from eight communities. For this study, we included participants with valid spirometry and respiratory symptom data from both, baseline (1991) and follow-up (2002) surveys (Online Figure 1).

The SAPALDIA cohort study complies with the Helsinki Declaration. Written informed consent was obtained from participants at both surveys. The study was approved by the central ethics committee of the Swiss Academy of Medical Sciences and the respective Cantonal Ethics Committees of the eight study regions.

Spirometry

The spirometry protocol was equivalent to the one of the European Community Respiratory Health Survey (ECRHS).[19] No bronchodilation was applied. Identical spirometers (Sensormedics model 2200, Yorba Linda, USA) and protocols were used at both surveys; comparability was assessed before and after each one.[20, 21] Three to eight forced expiratory lung function maneuvers were performed and at least two acceptable measurements of forced vital capacity (FVC) and forced expiratory volume in the first second (FEV_1) were obtained, complying with American Thoracic Society criteria.[22]

Obstruction to airflow and its severity

Spirometric obstruction was defined as FEV1/FVC <0.7 in pre-bronchodilation measurement. An incident case of obstruction was defined as a person with FEV1/FVC ratio ≥ 0.70 at baseline, but <0.70 at follow-up examination. Cases of non-persistence were defined inversely.

In measurements with FEV1/FVC <0.7, severity of obstruction was defined analogously to the GOLD guidelines,[4] applying the prediction equation of Quanjer et al.[23]: FEV1 values of 80% or more of the predicted value were classified stage I, values below this threshold as stage II and more, integrating stages III (FEV1 <50% predicted) and IV (FEV1 <30% predicted) into stage II.

Categories of change in obstruction severity during follow-up

Categories of change in severity of obstruction during follow-up were defined as follows: ‘incident stage I’ (normal FEV1/FVC ratio at baseline and stage I at follow-up, n=683), ‘incident stage II’ (normal FEV1/FVC at baseline and stage II at follow-up, n=85), ‘persistent stage I’ (stage I at baseline and follow-up, n=294), ‘stage I progressing’ (stage I at baseline and stage II at follow-up, n=56), ‘persistent stage II’ (stage II at both examinations, n=61) and ‘non-persistent’ (stage I or more at baseline and normal FEV1/FVC at follow-up, n=113). Cases of stage II at baseline but stage I at follow-up (n=16) were not analyzed.

Chronic bronchitis and shortness of breath

Chronic bronchitis was defined as self-report of cough or phlegm during the day or at night on most days for as much as 3 months each year for ≥ 2 years.

Shortness of breath was defined as affirmative answer to the question "Are you troubled by shortness of breath when hurrying on level ground or walking up a slight hill?".

Asthma status

Presence of asthma at baseline and follow-up, respectively, was defined by the question 'Have you ever had asthma?'. Asthma cases reporting attacks during the 12 months before interview or current use of asthma medication were classified as current asthma, the others as non-current. To identify hidden asthma, we considered subjects reporting wheezing without cold in the 12 months preceding each interview.

Smoking status

Ever smokers reported smoking ≥ 20 packs of cigarettes or ≥360g of tobacco in their lifetime at baseline,[24] former smokers quitting smoking at least 1 month before, and current smokers reported active smoking. Smoking intensity was assessed by pack-years smoked up to baseline and classified *a priori* into ≥ 15 and < 15 pack-years for heavy and light smoking, respectively.

Health service use for respiratory problems

Health service use for respiratory problems was defined as positive answer to one of the following questions: "Have you visited a hospital casualty department or emergency room because of breathing problems in the last 12 months?", "Have you spent a night in hospital because of breathing problems in the last 12 months?", "Have you been seen by a general practitioner because of breathing problems or because of shortness of breath in the last 12 months?", "Have you seen a specialist (chest physician, allergy specialist, internal medicine specialist, ENT doctor) because of your breathing problems or shortness of breath in the last 12 months?".

Health service use for cardiovascular problems

Data from equivalent questions assessing health service use for cardiovascular problems at follow-up was used for sensitivity analysis only.

Statistical analysis

Baseline characteristics were compared between the whole SAPALDIA study population and participants included in the present study, and analogously between COPD transition categories.

Incidence rate of spirometric obstruction was estimated as the number of new cases per total person-years (PY) at risk in thousands. The non-persistence rate was calculated equivalently. Rate ratios for both outcomes were estimated using Poisson regression with the following baseline characteristics: sex, age (in categories of 18-30, >30-40, >40-50 and >50 years), smoking status (never smoker, light or heavy ever-smoker), symptoms of chronic bronchitis at baseline, educational level and study centre. Variables coding for asthma and wheezing at baseline and follow-up were included into the models to

assess their independent impact on the outcomes, and to adjust for overt and hidden asthma. The analysis was repeated using the 5th percentile (lower limit of normal, calculated as 1.645 residual standard deviations or more below predicted according to Quanjer et al.[23]) of the FEV1/FVC ratio distribution to define obstruction. Logistic regression was used to compare presence of shortness of breath and health care services utilization for respiratory problems at follow-up between categories of change in severity of obstruction. Models were adjusted for demographic characteristics (sex, age, education, examination area), baseline health service use for respiratory problems (only in health service utilization models), smoking habits (light/heavy smoker at baseline, pack-years between surveys), preexisting symptoms (chronic bronchitis, shortness of breath), and asthma or wheezing at either examination.

As sensitivity analysis, confounding by cardiovascular co-morbidity was assessed for health care utilization for respiratory problems and respiratory symptoms by including service utilization for cardiovascular problems at follow-up. Furthermore, study participants having only baseline spirometry were compared to the present study sample to predict the probability of participation for each individual. A dichotomous variable coding participation was regressed on baseline covariates used in the regression models. Regression analyses were then repeated using the inversed participation probabilities as weights.

The statistical analysis was performed using SAS Software, Version 9.1 (SAS Institute Inc., Cary, North Carolina, USA) and STATA version 9.2 (StataCorp, College Station, Texas, USA).

RESULTS

Baseline characteristics

Baseline characteristics of SAPALDIA participants and subjects included in the current analysis are presented in online table O1. 53% of the participants were women and the average age at baseline was 41.1 years (range 18-62 years). 30% of the study population was actively smoking at baseline, 52% had ever smoked. Missing at follow-up examination was more frequent in participants with higher obstruction stages (online table O2). As previously described in detail, women, never smokers, well educated subjects, and people with good respiratory health and no atopy were slightly overrepresented among follow-up participants and therefore in this study.[17]

Baseline characteristics according to categories of change in severity of obstruction are presented in table 1. The proportion of females was markedly decreased in all categories but 'persistently normal' and 'incident stage I'. Lung function values presented a pattern expected from the severity definitions, except for categories 'persistent stage I' and 'non-persistent' which had a mean FEV1 close to 100% of the predicted value and the highest FVC values (125.9% and 122.4% predicted respectively). Both categories also had the highest absolute FVC values (4.97L and 4.84L respectively, online Table O3). The proportion of never smokers was lowest in categories 'stage I progressing' (19.6%), 'persistent stage II' (31.1%) and 'incident stage II' (31.8%).

Determinants of incidence and non-persistence of obstruction

To assess determinants of incidence and non-persistence of obstruction, we stratified the study sample by baseline FEV1/FVC ratio (FEV1/FVC<0.70 vs. FEV1/FVC≥0.70)

From the 4945 participants with baseline FEV1/FVC≥0.70, 765 had incident obstruction at follow-up (table 2). This corresponds to a cumulative incidence of 15.5% and an incidence rate of 14.2 cases/1000 person years (PY). Incidence rates were 23.1 and 28.0 cases/1000 PY for participants with non-current and current asthma at baseline respectively, but only 13.4 cases/1000 PY for subjects without. In participants never reporting asthma or wheezing at either examination, the rate was 12.3 cases/1000 PY. Determinants of incidence were (relative rate (RR) and 95%-confidence interval (95%-CI)): older age (RR 1.38 per 10 years, 95%-CI 1.29-1.47), heavy smoking at baseline (RR 1.51, 95%-CI 1.29-1.77), chronic bronchitis at baseline (RR 1.23, 95%-CI 1.00-1.51), non-current asthma at baseline (RR 1.39, 95%-CI 1.01-1.92), current asthma at follow-up (RR 1.68, 95%-CI 1.13-2.50), and wheezing without cold at follow-up (RR 1.95, 95%-CI 1.57-2.42). Among participants with FEV1/FVC<0.70 at baseline (n=540), 113 (20.9%) presented a normal value at follow-up, giving a non-persistence rate of 19.2 cases/1000 PY (online table O4). 93.8% of non-persistent cases classified as stage I obstruction at baseline. Participants with current asthma at follow-up had a significantly lower rate of non-persistence (4.9 cases/1000 PY). In participants never reporting asthma or wheezing at either examination, the rate was 22.8 cases/1000 PY. Heavy smokers at baseline and wheezers at follow-up showed lower rates of non-persistence (14.0 and 9.8 cases/1000 PY respectively), but the effects did not reach statistical significance after adjustment for all asthma variables.

When using the lower limit of normal of the FEV1/FVC ratio to define obstruction, lower incidence (7.2 cases/1000 PY) and higher non-persistence (31.5 cases/1000 PY) rates were observed (online table O5). Additionally, female sex was associated with incidence (RR 1.62, 95%-CI 1.32-1.98). The effects for the other determinants were comparable to the previous analyses (reported in tables 2 and O4).

Categories of change in severity of obstruction and shortness of breath at follow-up

All transition categories except ' non-persistent obstruction' were associated with shortness of breath at follow-up in the crude model (table 3). The association was strongest for categories 'stage I progressing' (odds ratio (OR) 3.76, 95%-CI 2.18-6.48) and 'persistent stage II' (OR 5.43, 95%-CI 3.15-9.37). After adjusting for baseline covariates sex, age, education, smoking, chronic bronchitis, shortness of breath and area, only categories 'stage I progressing' and 'persistent stage II' remained statistically significant (OR 2.21, 95%-CI 1.10-4.45 and OR 4.38, 95%-CI 2.19-8.75 respectively). Adjustment for current or non-current asthma and wheezing without a cold at either examination made the estimate for 'stage I progressing' statistically non-significant (OR 1.71, 95%-CI 0.83-3.54) and decreased effect sizes.

Categories of change in severity of obstruction and health service utilization for respiratory problems at follow-up

The only two transition categories significantly associated with health service use for respiratory problems at follow-up were 'stage I progressing' and 'persistent stage II', irrespective of covariates included in the logistic model (Figure 1; online table O6). After adjustment for sex, age, education, area, baseline health service use for respiratory problems, smoking, baseline respiratory symptoms (chronic bronchitis, shortness of breath) as well as asthma, subjects progressing from stage I to stage II obstruction during follow-up were 2.5 times (OR 2.49, 95%-CI 1.02-6.10) and those persistently in stage II 4.2 times (OR 4.17, 95%-CI 1.91-9.13) more likely to utilize health services for respiratory problems than subjects with normal spirometry. The association with category 'non-persistent obstruction' was marginally significant (OR 2.28, 95%-CI 0.98-5.27, p=0.054) and remained largely unaltered by asthma adjustment.

Sensitivity analysis

Inclusion of health service use for cardiovascular problems at follow-up did not alter the associations of categories of change in obstruction severity with health service use for respiratory problems or respiratory symptoms at follow-up.

Weighted regression analyses yielded the same determinants of incidence and non-persistence, and the same associations between longitudinal obstruction categories and shortness of breath or health service use for respiratory problems at follow-up (data not shown).

Table 1 Baseline characteristics according to change in severity of obstruction[1] during follow-up

	Persistently normal	Incident stage I	Incident stage II	Persistent stage I	Stage I progressing	Persistent stage II	Non-persistent
	n=4181	*n=683*	*n=85*	*n=294*	*n=56*	*n=61*	*n=113*
Female sex (%)	**54.7**	**54.2**	**44.7**	**39.1**	**33.9**	**32.8**	**40.7**
Age *in years* **(**mean/sd)	**39.2** / 11.2	**45.3** / 10.3	**44.5** / 11.9	**48.8** / 9.6	**48.5** / 9.2	**49.5** / 9.1	**47.0** / 9.1
No professional education (%)	**12.1**	**16.4**	**23.5**	**15.6**	**21.4**	**19.7**	**8.8**
FEV1 % of predicted value (mean/sd)	**109.9** / 13.6	**107.4** / 12.5	**91.3** / 12.5	**101.3** / 10.9	**89.1** / 7.0	**67.4** / 10.4	**99.6** / 13.5
FVC % of predicted value mean/sd)	**114.0** / 0.2	**119.2** / 0.1	**100.6** / 0.1	**125.9** / 0.1	**116.0** / 0.1	**96.0** / 0.1	**122.4** / 0.2
FEV1/FVC % of predicted value (mean/sd)	**100.9** / 0.1	**94.6** / 0.1	**94.9** / 0.1	**84.1** / 0.0	**80.2** / 0.1	**73.3** / 0.1	**84.8** / 0.0
Never smoker (%)	**49.9**	**44.7**	**35.3**	**36.7**	**19.6**	**31.1**	**38.9**
Light smoker at baseline (<15 PY)[2] (%)	**28.7**	**19.3**	**13.6**	**20.9**	**9.3**	**11.8**	**24.5**
Heavy smoker at baseline (>=15 PY)[2] (%)	**18.2**	**31.8**	**42.4**	**38.4**	**62.5**	**52.5**	**29.2**
Shortness of breath at baseline (%)	**21.7**	**25.0**	**42.4**	**25.5**	**44.6**	**47.5**	**14.2**

Chronic bronchitis (%) **at baseline**	**7.3**	**11.6**	**20.0**	**13.3**	**19.6**	**27.9**	**9.7**
Wheezing in last 12 months at baseline	**4.8**	**7.8**	**22.4**	**9.9**	**28.6**	**20.0**	**8.0**
Non-current Asthma (%) **at baseline**	**5.6**	**10.5**	**17.6**	**12.2**	**21.4**	**27.9**	**9.7**
Current asthma at (%) **baseline**	**1.8**	**3.4**	**11.8**	**4.4**	**16.1**	**16.4**	**4.4**
Health service use (%) for respiratory problems at baseline	**18.0**	**22.0**	**27.1**	**26.9**	**33.9**	**42.6**	**23.0**

[1] Obstruction was defined as FEV1/FVC<0.70 based on pre-bronchodilation spirometry.

[2] Numbers do not add up to 100.0% due to smokers with missing pack-year information

Table 2 Incidence rate of obstruction (FEV1/FVC<0.7) using pre-bronchodilator spirometry during 11 years of follow-up according to a set of baseline characteristics

Predictor at baseline	Person-years at risk (in 1000)	Number of cases	Incidence rate (cases per 1000 person years)	(95% CI)	Crude incidence rate ratio	(95% CI)	Adjusted[3] incidence rate ratio	(95% CI)
All (N=4945) [1]	**54.00**	**765**	**14.17**	**(13.20 - 15.21)**				
Gender:								
· **men**	24.55	357	**14.54**	(13.11 - 16.13)	**1.00**	(Ref)	**1.00**	(Ref)
· **women**	29.45	408	**13.85**	(12.57 - 15.27)	**0.95**	(0.84 - 1.09)	**1.03**	(0.90 - 1.18)
Age (years) at baseline								
· **18-30**	11.62	70	**6.02**	(4.77 - 7.61)	**1.00**	(Ref)	**1.00**	(Ref)
· **>30-40**	14.70	165	**11.22**	(9.64 - 13.07)	**1.86**	(1.43 - 2.43)	**1.72**	(1.33 - 2.24)
· **>40-50**	15.81	253	**16.00**	(14.15 - 18.10)	**2.66**	(2.08 - 3.40)	**2.38**	(1.85 - 3.06)
· **>50**	11.86	277	**23.35**	(20.75 - 26.27)	**3.88**	(3.05 - 4.93)	**3.77**	(2.94 - 4.83)
Smoking status at baseline[2]:								
· **never smoker**	28.02	362	**12.92**	(11.65 - 14.32)	**1.00**	(Ref)	**1.00**	(Ref)
· **ever smoker:**								
<15 packyrs.	14.94	153	**10.24**	(8.74 - 12.00)	**0.79**	(0.66 - 0.95)	**0.87**	(0.73 - 1.04)
≥15 packyrs	11.03	250	**22.66**	(20.02 - 25.65)	**1.75**	(1.52 - 2.03)	**1.51**	(1.29 - 1.77)

Chronic bronchitis at baseline:								
· **absent**	49.62	671	**13.52**	(12.54 - 14.59)	**1.00**	(Ref)	**1.00**	(Ref)
· **present**	4.38	94	**21.45**	(17.53 - 26.26)	**1.59**	(1.30 - 1.93)	**1.23**	(1.00 - 1.51)
Asthma at baseline								
· **absent**	50.47	678	**13.43**	(12.46 - 14.48)	**1.00**	(Ref)	**1.00**	(Ref)
· **present, but non-current**	2.34	54	**23.12**	(17.70 - 30.18)	**1.72**	(1.33 - 2.22)	**1.39**	(1.01 - 1.92)
· **present, current**	1.18	33	**28.02**	(19.92 - 39.41)	**2.09**	(1.51 - 2.88)	**0.79**	(0.51 - 1.23)
Asthma at Follow-up								
· **absent**	49.75	662	**13.31**	(12.33 - 14.36)	**1.00**	(Ref)	**1.00**	(Ref)
· **present, but non-current**	2.64	49	**18.54**	(14.01 - 24.53)	**1.39**	(1.07 - 1.82)	**1.19**	(0.85 - 1.65)
· **present, current**	1.60	54	**33.77**	(25.86 - 44.09)	**2.54**	(1.97 - 3.28)	**1.68**	(1.13 - 2.50)
Wheezing without a cold at baseline								
· **absent**	50.81	689	**13.56**	(12.59 - 14.61)	**1.00**	(Ref)	**1.00**	(Ref)
· **present**	2.96	71	**23.99**	(19.01 - 30.27)	**1.77**	(1.41 - 2.21)	**1.04**	(0.81 - 1.35)
Wheezing without a cold at follow-up								
· **absent**	50.34	655	**13.01**	(12.05 - 14.05)	**1.00**	(Ref)	**1.00**	(Ref)
· **present**	3.65	110	**30.10**	(24.97 - 36.29)	**2.31**	(1.92 - 2.79)	**1.95**	(1.57 - 2.42)

[1] additional reduction of sample size due to exclusion of participants with >120 pack years at baseline or > 150 at follow-up

[2] smoking status at baseline: never smokers: <20 packs of cigarettes and <360 g of tobacco in lifetime

[3] adjusted for study area, educational level, and all predictors listed in the table

Table 3 Association[1] of categories of change in severity of obstruction[2] with shortness of breath while walking at follow-up

	Crude Model		Adjusting for all but asthma covariates[3]		Adjusting for asthma & wheezing at base-line or follow-up	
Variable	**Rel. rate** 95%-CI	**p-value**	**Rel.rate** 95%-CI	**p-value**	**Rel.rate** 95%-CI	**p-value**
Incident stage I	**1.59**	**0.000**	**1.24**	0.056	**1.12**	0.344
(n= 683)	(1.32 - 1.91)		(0.99 - 1.56)		(0.89 - 1.41)	
Incident stage II	**2.74**	**0.000**	**1.43**	0.192	**1.10**	0.743
(n= 85)	(1.74 - 4.30)		(0.84 - 2.45)		(0.63 - 1.90)	
Persistent stage I	**1.48**	**0.004**	**1.14**	0.428	**1.02**	0.903
(n=294)	(1.13 - 1.94)		(0.82 - 1.60)		(0.73 - 1.44)	
Stage I progressing	**3.76**	**0.000**	**2.21**	**0.026**	**1.71**	0.148
(n= 56)	(2.18 - 6.48)		(1.10 - 4.45)		(0.83 - 3.54)	
Persistent stage II	**5.43**	**0.000**	**4.38**	**0.000**	**3.11**	**0.002**
(n= 61)	(3.15 - 9.37)		(2.19 - 8.75)		(1.50 - 6.42)	
Non-persistent	**1.02**	0.947	**1.40**	0.237	**1.39**	0.246
(n= 113)	(0.64 - 1.62)		(0.80 - 2.44)		(0.80 - 2.43)	
Asthma at baseline						
non-current[4]					**1.08**	0.739
					(0.70 - 1.65)	
current[4]					**0.50**	**0.024**
					(0.27 - 0.91)	
Asthma at follow-up						
non-current[4]					**1.09**	0.667
					(0.73 - 1.63)	
current[4]					**2.18**	**0.004**
					(1.28 - 3.72)	
Wheezing without a					**1.41**	**0.034**
cold at baseline					(1.03 - 1.94)	

Wheezing without a	**2.07**	**0.000**
cold at follow-up	(1.55-2.75)	

95%-CI: 95%-confidence intervals; PY: pack-years

[1] Reference category: persistently without obstruction to the airflow

[2] Obstruction was defined as FEV1/FVC<0.70 based on pre-bronchodilation spirometry.

[3] Covariates adjusted for were: sex, age, smoking (light or heavy ever smoker), chronic bronchitis, shortness of breath while walking at baseline, education and area.

[4] Current asthma was defined as presence of asthma attacks in the 12 months prior to assessment or current asthma medication.

Non-current asthma cases were defined as self-declared asthma without attacks or asthma medication.

DISCUSSION

In our general population sample, we observed an incidence of modified GOLD COPD (obstruction based on pre-bronchodilation spirometry) of 14.2 cases per 1000 person years (PY). This estimate is at the higher end of comparable ones[25-28] which range between 3 and 16 cases/1000 PY depending on age distribution, smoking prevalence, follow-up time, and inclusion of asthmatics. This high incidence could only partly be explained by these factors. We replicated associations with age and smoking from previous studies[1, 6, 25-29], and found a significant association with chronic bronchitis, a finding not reported consistently so far.[27-29] Female sex was significantly associated with incidence only when the FEV1/FVC ratio lower limit of normal was used to define disease. Previous evidence regarding gender differences in obstruction rates is inconsistent,[25, 27-29] but our finding could support the currently debated hypothesis that women are more susceptible to COPD.[1, 30]

Our observation that 20.9 % of obstructive cases at baseline did not persist is noteworthy. Two factors likely explain non-persistence. The first is measurement error: Like the ECRHS study,[28] we observed that FEV1/FVC values close to the 0.70 cut-off are predictive of both, incidence and non-persistence (data not shown) and 93.8% of our non-persistent cases were mildly obstructive. Second, the use of pre-bronchodilator measurements prevents the identification of reversible obstruction (mostly undiagnosed asthma). The high FVC and normal FEV1 percent predicted values in our non-persistent cases support this possibility. Also, category 'non-persistent obstruction' was marginally associated with health service use for respiratory problems irrespective of asthma adjustment. We captured reversible obstruction as far as possible by considering wheezing without a cold (besides self-declared asthma), but hidden non-wheezing asthma cases might still be present.

Pre-bronchodilatator measurements in epidemiological studies might thus misclassify COPD, especially in mild GOLD I stages, but our results suggest their longitudinal course may predict future health events on a population level independently of pre-existing symptoms, smoking or health care use. While shortness of breath and respiratory care utilization was particularly high in participants progressing from stage I to stage II obstruction or persisting in stage II, those remaining in stage I did not have increased risks for either outcome at follow-up.

There is thus a need to better characterize the modified GOLD stage I category in epidemiological studies. In the past, epidemiological studies have omitted post-bronchodilation spirometry due to time and resource constraints, or in favour of broncho-challenge testing. The procedure is however essential to differentiate asthma from COPD in clinical practice. Future epidemiological studies will thus additionally need longitudinal post-bronchodilation measurements and consider characteristics such as medication intake and symptoms for asthma[32] or the BODE index for COPD[33, 34], which are important prognostic factors on the individual level, to define groups at high risk for adverse health outcomes or increased use of health services. Such extended assessments are foreseen in the third examination of SAPALDIA.

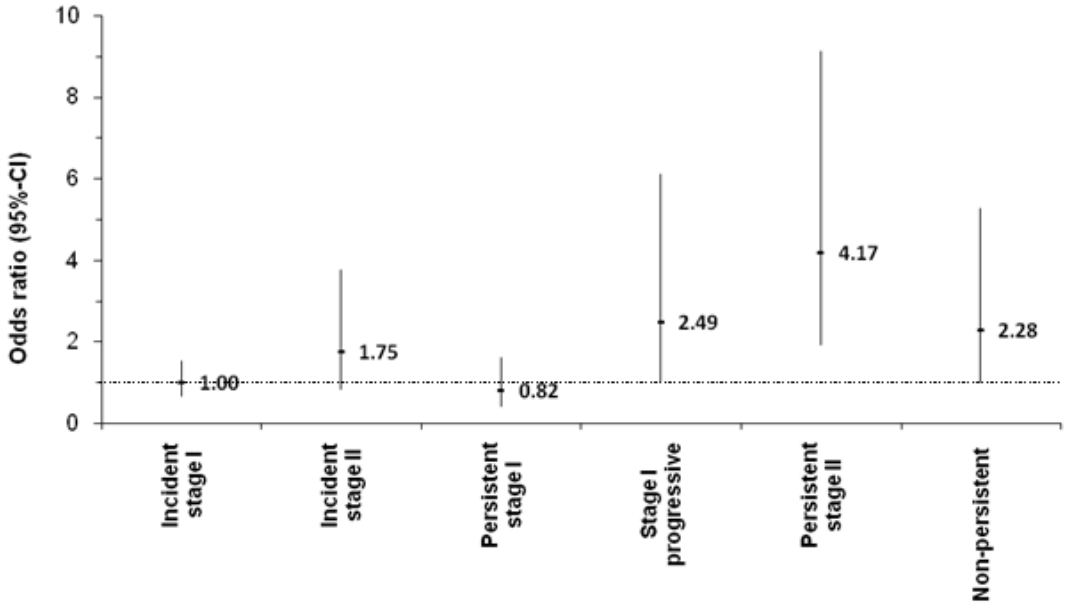
Our study benefited from stringent quality control in spirometry and detailed information on lifestyle-factors. As discussed above, a limitation is the use of pre-bronchodilator measurements. The associations of change in severity of obstruction with health service use for respiratory problems or shortness of breath were robust to cardiovascular co-morbidity. Finally, according to weighted regression analyses loss-to-follow-up was not a source of bias, although selection for lower stages of obstruction was detectable in our sample.

Conclusion

The observed non-persistence of obstruction suggests that pre-bronchodilation spirometry at only two time points in epidemiological studies might misclassify COPD. Still, our findings regarding shortness of breath and health service use for respiratory problems show that pre-bronchodilation spirometry, particularly its longitudinal course, has value in predicting health outcomes on a population level. To accurately identify risk groups, future epidemiological studies will have to consider both, pre- and post-bronchodilation spirometry as well as individual prognostic factors used in today's clinical practice.

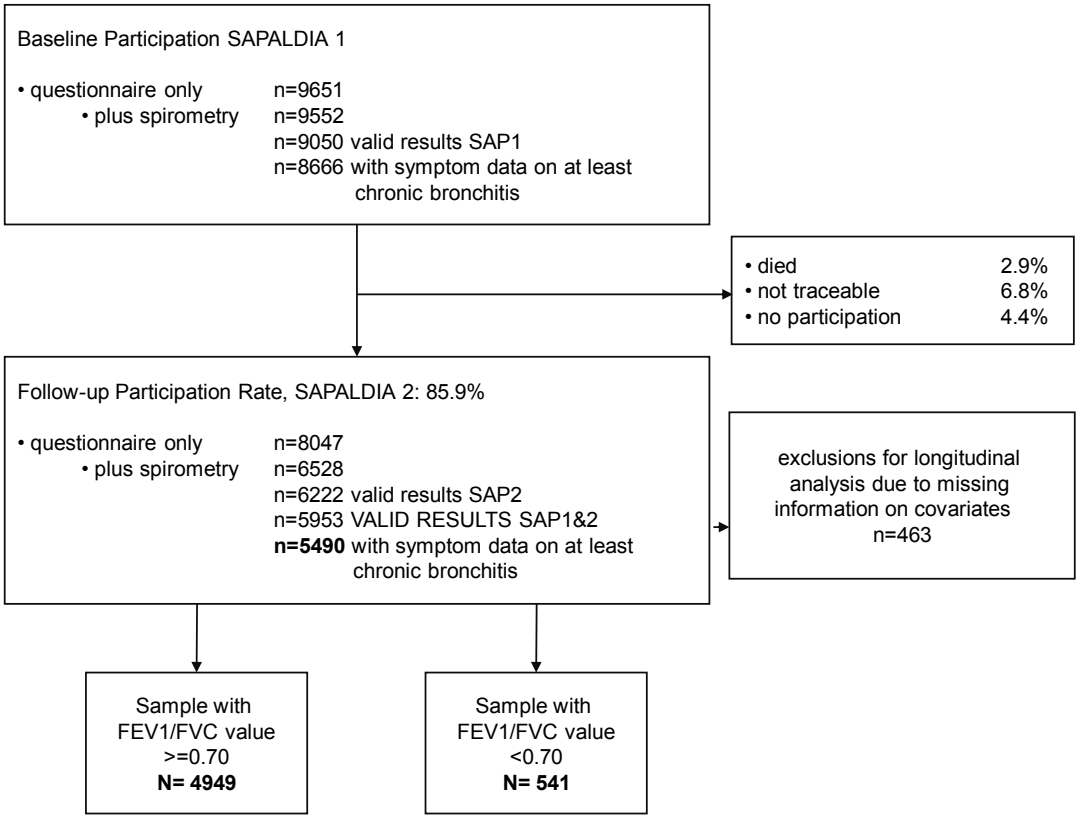
Figures

Figure 1 Association of categories of change in severity of obstruction with health service use for respiratory problems at follow-up[1]



[1] Obstruction was defined as FEV1/FVC<0.70 based on pre-bronchodilation spirometry.
Effect estimates are adjusted for: sex, age, smoking (light or heavy ever smoker), chronic bronchitis and shortness of breath at baseline, health service use for respiratory problems at baseline, asthma and wheezing at baseline or follow-up examination, education and area

Online figure 1 Participation in SAPALDIA1 (baseline) and SAPALDIA 2 (follow-up)



Acknowledgements

† **SAPALDIA TEAM**

Study directorate: T Rochat (p), U Ackermann-Liebrich (e), JM Gaspoz (c), N Künzli (e/exp), LJS Liu (exp), NM Probst Hensch (e/g), C Schindler (s).

Scientific team: JC Barthélémy (c), W Berger (g), R Bettschart (p), A Bircher (a), G Bolognini (p), O Brändli (p), M Brutsche (p), L Burdet (p), M Frey (p), MW Gerbase (p), D Gold (e/c/p), W Karrer (p), R Keller (p), B Knöpfli (p), U Neu (exp), L Nicod (p), M Pons (p), E Russi (p), P Schmid-Grendelmeyer (a), J Schwartz (e), P Straehl (exp), JM Tschopp (p), A von Eckardstein (cc), JP Zellweger (p), E Zemp Stutz (e).

Scientific team at coordinating centers: PO Bridevaux (p), I Curjuric (e), J Dratva (e), D Felber Dietrich (c), D Keidel (s), M Imboden (g), H Phuleria (exp), E Schaffner (s), GA Thun (g), A Ineichen (exp), M Ritter (exp).

(a) allergology, (c) cardiology, (cc) clinical chemistry, (e) epidemiology, (exp) exposure, (g) genetic and molecular biology, (m) meteorology, (p) pneumology, (s) statistics

The study could not have been done without the help of the study participants, technical and administrative support and the medical teams and field workers at the local study sites.

Local fieldworkers : Aarau: M Broglie, M Bünter, D Gashi, Basel: R Armbruster, T Damm, U Egermann, M Gut, L Maier, A Vögelin, L Walter, Davos: D Jud, N Lutz, Geneva: M Ares, M Bennour, B Galobardes, E Namer, Lugano: B Baumberger, S Boccia Soldati, E Gehrig-Van Essen, S Ronchetto, Montana: C Bonvin, C Burrus, Payerne: S Blanc, AV Ebinger, ML Fragnière, J Jordan, Wald: R Gimmi, N Kourkoulos, U Schafroth.

Administrative staff: N Bauer, D Baehler, C Gabriel, R Nilly, F Meier.

Competing interest: None

Ethical approval: The study was approved by the central ethics committee of the Swiss Academy of Medical Sciences and the respective Cantonal Ethics Committees of the eight study regions included in the study.

Funding

The Swiss National Science Foundation (grants no 33CSCO-108796, 3247BO-104283, 3247BO-104288, 3247BO-104284, 3247-065896, 3100-059302, 3200-052720, 3200-042532, 4026-028099), the Federal Office for Forest, Environment and Landscape, the Federal Office of Public Health, the Federal Office of Roads and Transport, the canton's government of Aargau, Basel-Stadt, Basel-Land, Geneva, Luzern, Ticino, Zurich, the Swiss Lung League, the canton's Lung League of Basel Stadt/ Basel Landschaft, Geneva, Ticino and Zurich.

References

1 Mannino DM and Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. *Lancet* 2007;370:765-73.

2 Engstrom CP, Persson LO, Larsson S, et al. Health-related quality of life in COPD: why both disease-specific and generic measures should be used. *Eur Respir J* 2001;18:69-76.

3 Lindberg A, Jonsson AC, Ronmark E, et al. Prevalence of chronic obstructive pulmonary disease according to BTS, ERS, GOLD and ATS criteria in relation to doctor's diagnosis, symptoms, age, gender, and smoking habits. *Respiration* 2005;72:471-9.

4 Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease (updated 2008). In GOLDReport08_Clean.pdf on www.goldcopd.com, accessed December 5th, 2008.

5 Akazawa M, Halpern R, Riedel AA, et al. Economic burden prior to COPD diagnosis: a matched case-control study in the United States. *Respir Med* 2008;102:1744-52.

6 Celli BR, Halbert RJ, Isonaka S, et al. Population impact of different definitions of airway obstruction. *Eur Respir J* 2003;22:268-73.

7 Halbert RJ, Natoli JL, Gano A, et al. Global burden of COPD: systematic review and meta-analysis. *Eur Respir J* 2006;28:523-32.

8 Johannessen A, Omenaas ER, Bakke PS, et al. Implications of reversibility testing on prevalence and risk factors for chronic obstructive pulmonary disease: a community study. *Thorax* 2005;60:842-7.

9 Kim SJ, Suk MH, Choi HM, et al. The local prevalence of COPD by post-bronchodilator GOLD criteria in Korea. *Int J Tuberc Lung Dis* 2006;10:1393-8.

10 Mannino DM, Buist AS, Petty TL, et al. Lung function and mortality in the United States: data from the First National Health and Nutrition Examination Survey follow up study. *Thorax* 2003;58:388-93.

11 Stavem K, Sandvik L and Erikssen J. Can global initiative for Chronic Obstructive Lung Disease stage 0 provide prognostic information on long-term mortality in men? *Chest* 2006;130:318-25.

12 Ekberg-Aronsson M, Pehrsson K, Nilsson JA, et al. Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. *Respir Res* 2005;6:98.

13 Carrasco Garrido P, de Miguel Diez J, Rejas Gutierrez J, et al. Negative impact of chronic obstructive pulmonary disease on the health-related quality of life of patients. Results of the EPIDEPOC study. *Health Qual Life Outcomes* 2006;4:31.

14 Mannino DM, Doherty DE and Sonia Buist A. Global Initiative on Obstructive Lung Disease (GOLD) classification of lung disease and mortality: findings from the Atherosclerosis Risk in Communities (ARIC) study. *Respir Med* 2006;100:115-22.

15 Bridevaux PO, Gerbase MW, Probst-Hensch NM, et al. Long-term decline in lung function, utilisation of care and quality of life in modified GOLD stage 1 COPD. *Thorax* 2008;63:768-74.

16 Martin BW, Ackermann-Liebrich U, Leuenberger P, et al. SAPALDIA: methods and participation in the cross-sectional part of the Swiss Study on Air Pollution and Lung Diseases in Adults. *Soz Praventivmed* 1997;42:67-84.

17 Ackermann-Liebrich U, Kuna-Dibbert B, Probst-Hensch NM, et al. Follow-up of the Swiss Cohort Study on Air Pollution and Lung Diseases in Adults (SAPALDIA 2) 1991-2003: methods and characterization of participants. *Soz Praventivmed* 2005;50:245-63.

18 Downs SH, Schindler C, Liu LJ, et al. Reduced exposure to PM10 and attenuated age-related decline in lung function. *N Engl J Med* 2007;357:2338-47.

19 Burney PG, Luczynska C, Chinn S, et al. The European Community Respiratory Health Survey. *Eur Respir J* 1994;7:954-60.

20 Kunzli N, Ackermann-Liebrich U, Keller R, et al. Variability of FVC and FEV1 due to technician, team, device and subject in an eight centre study: three quality control studies in SAPALDIA. Swiss Study on Air Pollution and Lung Disease in Adults. *Eur Respir J* 1995;8:371-6.
21 Kunzli N, Kuna-Dibbert B, Keidel D, et al. Longitudinal validity of spirometers--a challenge in longitudinal studies. *Swiss Med Wkly* 2005;135:503-8.
22 Standardization of Spirometry, 1994 Update. American Thoracic Society. *Am J Respir Crit Care Med* 1995;152:1107-36.
23 Quanjer PH, Tammeling GJ, Cotes JE, et al. 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. *Eur Respir J Suppl* 1993;16:5-40.
24 Epidemiology standardization project ec, American Thoracic Society. Recommended respiratory disease questionnaires for use with adults and children in epidemiological research. *Am Rev Respir Dis* 1978;118:7-52.
25 Johannessen A, Omenaas E, Bakke P, et al. Incidence of GOLD-defined chronic obstructive pulmonary disease in a general adult population. *Int J Tuberc Lung Dis* 2005;9:926-32.
26 Lokke A, Lange P, Scharling H, et al. Developing COPD: a 25 year follow up study of the general population. *Thorax* 2006;61:935-9.
27 Lindberg A, Eriksson B, Larsson LG, et al. Seven-year cumulative incidence of COPD in an age-stratified general population sample. *Chest* 2006;129:879-85.
28 de Marco R, Accordini S, Cerveri I, et al. Incidence of chronic obstructive pulmonary disease in a cohort of young adults according to the presence of chronic cough and phlegm. *Am J Respir Crit Care Med* 2007;175:32-9.
29 Vestbo J and Lange P. Can GOLD Stage 0 provide information of prognostic value in chronic obstructive pulmonary disease? *Am J Respir Crit Care Med* 2002;166:329-32.
30 Han MK, Postma D, Mannino DM, et al. Gender and chronic obstructive pulmonary disease: why it matters. *Am J Respir Crit Care Med* 2007;176:1179-84.
31 Schermer T, Heijdra Y, Zadel S, et al. Flow and volume responses after routine salbutamol reversibility testing in mild to very severe COPD. *Respir Med* 2007;101:1355-62.
32 de Marco R, Marcon A, Jarvis D, et al. Prognostic factors of asthma severity: a 9-year international prospective cohort study. *J Allergy Clin Immunol.* 2006;117:1249-56.
33 Celli BR, Cote CG, Marin JM, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med.* 2004;350:1005-12.
34 Ong KC, Earnest A and Lu SJ. A multidimensional grading system (BODE index) as predictor of hospitalization for COPD. *Chest.* 2005;128:3810-6.

5. Paper 2:
Assessing the contribution of ambient air pollution to the COPD burden after accounting for known determinants. (aim II)

5.1. HMOX1 and GST variants modify attenuation of FEF25-75% decline due to PM10 reduction.

(published in: **Curjuric I**, Imboden M, Schindler C et al. HMOX1 and GST variants modify attenuation of FEF25-75% decline due to PM10 reduction. Eur Respir J. 2010; 35(3):505-14.
Complying with European Respiratory Society copyright policy the article text is not included into the online thesis version, but is available under: http://erj.ersjournals.com/content/35/3/505.full.pdf+html?sid=48efc88f-50aa-4dcb-a4c5-f7dd242a93a6)

6. Paper 3:
Comparison of the impacts of ambient air pollution and tobacco smoke exposure on COPD on the population level (aim III)

6.1. Different genes interact with particulate matter and tobacco smoke exposure in affecting lung function decline in the general population.

(published on July 6th 2012 in: **Curjuric I**, Imboden M, Nadif R, et al. Different Genes Interact with Particulate Matter and Tobacco Smoke Exposure in Affecting Lung Function Decline in the General Population. PLoS ONE 2012; 7(7): e40175. doi: 10.1371/journal.pone. 0040175

Article link: http://dx.plos.org/10.1371/journal.pone.0040175)

Different Genes Interact with Particulate Matter and Tobacco Smoke Exposure in Affecting Lung Function Decline in the General Population

Ivan Curjuric[1,2], Medea Imboden[1,2], Rachel Nadif[3,4], Ashish Kumar[1,2,5], Christian Schindler[1,2], Margot Haun[6], Florian Kronenberg[6], Nino Künzli[1,2], Harish Phuleria[1,2], Dirkje S Postma[7], Erich W Russi[8], Thierry Rochat[9], Florence Demenais[10,11,12], Nicole M Probst-Hensch[1,2]

[1] Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute SwissTPH,
Basel, Switzerland

[2] University of Basel, Basel, Switzerland

[3] INSERM, U1018, CESP Centre for research in Epidemiology and Population Health, Respiratory and Environmental Epidemiology Team, Villejuif, France

[4] Université Paris-Sud 11, UMRS 1018, Villejuif, France

[5] Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

[6] Division of Genetic Epidemiology, Department of Medical Genetics, Molecular and Clinical Pharmacology, Innsbruck Medical University, Innsbruck, Austria

[7] Department of Pulmonary Medicine and Tuberculosis, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[8] Division of Pulmonary Medicine, University Hospital Zürich, Zürich, Switzerland

[9] Division of Pulmonary Medicine, Geneva University Hospitals, Geneva, Switzerland

[10] INSERM, U946, Genetic Variation and Human Diseases Unit, Paris, France

[11] Fondation Jean Dausset - Centre d'Etude du Polymorphisme Humain (CEPH), Paris, France

[12] Université Paris Diderot, Sorbonne Paris Cité, Institut Universitaire d'Hématologie, Paris, France

Corresponding author:

Prof. Dr. Nicole Probst-Hensch, Dr.phi.II, PhD, MPH

Head Unit Chronic Disease Epidemiology

Swiss Tropical and Public Health Institute

Socinstrasse 57, P.O. Box, 4002 Basel

Switzerland

TEL 0041-61-284-83-78 (direct)

FAX 0041-61-284-81-05

EMAIL: Nicole.Probst@unibas.ch

Keywords:

General population, lung function decline, particulate matter, smoking, oxidative stress, genes, gene-environment interaction, pathway analysis

Abstract

Background:

Oxidative stress related genes modify the effects of ambient air pollution or tobacco smoking on lung function decline. The impact of interactions might be substantial, but previous studies mostly focused on main effects of single genes.

Objectives:

We studied the interaction of both exposures with a broad set of oxidative-stress related candidate genes and pathways on lung function decline, and contrasted interactions between exposures.

Methods:

For 12679 single nucleotide polymorphisms (SNPs), change in forced expiratory volume in one second (FEV_1), FEV_1 over forced vital capacity (FEV_1/FVC), and mean forced expiratory flow between 25 and 75% of the FVC (FEF_{25-75}) was regressed on interval exposure to particulate matter <10µm in diameter (PM10) or packyears smoked (a), additive SNP effects (b), and interaction terms between a) and b) in 669 adults with GWAS data. Interaction p-values for 152 genes and 14 pathways were calculated by the adaptive rank truncation product (ARTP) method, and compared between exposures. Interaction effect sizes were contrasted for the strongest SNPs of nominally significant genes ($p_{interaction}<0.05$). Replication was attempted for SNPs with MAF>10% in 3320 SAPALDIA participants without GWAS.

Results:

On the SNP-level, rs2035268 in gene *SNCA* accelerated FEV_1/FVC decline by 3.8% ($p_{interaction}=2.5x10^{-6}$), and rs12190800 in *PARK2* attenuated FEV1 decline by 95.1ml $p_{interaction}=9.7x10^{-8}$) over 11 years, while interacting with PM10. Genes and pathways nominally interacting with PM10 and packyears exposure differed substantially. Gene *CRISP2* presented a significant interaction with PM10 ($p_{interaction}= 3.0x10^{-4}$) on FEV_1/FVC decline. Pathway interactions were weak. Replications for the strongest SNPs in *PARK2* and *CRISP2* were not successful.

Conclusions:

Consistent with a stratified response to increasing oxidative stress, different genes and pathways potentially mediate PM10 and tobacco smoke effects on lung function decline. Ignoring environmental exposures would miss these patterns, but achieving sufficient sample size and comparability across study samples is challenging.

Introduction

Lung function is an important determinant of respiratory health and life expectancy [1,2,3,4]. Its longitudinal course is affected by different environmental exposures such as active tobacco smoking, environmental tobacco smoke exposure [5], possibly workplace exposures to dusts and fumes [6,7,8,9] as well as ambient air pollution [10]. Both air pollution and tobacco smoke are known to contain free radicals and to induce their direct formation at the tissue level causing damage of cell walls, proteins and DNA, and chronic tissue inflammation and remodeling in the long run [11,12]. Upon exposure, different protein systems including those scavenging reactive oxygen species (ROS) are up-regulated, and the level of response is influenced by variation in underlying genes. Likewise, polymorphisms in oxidative stress related candidate genes like *gluthathione s-transferases (GSTs)*, *microsomal epoxide hydroxylase (EPHX)*, or *heme-oxygenase 1 (HMOX-1),* have been associated with lung function decline and chronic obstructive pulmonary disease (COPD), a disease characterized by accelerated, progressive lung function loss [13,14,15,16]. But most of these candidate genes have not been consistently replicated across studies and populations according to a recent review [15]. Similarly, genome-wide association studies (GWAS) of lung function partially struggled with replication [17,18]. Further, in GWAS on lung function level or COPD prevalence [17,18,19,20,21,22] association signals in known oxidative-stress genes were not strong [23].

Reasons for non-replication could be genetic heterogeneity across populations, or also sub-phenotypes of disease [24]. However, it is also possible that differences in environmental factors, and hence presence of gene-environment interaction play a role. To the best of our knowledge, only one published genome-wide interaction study examining the effect of farming exposure on childhood asthma has taken into account gene-environment interaction in respiratory disease to date [25]. This gap in the scientific literature is probably due to increased sample size requirements when assessing gene-environment interactions with classical analysis methods [26,27]. However, their importance in respiratory disease has previously been shown in candidate gene studies focusing on single genes and SNPs therein [28,29,30], as well as follow-up studies of GWAS [31,32].

Analysis methods such as pathway- or gene-set analyses [33] can at least partly overcome sample size restrictions by reducing the dimensionality of the data, and thus offer a promising alternative study approach. Based on biological knowledge of genes and their organization into molecular pathways, the longitudinal course of lung function might be better explained by accumulating interaction signals between environmental exposures and multiple SNPs of the same gene, or different genes involved in the same canonical pathway contributing to a functional entity in the organism.

We thus aimed to investigate to which extent oxidative-stress related genes and pathways interact significantly with interval exposure to ambient particulate matter of mean diameter < 10 µm (PM10) or active tobacco smoking on natural lung function decline using genome-wide data from non-asthmatic adults of the Swiss Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA). SNP-level interaction signals were integrated onto upper biological levels to identify significantly interacting genes and pathways. The impact of PM10 exposure on lung function decline was contrasted to tobacco smoking by comparing patterns of associations at the

gene- and pathway level, as well as interaction effect sizes for the strongest interacting SNP within genes.

Methods

Ethics statement

All participants gave written informed consent. The study was approved by the Overall Regional Ethics Commission for Clinical Medicine (Swiss Academy of Medical Sciences, Basel, Switzerland) and the responsible cantonal ethics committees of each study centre (Ethics Commissions of the cantons Aarau, Basel, Geneva, Grisons, Ticino, Valais, Vaud, and Zürich).

Study population

SAPALDIA is a population-based cohort study established in 1991 to assess the effects of long-term exposure to ambient air pollution on respiratory health, with a first follow-up examination in 2002. Participants were residents from 8 communities throughout Switzerland aged 18-60 years at baseline. Details of the study design and methodology were published elsewhere [10,34,35].

The current work is based on up to 669 non-asthmatic participants with genome-wide data fulfilling quality control criteria and complete data on sex, age, height, PM10- and smoking exposure (**see supporting files Figure S1**). Participants without genome-wide data served as replication sample.

Spirometric measurements

Spirometry was performed without bronchodilation. Identical spirometry protocols and devices (Sensormedics model 2200, Yorba Linda, USA) were used in 1991 and 2002 [36,37]. Participants were in an upright sitting position and performed three to eight forced expiratory lung function maneuvers according to American Thoracic Society quality criteria [38]. At least two acceptable measurements of forced vital capacity (FVC) and forced expiratory volume in the first second (FEV_1) were obtained. Forced expiratory flow between 25 and 75% of the FVC (FEF_{25-75}) was recorded.

In the present study we studied the decline of FEV_1, the ratio FEV_1/FVC and FEF_{25-75} between 1991 and 2002, as measures of airway obstruction, calculated by subtracting the first measurement from the second (measurement at SAPALDIA2 – measurement at SAPALDIA1).

Health Questionnaire data

Smoking information was assessed by questionnaire. At each examination, never smokers were defined as having smoked less than 20 packs of cigarettes or 360g of tobacco in their life, ex-smokers as having quit smoking at least 30 days before the interview, and current smokers as those who reported active smoking [39]. Packyears smoked between baseline and follow-up examination were used for comparison with interval PM10 exposure, and were calculated by dividing the number of cigarettes per day by 20 (giving number of cigarette packs) and multiplying the result with years of exposure.

Air pollution exposure

Similarly to calculating packyears, interval PM10 exposure was defined by summing individual average home outdoor exposure to PM10 over each year of follow-up, giving estimates in ($\mu g/m^3$) * years. Annual average exposures were calculated by using exposure estimates from Gaussian Dispersion models on a 200m x 200m grid throughout Switzerland for years 1990 and 2000, and interpolating historical trends from fixed air pollution monitoring stations. Participants were assigned individual annual exposure estimates via their geo-referenced residence addresses, taking account of residence changes during follow-up. Details on exposure modeling are given elsewhere [40].

SNP genotyping and imputation

Blood for DNA-analysis was drawn in 2002 in participants giving consent to genetic analyses [34].

Genome-wide genotyping was done on the Illumina Human 610quad BeadChip in the framework of the EU-funded GABRIEL study [41], a large consortium aiming to uncover genetic and environmental causes of asthma. The current work focused on the non-asthmatic portion of participants.

567'589 successfully genotyped autosomal SNPs were imputed to 2.5 Mio using MACH v 1.0 software [42] and the HapMap v22 CEPH reference panel of Utah residents with ancestry from northern and western Europe [43].

Strict quality control (QC) was applied by excluding samples with <97% genotyping success rate, non-European origin, cryptic relatedness or sex-inconsistencies, as well as SNPs with Hardy-Weinberg equilibrium p-value<10^{-4}, call rate <97%, minor allele frequency (MAF) <5% or low imputation quality (Rsq<0.5). A total of 2'168'681 SNPs withstood QC, and genome-wide data was finally available in 669 non-asthmatic individuals with environmental exposure data.

Replication genotyping was attempted for two interacting SNPs (rs360563 in gene *CRISP2*, and rs12190800 in *PARK2*) with MAF>10%. Genotyping was done using the iPLEX Gold MassARRAY (SEQUENOM, San Diego, USA) on the whole SAPALDIA study population including the analysis sample, as the costs for manual sample selection outweigh those of additional genotyping. The replication sample consisted of 3320 successfully genotyped participants with complete data for covariates and all three lung function parameters (**see supporting files Figure S1**).

Definition of oxidative-stress genes and pathways

Oxidative stress related genes were defined as either coding proteins that directly scavenge or endogenously produce ROS, their immediate regulators, or key genes in cascades triggered by oxidative stress. They were identified by searching the Gene Ontology database [44] with the term "response to oxidative stress" and GeneCards with "oxidative stress" in the pathway field of the advanced search option (http://www.genecards.org/index.php?path=/Search/Advanced/, accessed November 2010). Resulting gene lists were further enriched by literature reviews [45,46,47,48,49]. By feeding the gene lists into Ingenuity Pathway Analysis (Ingenuity® Systems,

www.ingenuity.com), 14 molecular pathways related to oxidative stress and environmental exposures of interest were identified (**Table 1**).

Gene regions were defined by retrieving transcription start and end positions in the 'gene track' of the UCSC browser (http://genome.ucsc.edu/) [50], genome build 18 (March 2006), and adding 20 kilo-bases to each end. Referring to dbSNP version 126, available SNP data was matched to gene regions. Data was available for 152 autosomal genes (**Table 1**), of which 46 mapped once to a pathway, 33 twice, and 37 three times or more. Thirty-six genes did not map to one of the 14 pathways, but were related to oxidative stress based on their function. Details on gene size, SNP-coverage and pathway mapping are given in supplemental Table S1 (**see supporting files, Table S1**). Gene specific allele dosage files in MACH format were used for analysis.

Statistical analysis

Characterization of study population

The distribution of sex, age, baseline lung function parameters, their change during follow-up as well as packyears exposure during follow-up was tabulated according to categories of smoking status (never, former and current smokers) and interval PM10 exposure (high versus low exposure, defined by the median value). To assess a potential impact of loss to follow-up on our results, our study population consisting of up to 669 non-asthmatic adults with high quality genome-wide data and complete information on model covariates was compared to non-asthmatic participants examined at follow-up without genome-wide data (n=3833), and to those completing only baseline examination (n=1299) by means of descriptive tables and tests for independent samples (see supporting files, **Table S2**).

Gene- and pathway-environment interaction analysis

The interaction of genetic variation and exposure to PM10 or tobacco smoke on lung function decline was assessed in different stages.

First, SNP level analyses on decline in FEV_1, FEV_1/FVC and FEF_{25-75} were done for each gene separately using multiple linear regression in ProbABEL v0.1.3 (http://www.genabel.org) [51] with robust sandwich-estimation of standard errors. Models specified an additive SNP-effect, main effects for packyears smoked and interval PM10 exposure between surveys, and an interaction term between the SNP-variable and either exposure. They adjusted for sex, age and height at follow-up, packyears smoked up to baseline, principal components of population ancestry, and study area. No adjustment for ageing during follow-up was made, since follow-up time was 11 years for all participants. Complete data including covariates and environmental exposures was available on 669 participants for FEV_1 decline, and on 650 for FEV_1/FVC and FEF_{25-75} decline.

We used a slightly modified version of the Adaptive Rank Truncation Product (ARTP) method described by Yu and colleagues [52] to calculate gene- and pathway level p-values. Briefly, according to the method, SNPs are sorted in ascending order of interaction strength, and SNP-interaction p-values are multiplied up to several pre-specified truncation points which depend on

Table 1 Mapping of candidate oxidative-stress genes to molecular pathways of interest.

PATHWAY	GENES
Apoptosis Signaling	BCL2 BCL2L1 CASP6 CDK1 CHUK MAP2K1 NFKB1 PLCG1 PRKCA RELA TP53
Arachidonic Acid Metabolism	ALOX12 CYP1A1 CYP1A2 DHRS2 EPHX2 GPX1 GPX2 GPX3 GPX4 GPX5 GPX6 GPX7 GPX8 GSTK1 GSTT1 GSTZ1 MGST2 MGST3 PLA2G4A PRDX6 PTGS1 PTGS2
Aryl Hydrocarbon Receptor Signaling	ARNT CDKN1A CYP1A1 CYP1A2 EP300 FOS GSTK1 GSTM1 GSTM2 GSTM3 GSTM4 GSTM5 GSTO1 GSTO2 GSTP1 GSTT1 GSTT2 JUN MGST1 MGST2 MGST3 NFE2L2 NFKB1 NQO1 NQO2 RELA TP53
fMLP Signaling in Neutrophils	MAP2K1 NCF2 NFKB1 NOX3 NOX4 PLCB1 PRKCA RAC1 RELA
Glutathione Metabolism	GCLC GCLM GLRX GPX1 GPX2 GPX3 GPX4 GPX5 GPX6 GPX7 GPX8 GSR GSS GSTK1 GSTM1 GSTM2 GSTM3 GSTM4 GSTM5 GSTO1 GSTO2 GSTP1 GSTT1 GSTT2 GSTZ1 IDH1 MGST1 MGST2 MGST3 PRDX6
IL-6 Signaling	CHUK COL1A1 FOS GRB2 JAK2 JUN MAP2K1 MAPK14 NFKB1 RELA
Metabolism of Xenobiotics by Cytochrome P450	AKR1A1 CYP1A1 CYP1A2 DHRS2 EPHX1 GSTK1 GSTM1 GSTM2 GSTM3 GSTM4 GSTM5 GSTO1 GSTO2 GSTP1 GSTT1 GSTT2 GSTZ1 MGST1 MGST2 MGST3
Methane Metabolism	CAT EPX LPO MPO PRDX1 PRDX2 PRDX5 PRDX6 TPO
Mitochondrial Dysfunction	CAT GLRX2 GPX4 GPX7 GSR NDUFA12 NDUFA13 NDUFA6 NDUFS1 NDUFS2 NDUFS3 NDUFS4 NDUFS8 PARK2 PARK7 PRDX3 PRDX5 PSEN1 SNCA SOD2 TXN2 TXNRD2 UCP2
NF-κB Signaling	CHUK EGFR EP300 INSR NFKB1 RAC1 RAC2 RELA RIPK1 TGFBR2 TLR4
NRF2-mediated Oxidative Stress Response	ABCC1 AKR1A1 AKR7A2 AKR7A3 AOX1 CAT EP300 EPHX1 FOS FOSL1 GCLC GCLM GPX2 GSR GSTK1 GSTM1 GSTM2 GSTM3 GSTM4 GSTM5 GSTO1 GSTO2 GSTP1 GSTT1 GSTT2 HMOX1 JUN KEAP1 MAP2K1 MAPK14 MGST1 MGST2 MGST3 NFE2L2 NQO1 NQO2 PRDX1 PRKCA SOD1 SOD2 SOD3 TXN TXNRD1
Oxidative Phosphorylation	NDUFA12 NDUFA13 NDUFA6 NDUFS1 NDUFS2 NDUFS3 NDUFS4 NDUFS8
Production of Nitric Oxide and Reactive Oxygen Species in macrophages	CAT CHUK CYBA FOS JAK2 JUN MAP2K1 MAPK14 MPO NCF2 NFKB1 NOS2 PLCG1 PPP2CB PRKCA RAC1 RAC2 RELA STAT1 TLR4
Xenobiotic Metabolism Signaling	ARNT CAT CYP1A1 CYP1A2 EP300 FMO2 GCLC GSTK1 GSTM1 GSTM2 GSTM3 GSTM4 GSTM5 GSTO1 GSTO2 GSTP1 GSTT1 GSTT2 HMOX1 KEAP1 MAP2K1 MAPK14 MGST1 MGST2 MGST3 NFE2L2 NFKB1 NOS2 NQO1 NQO2 PPP2CB PRKCA RELA SOD3
NOT MAPPED TO PATHWAY	AATF AGT AGTR1 ATOX1 CCL5 CP CRISP2 CYGB DHCR24 DUSP1 ERCC1 GLRX3 GLRX5 GSTCD HMOX2 HP MSRA MT2A NAPRT1 NOS1 NOS3 NOX5 NOXO1 OGG1 OXR1 PNKP PSMB5 PTK2B PXDN PYCR1 SCARA3 SEPP1 SLC23A2 SRXN1 STK25 TXNIP

the number of SNPs in the gene. The statistical significance of these products is derived using the empirical distribution of products observed in the original and permutated datasets. For each gene, the strongest product p-value across all truncation points is readjusted using again its empirical distribution, to result in the gene-level p-value. Using the gene-level p-values in observed and permutated datasets, the same procedure can be applied to calculate pathway-level p-values. Details on the ARTP method, the applied modifications and truncation point definitions are presented in the supporting online material (**see supporting files, Figure S2 and Methods S1**).

SNP-level analyses were run 10000 times, always after having newly permuted gene-specific SNP-allele-dosages across participants. SNP-level interaction p-values of the observed and permutated datasets were used for calculating gene- and pathway-level p-values. According to Yu et al. [52], results from simulation studies suggest the ARTP-method yields type I error rates close to 5%. We thus additionally corrected for 152 tests at the gene and 14 tests at the pathway level in a first look. In a second line of investigation, a non-stringent nominal threshold of α=0.05 was chosen for further exploring gene- and pathway-level interaction signals due to our restricted sample size.

Comparing the impact of PM10 versus tobacco smoking

Emerging patterns of interaction were compared between exposures at the pathway- and gene-level. In pathways with nominally significant interactions, gene-level p-values were plotted against each other to identify the relative contributions to the pathway signal.

For the SNP with the strongest interaction signal in each nominally significant gene regression analyses were repeated with exposure centered to the median. Effect estimates were scaled to represent an exposure contrast of one interquartile range (IQR), and interaction effect sizes were compared between PM10 and tobacco smoke exposure. For SNP rs2035268 in gene *SNCA*, which was one of the top interaction signals in FEV_1/FVC decline, genotype specific estimates for PM10 and packyears exposure were calculated to exemplify the effect modification by genotype. To this purpose, imputed allele dosages were coded as genotypes as follows: dosage <0.5 genotype TT, 0.5≤ dosage <1.5 genotype GT, and dosage ≥1.5 genotype GG. Reparametrization of exposure variables into genotype specific ones was employed to avoid model-overspecification and instable estimation in small genotype strata (rs2035268: MAF 5%).

Statistical power

Power calculations were done using QUANTO software [53] version 1.2 specifying a gene-environment study on independent individuals. Details of the power calculation are given in the online supplement (**see supporting files, Methods S1**). The most important aspect of the calculation was that a two-sided significance threshold of 5% was used (i.e. no multiple testing correction was included), since all 12679 SNP-estimates were further processed for deriving gene- and pathway level p-values without filtering by association strength. In our first analysis with 650 subjects, we have at least 75% power to detect a SNP*environment interaction that accounts for 1% of the total variance and that power increases to 99% when the SNP*environment interaction accounts for 5% of the total variance. In the replication analysis with n=3320, estimated power is 99% in both cases. Statistical power is expected to be higher for the gene and pathway level analysis, but that increase in power could not be quantified since p-values for

interaction at the gene (or pathway) level are obtained from individual p-values for interactions with SNPs belonging to the gene (or pathway), and the effect of interaction may vary among SNPs.

Results

Characteristics of study population

Regarding the distribution of sex, age and lung function according to categories of smoking and PM10 exposure, our study sample on average presented decreasing lung function values and accelerated lung function decline with increased smoking (**Table 2**). The percentage of females decreased with smoking exposure. Compared to participants assessed only at baseline, our study sample had slightly better lung function values, substantially less current smokers, was slightly less exposed to PM10 and tobacco smoke, and was older and leaner (**see supporting files, Table S2**).

SNP-level analysis

A SNP-level analysis correcting for 12679 tests (α=0.05/12679=3.9 x 10^{-6}) detected an interaction between SNP rs2035268 in gene ***synuclein alpha*** (***SNCA***) on chromosome 4q21 and PM10 on FEV_1/FVC decline ($p_{interaction}$=2.5 x 10^{-6}). Compared to the baseline TT genotype, each G-allele was associated with a 3.8% (95% confidence interval (95%-CI) 2.2 to 5.4%) higher decline per 83.4 µg/m3*year PM10 exposure (IQR) over 11 years. Further, rs12190800 located in gene Parkinson disease protein 2 (*PARK2)* on chromosome 6q25.2 interacted with PM10 on FEV_1 decline. Compared to the TT-genotype each C-allele entailed an attenuation of 95.1 ml (95%-CI 60.1ml to 130.1ml) in FEV_1 decline per IQR of PM10 ($p_{interaction}$=9.8 x 10^{-8}). Exposure and outcome specific regression estimates for all 12679 SNPs are given in the supporting online information (**see supporting** files, **Data S1**).

Gene-level analysis

In the gene-level analysis, nominally interacting genes differed between PM10 and packyears exposure across the parameters of lung function decline (**Table 3**). Genes interacting with PM10 exposure partially overlapped for FEV_1/FVC and FEF_{25-75} decline (genes *CRISP2*, *ERCC1*, *LPO*, *MPO*, and *SNCA*). After correcting for performing 152 gene-level tests ($\alpha_{Bonferroni}$=0.05/152=3.29* 10^{-4}), the interaction between gene ***cysteine-rich secretory protein 2*** (***CRISP2***) located on chromosome 6p12.3 and interval PM10 exposure on FEV_1/FVC decline remained significant ($p_{interaction}$=3.0x10^{-4}). A marginally significant interaction was seen for gene *SNCA* on chromosome 4q21 with the same outcome and exposure ($p_{interaction}$=4.0x10^{-4}). Interactions observed for packyears exposure did not withstand multiple testing corrections.

P-values of interaction for all tested genes are given in the supporting online material (**see supporting files, Table S3**).

Table 2 Distribution of main characteristics by smoking status at follow-up and PM_{10} exposure during follow-up (N=650[1]) .

		Interval PM_{10} EXPOSURE *(median:239.0µg/m³*years)*					
SMOKING CATEGORY		**exposure < median**			**exposure≥ median**		
	variable	*value*[2]	*range*		*value*[2]	*range*	
Never	n	**152**			**145**		
Smoker	female sex [%]	**67.1**			**53.8**		
	age at follow-up [years]	**53.0**	31.1	71.5	**52.7**	29.8	71.8
	FEV1 [L]	**3.4**	2.0	5.2	**3.6**	2.4	6.0
	FEV1 decline [L]	**-0.3**	-1.1	1.0	**-0.3**	-1.5	0.5
	FEV_1/FVC [%]	**79.2**	61.4	98.0	**81.0**	62.5	99.8
	FEV_1/FVC decline [%]	**-3.9**	-13.2	9.4	**-4.4**	-21.6	8.2
	FEF_{25-75} [L/sec]	**3.2**	1.3	7.5	**3.6**	1.2	6.8
	FEF_{25-75} decl. [L/sec]	**-0.7**	-2.5	2.1	**-0.8**	-3.7	0.6
	pack years d. follow-up	**n.a.**			**n.a.**		
Former	n	**98**			**102**		
Smoker	female sex [%]	**40.8**			**52.9**		
	age at follow-up [years]	**54.1**	32.7	72.0	**54.4**	30.8	71.9
	FEV1 [L]	**3.8**	2.2	5.4	**3.6**	2.3	5.6
	FEV1 decline [L]	**-0.4**	-1.2	0.3	**-0.4**	-1.4	0.6
	FEV_1/FVC [%]	**78.8**	60.0	97.3	**80.6**	66.4	95.2
	FEV_1/FVC decline [%]	**-3.2**	-16.1	12.4	**-4.7**	-20.7	13.7
	FEF_{25-75} [L/sec]	**3.6**	1.4	7.7	**3.6**	1.5	7.2
	FEF_{25-75} decl. [L/sec]	**-0.8**	-2.8	1.7	**-0.8**	-3.6	1.6
	pack years d. follow-up	**0.0**	0.0	25.0	**0.0**	0.0	35.0
Current	n	**75**			**78**		
Smoker	female sex [%]	**37.3**			**46.2**		
	age at follow-up [years]	**52.8**	29.8	70.8	**49.7**	30.3	70.6
	FEV1 [L]	**3.6**	2.4	5.7	**3.7**	1.8	6.8
	FEV1 decline [L]	**-0.5**	-1.5	0.1	**-0.4**	-1.3	0.3
	FEV_1/FVC [%]	**77.6**	59.3	94.5	**79.0**	49.0	97.1
	FEV_1/FVC decline [%]	**-6.3**	-20.5	3.9	**-4.9**	-21.5	7.9
	FEF_{25-75} [L/sec]	**3.3**	1.4	7.4	**3.5**	0.7	7.3
	FEF_{25-75} decl. [L/sec]	**-1.0**	-3.0	0.4	**-0.8**	-2.4	0.7
	pack years d. follow-up	**10.9**	0.0	27.3	**9.0**	0.0	24.0

[1] restricted to sample with complete data on all three lung function parameters. Regarding FEV_1, sample size with complete covariate data would be n=669

[2] means for age and lung function values, medians for pack years exposures.

Table 3 Nominally significant gene-environment interactions by outcome and exposure.

Outcome	**decline FEV_1 (n=669)**		**decline FEV_1/FVC (n=650)**		**decline FEF_{25-75} (n=650)**	
Exposure	**interval PM10**	**packyears**	**interval PM10**	**packyears**	**interval PM10**	**packyears**
Gene ($p_{interaction}$)	CP (0.005)	BCL2 (0.003)	**CRISP2 (0.0003)**[a]	PSMB5 (0.003)	LPO (0.008)	TGFBR2 (0.006)
	PRDX3 (0.010)	PTK2B (0.017)	**SNCA (0.0004)**[b]	SOD2 (0.015)	ERCC1 (0.014)	PTK2B (0.033)
	ERCC1 (0.014)	PSEN1 (0.023)	ERCC1 (0.007)	MAP2K1 (0.019)	MPO (0.018)	TP53 (0.033)
	RAC1 (0.027)	NOXO1 (0.034)	ALOX12 (0.012)	NFKB1 (0.022)	SLC23A2 (0.022)	CASP6 (0.039)
	CYP1A2 (0.028)	AOX1 (0.044)	LPO (0.018)	HMOX2 (0.048)	CRISP2 (0.023)	OXR1 (0.045)
	PSMB5 (0.038)	MAP2K1 (0.046)	CHUK (0.035)		SNCA (0.025)	TXNRD2 (0.047)
	GLRX (0.046)		GPX5 (0.039)		GPX5 (0.026)	
	GLRX2 (0.048)		MPO (0.039)		COL1A1 (0.049)	
			EPX (0.040)			

Genes are sorted in ascending order of interaction p-value within outcome-exposure strata

[a] significant after Bonferroni-correction for testing 152 genes (α=.00033)

[b] marginally significant after Bonferroni-correction for testing 152 genes (α=.00033)

Pathway-level analysis

Pathways **"mitochondrial dysfunction"** and **"methane metabolism"** interacted nominally (α=0.05) with PM10 on FEV_1/FVC decline (p=0.017) and FEF_{25-75} decline (p=0.029), respectively. A further interaction signal was observed for pathway **"apoptosis"** and packyears exposure on FEV_1-decline (p=0.051). Inspecting the interaction p-values of pathway-specific genes revealed that the pathway signals mostly arose from single genes (*SNCA* in pathway "mitochondrial dysfunction") or single genomic loci (overlapping gene regions of genes *eosinophil peroxidase, EPX, lactoperoxidase*, *LPO*, and *myeloperoxidase, MPO* in pathway "methane metabolism") (**Figure 1, parts A-C**). P-values of interaction for all tested pathways are given in the supporting online material (**see supporting files, Table S3**).

Comparison of interactions with PM10 versus packyears exposure

The comparison of interaction effect sizes for PM10 and packyears exposure was based on regression estimates for the strongest interacting SNP only within each nominally significant gene. **Table 4** presents estimates for FEV_1/FVC decline, where significant and marginally significant gene-level interactions have been detected for genes *CRISP2* and *SNCA*, respectively. Estimates for decline in FEV_1 and FEF_{25-75} are presented in the supporting online material (**see supporting files, Table S4 and Table S5**).

The C-allele of SNP rs360563 in gene *CRISP2* accelerated FEV_1/FVC decline by 1.1 % per IQR change in PM10 exposure over 11 years (**Table 4**). Similarly, the G-allele of SNP rs2035268 in *SNCA* was associated with an accelerated decline by 3,8% per allele and IQR change in exposure. Genotype specific exposure estimates were calculated for rs2035268. Within genotypes GT and GG of SNP rs2035268, a change in IQR of PM10 was associated with a signficant acceleration of FEV_1/FVC decline by 3.9%, opposed to a small and non-signficiant acceleration by 0.2% in baseline genotype TT (**Table 5**). In contrast, a change in IQR of packyears smoked was associated with a significant acceleration by 1.1% in the baseline TT genotype stratum, but not in the GT/GG strata.

For FEV_1- and FEF_{25-75} decline, interaction effect sizes for the strongest interacting SNPs in nominally significant genes tended to be considerably larger with packyears compared to PM10 exposure. Further, packyears exposure frequently presented significant main effects besides the interaction with SNPs (**supporting files, Table S4 and Table S5**).

In models including only main effects but no interaction between SNPs and exposure, an IQR of 9.8 packyears was significantly associated with accelerated decline in FEV_1/FVC by 1%, and in FEV_1 by 50ml (data not shown). Respective estimates for PM10 were non-significant. SNP main effects remained non-significant and their beta estimates largely unaffected by the exclusion of interaction terms.

Table 4 Effect estimates of the strongest interacting SNP from each nominally significant gene on FEV_1/FVC decline (n=650).

Exposure	Chrom	Position	Gene	SNP	type	All1	All2	Freq All1	$Beta_{interaction}$ (SE), P	$Beta_{SNP}$ (SE), P	$Beta_{exposure}$ (SE), P
PM10	4	90975104	***SNCA***	rs2035268	imp	G	T	0.05	**-3.8** (0.8), *2.54E-06**	-0.7 (0.6), *0.254*	-0.2 (0.7), *0.786*
(IQR. 83.4	6	49766228	***CRISP2***	rs360563	imp	C	T	0.50	**-1.1** (0.3), *3.78E-05*	0.0 (0.3), *0.975*	0.6 (0.7), *0.375*
ug/m3 * y)	17	6840800	***ALOX12***	rs2073438	gen	A	G	0.26	**1.0** (0.3), *2.38E-04*	0.4 (0.3), *0.181*	-1.0 (0.7), *0.144*
	17	53675156	***LPO***	rs8178290	imp	A	C	0.18	**1.1** (0.3), *9.61E-04*	-0.2 (0.3), *0.582*	-1.0 (0.7), *0.153*
	17	53629132	***EPX***	rs3785496	gen	A	G	0.80	**-1.1** (0.3), *0.001*	-0.1 (0.3), *0.773*	1.1 (0.8), *0.140*
	17	53699864	***MPO***	rs8178409	imp	A	G	0.18	**1.1** (0.3), *0.001*	-0.2 (0.3), *0.523*	-1.0 (0.7), *0.158*
	19	50600888	***ERCC1***	rs1005165	imp	C	T	0.83	**-1.3** (0.4), *0.002*	0.1 (0.4), *0.765*	1.5 (0.9), *0.084*
	6	28629296	***GPX5***	rs393414	gen	C	T	0.79	**-0.9** (0.3), *0.003*	0.3 (0.3), *0.385*	1.0 (0.8), *0.219*
	10	101996416	***CHUK***	rs4919438	imp	C	T	0.50	**-0.8** (0.3), *0.003*	-0.1 (0.2), *0.669*	0.2 (0.7), *0.813*
packyears	14	22552780	***PSMB5***	rs12590429	imp	A	G	0.09	**-3.8** (0.9), *1.06E-05*	0.3 (0.5), *0.540*	-0.5 (0.5), *0.265*
(IQR: 9.8	6	160020288	***SOD2***	rs7855	imp	A	G	0.94	**2.7** (0.7), *2.17E-04*	-0.4 (0.8), *0.620*	-6.1 (1.4), *1.64E-05*
PY)	15	64583032	***MAP2K1***	rs8043062	imp	A	G	0.15	**1.9** (0.6), *0.001*	-0.1 (0.3), *0.741*	-1.6 (0.5), *0.003*
	4	103676616	***NFKB1***	rs230528	gen	G	T	0.38	**-1.7** (0.6), *0.003*	0.1 (0.3), *0.775*	0.2 (0.6), *0.693*
	16	4466293	***HMOX2***	rs2270363	imp	A	G	0.25	**1.1** (0.4), *0.013*	0.0 (0.3), *0.935*	-1.8 (0.6), *0.002*

SNP-estimates are based on an additive model. Beta-estimates represent percentages of decline in FEV_1/FVC over 11 years per effect allele and/or for an exposure contrast of one interquartile range (IQR). All estimates are taken from the same interaction model. Positive values mean an attenuation, and negative ones an acceleration of FEV_1/FVC decline. Rows are sorted according to ascending interaction p-values.

*significant after Bonferroni correction for testing 12679 SNPs (α=3.9 x 10E-6)

gen: genotyped SNP; imp: imputed SNP; All1: allele 1 (effect allele); All2: allele 2 (baseline allele); FreqAll1: frequency of allele 1

Table 5 rs2035268 genotype specific estimates of the effect of interval PM10 and pack years exposure on percentage decline in FEV_1/FVC ratio during 11 years of follow-up.

exposure	rs2035268 genotype	effect estimate[a]	(95%-confidence interval)	p-value	$p_{interaction}$[b]
interval PM10	wild-type (TT)	**-0.2**	(-1.7 to 1.4)	0.827	**7.35E-07**
(IQR 83.4ug/m^3* y)	mutant (GT/GG)	**-3.9**	(-5.9 to -1.8)	2.25E-04	
packyears	wild-type (TT)	**-1.1**	(-2.0 to -0.1)	0.024	0.909
(IQR 9,8 PY)	mutant (GT/GG)	**-0.7**	(-2.3 to 1.0)	0.434	

[a] Environmental effect estimates are based on a multiple linear model with sample size n=650 adjusting for sex, age and height at follow-up, packyears smoked up to baseline, population ancestry, and study area. PM10 and packyears exposure was reparametrized into genotype specific exposure variables to avoid model overspecification with instable estimates in the genotypic risk stratum (rs2035268 has minor allele frequency of 0.05). Estimates are in units of percentage decline in FEV_1/FVC

[b] p-value of interaction between environmental exposure and genotypes of rs2035268 (TT vs GT/GG)

Replication of significant associations

Replication genotyping was done for *CRISP2* SNP rs360563 (MAF of 49.8%) and rs12190800 in *PARK2* (MAF 16%), but their interaction with PM10 exposure on FEV_1/FVC and FEV_1 decline could not be confirmed in the remainder of the SAPALDIA population ($p_{interaction}$=0.63 and 0.50 respectively, n=3320 for both). Thereby, MAFs in the replication sample corresponded to those in the discovery sample, and both SNPs were in Hardy-Weinberg equilibrium.

Discussion

To the best of our knowledge, this is the first study assessing gene-environment interactions on lung function decline using analysis methods that accumulate interaction effects along a broadly defined set of candidate genes and pathways. Our results suggest that different oxidative stress genes could be involved in mediating the adverse effects of ambient air pollution and tobacco smoke exposure on lung function decline.

We can currently only hypothesize about the reason for observing different patterns of interaction between the two environmental exposures. A possible explanation would be that ambient particulate matter pollution and tobacco smoke, although sharing many constituents, also differ in their composition, which possibly affects the overall and relative relevance of the different pathways. A probably more important explanation is that levels of oxidative stress imposed by ambient PM10 exposure are much lower than those induced by active tobacco smoking. Experimental studies have shown that different levels of oxidative stress trigger dose-dependent, specific activations of pathways on the cellular level in response to the oxidant burden [54]. Li and colleagues delineated a **stratified oxidative stress model** while studying the biological effects of particulate matter exposure on human and mouse cell lines exposed to solutions of Diesel exhaust particles (DEP) and concentrated ambient air particles (CAP) sampled in a highly polluted area [55,56]. According to their observations, at the lower end of exposure pivotal ROS-scavenging enzymes like heme oxygenase-1 are induced, representing the activation of protective cell-mechanisms. Intermediate exposure levels trigger inflammatory pathways via signal transduction cascades (increased expression of interleukin-8 and Jun kinase), while high exposure levels impact on mitochondrial permeability, and result in cytotoxicity and apoptosis. Thereby CAP were mostly representing the lower to mid-level of exposure, inducing oxidative-stress enzymes and inflammation, but not apoptosis (as observed with DEP). In contrast, tobacco smoke exposure is known to induce the whole spectrum of cellular reactions, from oxidative stress response and inflammation [57,58] up to DNA-damage [58], apoptosis [59,60,61,62] as well as cellular necrosis [61]. Although in the light of limited sample size, we cannot provide statistical evidence of exposure-specific interaction patterns with genes and pathways in our current study, it is interesting to see that many of the top-ranking genes interacting with packyears exposure are involved in signal transduction or apoptosis (**Table 3** *BCL2, CASP6, MAP2K1, NFκB1, TGFBR2, TP53*). Only two such genes showed interaction signals with PM10 exposure (*CHUK, RAC1*), and many of the others related to scavenging or production of ROS (*CRISP2, CYP1A2, EPX, GLRX, GLRX2, GPX5, LPO, MPO, PRDX3*). These observations are consistent with the stratified oxidative stress model. The observation of larger interaction effect sizes at the level of SNPs for FEV_1 and FEF_{25-75} decline, as well as the frequent presence of significant main effects further support higher oxidative stress levels induced by tobacco smoke than PM10 exposure.

Another important observation was that the effect of genetic variation related to oxidative stress appeared to be mediated predominantly by the interactions with environmental exposures, as hardly any SNP main-effects were observed. This is in line with the findings of genome-wide studies on lung function performed to date [17,18,21], where oxidative-stress related candidate genes did not produce strong signals. But their design was cross-sectional and importantly, these analyses focused on SNP-main effects. Exposure specific gene-effects might thus be missed as they can cancel out when averaged over the whole population (which happens in a gene main

effect analysis). Disregarding gene-environment interaction might also explain part of the missing heritability in complex disease genetics.

Our study had several limitations. First, the limited number of non-asthmatic adults with available genome-wide data restricted our power to detect associations at the gene and pathway levels. In this context, we faced the problem of finding studies with genome-wide genotyping and comparable data on both phenotypes and environmental exposures. This issue is particularly imminent regarding ambient air pollution exposure. As a consequence, small sample size did not allow us to identify further strong interaction signals to follow-up, while the observed ones could not be replicated in the remainder of the study population. Our gene and pathway level results are thus of more exploratory nature. Limited power is also known to inflate effect estimates when the strongest association signals are selected for further follow-up (so-called "winner's curse" [63]), thus our interaction effect estimates on the SNP-level are likely overestimated for both exposures. But the relative difference in effect size between exposures is probably less affected by this phenomenon. In case of differential overestimation, the true difference would likely be larger, as observed PM10 effects were smaller and therefore would be more affected than packyears effects. Further, follow-up participants were healthier than those completing only baseline examination. Our results are thus applicable to an adult general population sample of good health. Environmental exposure and genetic susceptibility might possibly have affected health and thus participation of our study subjects. But in this case, true effects would likely be underestimated in our present study [64]. Finally, SNP-coverage was low for certain genes (**see supporting files, Table S1**), and the well-known gene-deletions in glutathione S-transferases are difficult to tag by SNP-genotyping as they represent copy number variations. This makes it difficult to interpret respective results. On the other side, a comparison of imputed SNP data for rs360563 (gene *CRISP2*) with genotypes measured during replication in the initial study sample showed a high concordance indicating high imputation quality (**see supporting files, Table S6**). The absence of strong interactions on the pathway level is likely due to our primary focus on function while selecting candidate genes, which limited pathway coverage. But genes in a pathway may also differently interact with exposure, or compensate for each other. Further, regulatory genomic regions could be located farer away than the chosen flanking segments of 20 kilobases. Detecting interactions in pathways is thus more challenging.

Strengths of our study were the population based design comprising non-asthmatic adults of a wide age-spectrum, the detailed data on individual tobacco smoke and particularly PM10 exposure, and the high quality of longitudinal lung function data. Finally, the application of analysis methods which exploit interaction signals below the significance threshold of a pure SNP-level analysis provided new insight into a possible differential involvement of genes according to exposure specific oxidative stress levels.

Conclusions

Applying a gene- and pathway-level analysis, we observed that PM10 and packyears exposure potentially interact with different genes on lung function decline, consistent with a stratified response to different oxidative stress levels. Our study thus points to the importance of considering interactions with environmental factors in the search for molecular pathways underlying lung function decline in response to exogenous inhalants. But it is also a good example of the challenges faced by gene-environment interaction studies today: While studies with partial genome-wide data, and hence often small sample size, can beneficially use the remainder of the study population as highly comparable replication sample, their potential to identify sufficient variants to follow-up is limited. In contrast, large studies or study consortia are more powerful in the discovery stage, but suffer from data heterogeneity as finding suitable replication studies with comparable phenotypic, genetic and environmental exposure data is difficult. This results in a challenging trade-off between sample size and data homogeneity.

Acknowledgements:

The **SAPALDIA study** could not have been done without the help of the study participants, technical and administrative support, scientific teams at coordinating and local study sites, and the medical teams and field workers at the local study sites. An overview of the SAPALDIA team as of July 2011 is given in the online appendix (**supporting files, Text S1**).

SAPALDIA was part of the **GABRIEL** Consortium, a European 6th Framework Research project on asthma genetics, which allowed us to obtain the genotype information used in this analysis. We acknowledge **Dr. Kai Yu**, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville MD, USA, and **Prof. Roel Vermeulen**, University Medical Centre, Utrecht University, The Netherland for their help in making the ARTP-method available. We also acknowledge **Philippe Dessen**, Plateforme de Biologie intégrée, Bioinformatique, Institut Gustave Roussy IGR, Villejuif, France for his help with the work related to Ingenuity Pathways Analyis®.

References

1. Burney PG, Hooper R (2011) Forced vital capacity, airway obstruction and survival in a general population sample from the USA. Thorax 66: 49-54.
2. Stavem K, Aaser E, Sandvik L, Bjørnholt JV, Erikssen G, et al. (2005) Lung function, smoking and mortality in a 26-year follow-up of healthy middle-aged males. European Respiratory Journal 25: 618-625.
3. Schünemann HJ, Dorn J, Grant BJB, Winkelstein W, Trevisan M (2000) Pulmonary Function Is a Long-term Predictor of Mortality in the General Population. Chest 118: 656-664.
4. Lange P, Nyboe J, Appleyard M, Jensen G, Schnohr P (1990) Spirometric findings and mortality in never-smokers. Journal of Clinical Epidemiology 43: 867-873.
5. Eisner MD, Wang Y, Haight TJ, Balmes J, Hammond SK, et al. (2007) Secondhand Smoke Exposure, Pulmonary Function, and Cardiovascular Mortality. Annals of Epidemiology 17: 364-373.
6. Carta P, Aru G, Barbieri MT, Avataneo G, Casula D (1996) Dust exposure, respiratory symptoms, and longitudinal decline of lung function in young coal miners. Occupational and Environmental Medicine 53: 312-319.
7. Kauffmann F, Drouet D, Lellouch J, Brille D (1982) Occupational exposure and 12-year spirometric changes among Paris area workers. British Journal of Industrial Medicine 39: 221-232.
8. Johnsen HL, Hetland SM, Benth JS, Kongerud J, Soyseth V (2010) Dust Exposure Assessed by a Job Exposure Matrix Is Associated with Increased Annual Decline in FEV1: A 5-year Prospective Study of Employees in Norwegian Smelters. American Journal of Respiratory and Critical Care Medicine 181: 1234-1240.
9. Sunyer J, Zock JP, Kromhout H, Garcia-Esteban R, Radon K, et al. (2005) Lung function decline, chronic bronchitis, and occupational exposures in young adults. American Journal of Respiratory and Critical Care Medicine 172: 1139-1145.
10. Downs SH, Schindler C, Liu LJ, Keidel D, Bayer-Oglesby L, et al. (2007) Reduced exposure to PM10 and attenuated age-related decline in lung function. New England Journal Medicine 357: 2338-2347.
11. Rahman I, Adcock IM (2006) Oxidative stress and redox regulation of lung inflammation in COPD. European Respiratory Journal 28: 219-242.
12. MacNee W (2005) Pulmonary and Systemic Oxidant/Antioxidant Imbalance in Chronic Obstructive Pulmonary Disease. Proceedings of the American Thoracic Society 2: 50-60.
13. Hall IP, Lomas DA (2010) The genetics of obstructive lung disease: big is beautiful. Thorax 65: 760-761.
14. Silverman EK, Spira A, Pare PD (2009) Genetics and Genomics of Chronic Obstructive Pulmonary Disease. Proceedings of the American Thoracic Society 6: 539-542.
15. Smolonska J, Wijmenga C, Postma DS, Boezen HM (2009) Meta-analyses on suspected chronic obstructive pulmonary disease genes: a summary of 20 years' research. American Journal of Respiratory and Critical Care Medicine 180: 618-631.
16. He JQ, Connett JE, Anthonisen NR, Pare PD, Sandford AJ (2004) Glutathione S-transferase variants and their interaction with smoking on lung function. American Journal of Respiratory and Critical Care Medicine 170: 388-394.
17. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, et al. (2010) Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. Nature Genetics 42: 45-52.
18. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, et al. (2010) Genome-wide association study identifies five loci associated with lung function. Nature Genetics 42: 36-44.
19. Pillai SG, Ge D, Zhu G, Kong X, Shianna KV, et al. (2009) A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. PLoS Genetics 5: e1000421.

20. Cho MH, Boutaoui N, Klanderman BJ, Sylvia JS, Ziniti JP, et al. (2010) Variants in FAM13A are associated with chronic obstructive pulmonary disease. Nature Genetics 42: 200-202.
21. Wilk JB, Chen TH, Gottlieb DJ, Walter RE, Nagle MW, et al. (2009) A genome-wide association study of pulmonary function measures in the Framingham Heart Study. PLoS Genetics 5: e1000429.
22. Artigas MS, Loth DW, Wain LV, Gharib SA, Obeidat M, et al. (2011) Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. Nature Genetics. Epub ahead of print, doi: 10.1038/ng.941.
23. Obeidat Me, Wain LV, Shrine N, Kalsheker N, Artigas MS, et al. (2011) A Comprehensive Evaluation of Potential Lung Function Associated Genes in the SpiroMeta General Population Sample. PLoS One 6: e19382.
24. Pillai SG, Kong X, Edwards LD, Cho MH, Anderson WH, et al. (2010) Loci identified by genome-wide association studies influence different disease-related phenotypes in chronic obstructive pulmonary disease. American Journal of Respiratory and Critical Care Medicine 182: 1498-1505.
25. Ege MJ, Strachan DP, Cookson WO, Moffatt MF, Gut I, et al. (2011) Gene-environment interaction for childhood asthma and exposure to farming in Central Europe. The Journal of allergy and clinical immunology 127: 138-144, 144 e131-134.
26. Thomas D (2010) Methods for Investigating Gene-Environment Interactions in Candidate Pathway and Genome-Wide Association Studies. Annual Review of Public Health 31: 21-36.
27. Le Marchand L, Wilkens LR (2008) Design Considerations for Genomic Association Studies: Importance of Gene-Environment Interactions. Cancer Epidemiology Biomarkers & Prevention 17: 263-267.
28. Yang IA, Fong KM, Zimmerman PV, Holgate ST, Holloway JW (2008) Genetic susceptibility to the respiratory effects of air pollution. Thorax 63: 555-563.
29. Curjuric I, Imboden M, Schindler C, Downs SH, Hersberger M, et al. (2010) HMOX1 and GST variants modify attenuation of FEF25–75% decline due to PM10 reduction. European Respiratory Journal 35: 505-514.
30. Siedlinski M, Postma DS, van Diemen CC, Blokstra A, Smit HA, et al. (2008) Lung function loss, smoking, vitamin C intake, and polymorphisms of the glutamate-cysteine ligase genes. American Journal of Respiratory and Critical Care Medicine 178: 13-19.
31. Bouzigon E, Corda E, Aschard H, Dizier MH, Boland A, et al. (2008) Effect of 17q21 variants and smoking exposure in early-onset asthma. The New England Journal of Medicine 359: 1985-1994.
32. Van Durme YM, Eijgelsheim M, Joos GF, Hofman A, Uitterlinden AG, et al. (2010) Hedgehog-interacting protein is a COPD susceptibility gene: the Rotterdam Study. The European Respiratory Journal 36: 89-95.
33. Fridley BL, Biernacka JM (2011) Gene set analysis of SNP data: benefits, challenges, and future directions. European Journal of Human Genetics 19: 837-843.
34. Ackermann-Liebrich U, Kuna-Dibbert B, Probst-Hensch NM, Schindler C, Felber Dietrich D, et al. (2005) Follow-up of the Swiss Cohort Study on Air Pollution and Lung Diseases in Adults (SAPALDIA 2) 1991-2003: methods and characterization of participants. Sozial- und Präventivmedizin 50: 245-263.
35. Martin BW, Ackermann-Liebrich U, Leuenberger P, Kunzli N, Stutz EZ, et al. (1997) SAPALDIA: methods and participation in the cross-sectional part of the Swiss Study on Air Pollution and Lung Diseases in Adults. Sozial- und Präventivmedizin 42: 67-84.
36. Kunzli N, Ackermann-Liebrich U, Keller R, Perruchoud AP, Schindler C (1995) Variability of FVC and FEV1 due to technician, team, device and subject in an eight centre study: three quality control studies in SAPALDIA. Swiss Study on Air Pollution and Lung Disease in Adults. European Respiratory Journal 8: 371-376.
37. Kunzli N, Kuna-Dibbert B, Keidel D, Keller R, Brandli O, et al. (2005) Longitudinal validity of spirometers--a challenge in longitudinal studies. Swiss Medical Weekly 135: 503-508.

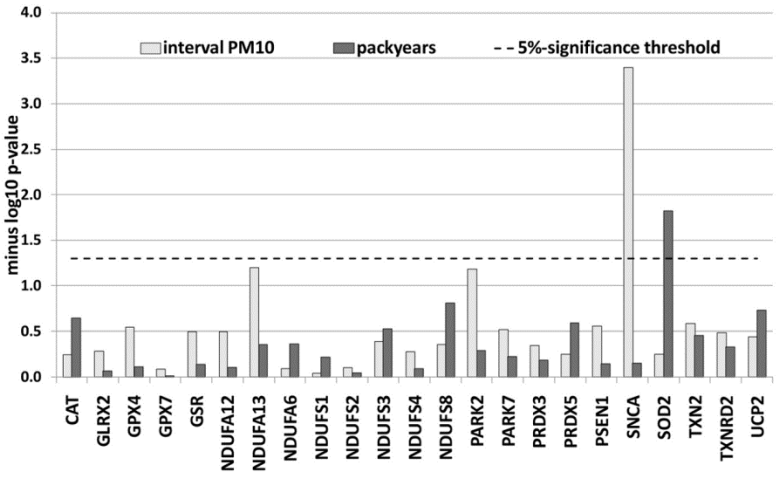
38. American Thoracic Society (1995) Standardization of Spirometry, 1994 Update. American Thoracic Society. American Journal of Respiratory and Critical Care Medicine 152: 1107-1136.
39. American Thoracic Society (1978) Epidemiology Standardization Project. e.c. American Thoracic Society. Recommended respiratory disease questionnaires for use with adults and children in epidemiological research. The American Review of Respiratory Disease 118: 7-52.
40. Liu LJ, Curjuric I, Keidel D, Heldstab J, Kunzli N, et al. (2007) Characterization of source-specific air pollution exposure for a large population-based Swiss cohort (SAPALDIA). Environmental Health Perspectives 115: 1638-1645.
41. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, et al. (2010) A large-scale, consortium-based genomewide association study of asthma. The New England Journal of Medicine 363: 1211-1221.
42. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genetic Epidemiology 34: 816-834.
43. International HapMap Consortium (2005) A haplotype map of the human genome. Nature 437: 1299-1320.
44. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics 25: 25-29.
45. Elliott NA, Volkert MR (2004) Stress induction and mitochondrial localization of Oxr1 proteins in yeast and humans. Molecular and Cellular Biology 24: 3180-3187.
46. Kaimul Ahsan M, Nakamura H, Tanito M, Yamada K, Utsumi H, et al. (2005) Thioredoxin-1 suppresses lung injury and apoptosis induced by diesel exhaust particles (DEP) by scavenging reactive oxygen species and by inhibiting DEP-induced downregulation of Akt. Free Radical Biology & Medicine 39: 1549-1559.
47. Nickel C, Trujillo M, Rahlfs S, Deponte M, Radi R, et al. (2005) Plasmodium falciparum 2-Cys peroxiredoxin reacts with plasmoredoxin and peroxynitrite. Biological Chemistry 386: 1129-1136.
48. Tomita M, Okuyama T, Katsuyama H, Hidaka K, Otsuki T, et al. (2006) Gene expression in rat lungs during early response to paraquat-induced oxidative stress. International Journal of Molecular Medicine 17: 37-44.
49. Tseng CF, Huang HY, Yang YT, Mao SJ (2004) Purification of human haptoglobin 1-1, 2-1, and 2-2 using monoclonal antibody affinity chromatography. Protein expression and Purification 33: 265-273.
50. Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, et al. (2004) The UCSC Table Browser data retrieval tool. Nucleic Acids Research 32: D493-496.
51. Aulchenko Y, Struchalin M, van Duijn C (2010) ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 11: 134-134.
52. Yu K, Li Q, Bergen AW, Pfeiffer RM, Rosenberg PS, et al. (2009) Pathway analysis by adaptive combination of P-values. Genetic Epidemiology 33: 700-709.
53. Gauderman W, Morrison J (2006) QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies, http://hydra.usc.edu/gxe.
54. Araujo JA, Nel AE (2009) Particulate matter and atherosclerosis: role of particle size, composition and oxidative stress. Particle and Fibre Toxicology 6: 24.
55. Li N, Hao M, Phalen RF, Hinds WC, Nel AE (2003) Particulate air pollutants and asthma. A paradigm for the role of oxidative stress in PM-induced adverse health effects. Clinical Immunology 109: 250-265.
56. Li N, Kim S, Wang M, Froines J, Sioutas C, et al. (2002) Use of a stratified oxidative stress model to study the biological effects of ambient concentrated and diesel exhaust particulate matter. Inhalation Toxicology 14: 459-486.
57. Comandini A, Marzano V, Curradi G, Federici G, Urbani A, Saltini C (2010) Markers of anti-oxidant response in tobacco smoke exposed subjects: a data-mining review. Pulmonary Pharmacology & Therapeutics 23(6): 482-92.

58. Faux SP, Tai T, Thorne D, Xu Y, Breheny D, et al. (2009) The role of oxidative stress in the biological responses of lung epithelial cells to cigarette smoke. Biomarkers : Biochemical Indicators of Exposure, Response, and Susceptibility to Chemicals 14 Suppl 1: 90-96.
59. Aoshiba K, Tamaoki J, Nagai A (2001) Acute cigarette smoke exposure induces apoptosis of alveolar macrophages. American Journal of Physiology - Lung Cellular and Molecular Physiology 281: L1392-1401.
60. Kim HP, Wang X, Chen ZH, Lee SJ, Huang MH, et al. (2008) Autophagic proteins regulate cigarette smoke-induced apoptosis: protective role of heme oxygenase-1. Autophagy 4: 887-895.
61. Slebos DJ, Ryter SW, van der Toorn M, Liu F, Guo F, et al. (2007) Mitochondrial localization and function of heme oxygenase-1 in cigarette smoke-induced cell death. American Journal of Respiratory Cell and Molecular Biology 36: 409-417.
62. van der Vaart H, Postma DS, Timens W, ten Hacken NH (2004) Acute effects of cigarette smoke on inflammation and oxidative stress: a review. Thorax 59: 713-721.
63. Kraft P (2008) Curses--winner's and otherwise--in genetic epidemiology. Epidemiology 19: 649-651; discussion 657-648.
64. Sacks JD, Stanek LW, Luben TJ, Johns DO, Buckley BJ, et al. (2011) Particulate matter-induced health effects: who is susceptible? Environmental Health Perspectives 119: 446-454.

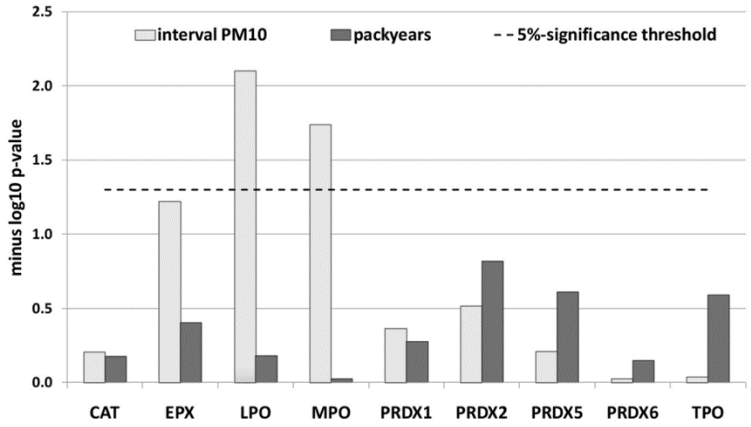
Figures

Figure 1 Distribution of interaction p-values across genes mapping to pathways with weak interaction signals.

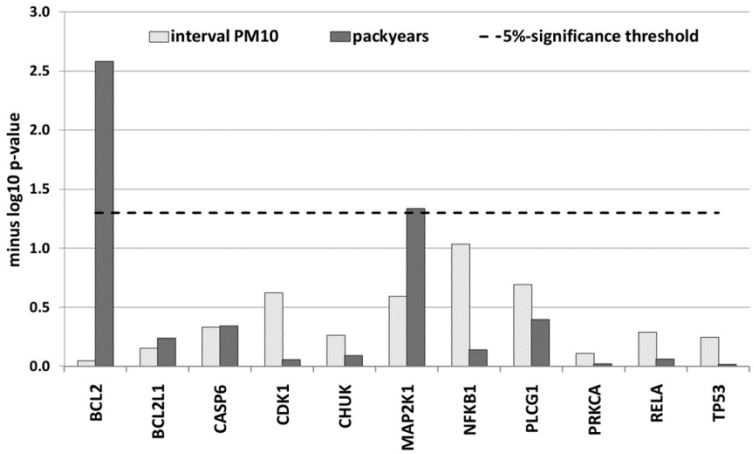
A) Mitochondrial dysfunction pathway genes & decline in FEV1/FVC



B) Methane metabolism pathway genes & decline in FEF_{25-75}



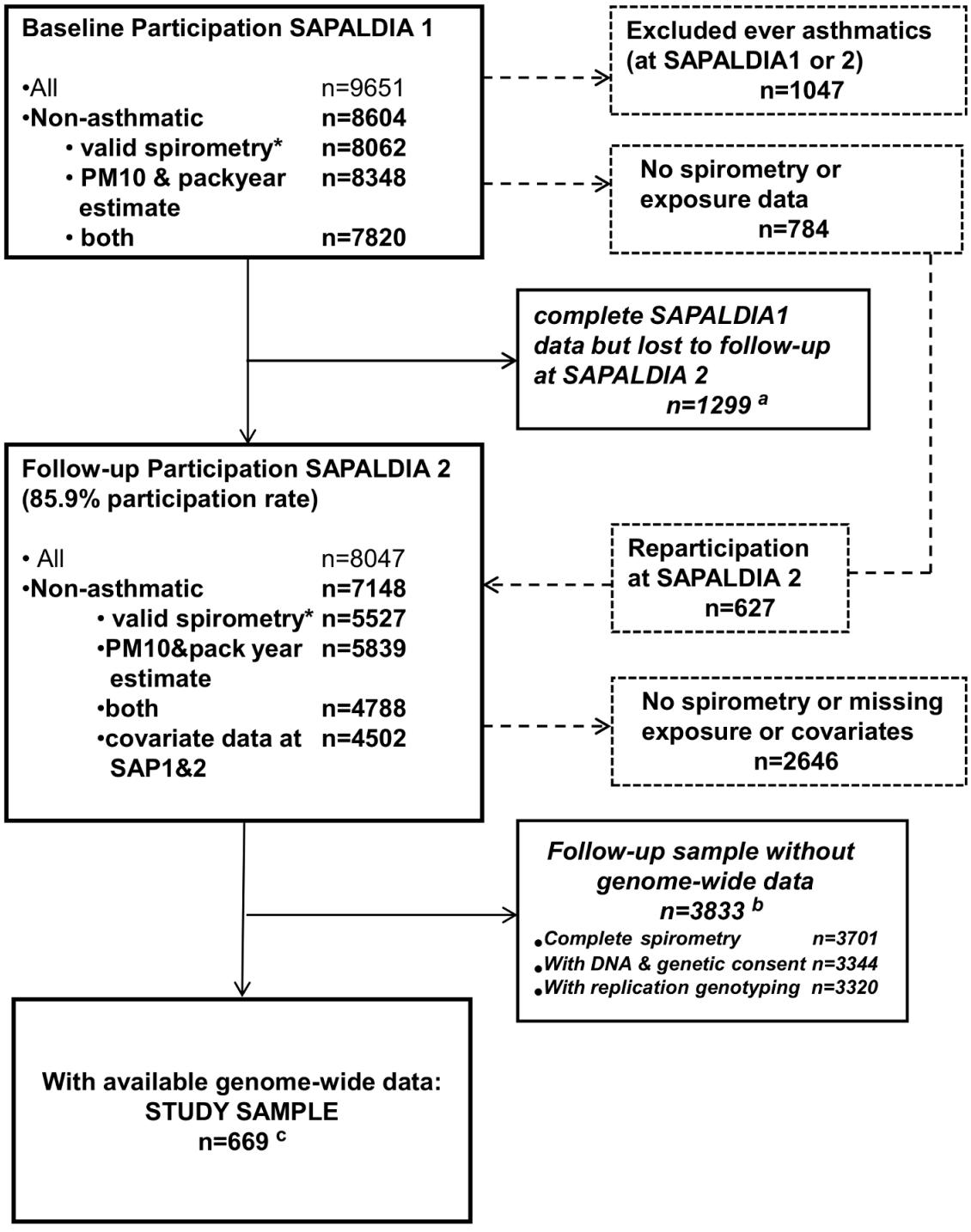
C) Apoptosis pathway genes & decline in FEV_1



P-values of interaction on the gene-level are given on a minus $\log_{10}$ scale (y-axis), i.e. higher bars represent smaller interaction p-values

A) Genes of the mitochondrial dysfunction pathway interacting with PM10 and packyears exposure between surveys on FEV_1/FVC decline.
B) Genes of the methane metabolism pathway interacting with PM10 and packyears exposure between surveys on FEF_{25-75} decline.
C) Genes of the apoptosis signaling pathway interacting with PM10 and packyears exposure between surveys on FEV_1 decline

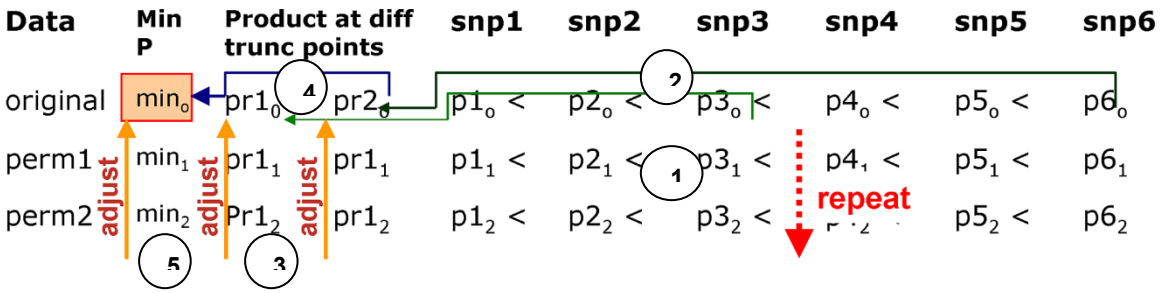
Figure S1 Follow-up of participants and selection of study population.



* Valid measurement of at least FEV_1

[a] Baseline sample, [b] Follow-up sample, and [c]Analysis sample in comparisons of supporting files, Table S2

Figure S2 Scheme of analysis steps in the ARTP-method.



The ARTP method developed by Yu and colleagues [52] assumes that an analysis at the SNP-level has been performed on the originally observed data, followed by a reanalysis on permutated datasets, i.e. p-values of association for original and permutated datasets are available for each SNP. The ARTP procedure then entails the following 4 steps:

1. **Order** p-values from single SNP analysis in ascending order
2. Calculate **products** of ranked p-values at different truncation points depending on gene length
3. **Adjust** product p-values using permutation distribution
4. **Select** the **minimum** of the adjusted products
5. **Readjust** the product minimum.

The **readjusted product minimum** represents the **gene-level p-value**. For each permutated dataset, an adjusted product minimum can be calculated as well. The procedure can then be repeated using the resulting, original and permutation gene-level p-values to yield p-values of the pathway.

7. Summary of findings

7.1. Paper 1: Longitudinal change of prebronchodilator spirometric obstruction and health outcomes: results from the SAPALDIA cohort

In this article, we estimated the incidence of airway obstruction over a follow-up of 11 years using a modified GOLD-definition of COPD, which applied the GOLD criteria to pre-bronchodilation spirometry values. We found an incidence of 14.2 cases/1000 person years (PY), which was at the higher end of estimates from comparable population-based cohorts ranging from 3-16 cases/1000 PY. The high incidence could only partly be explained by differences in age structure, smoking distribution or length of follow-up. Baseline characteristics identified as determinants of incidence were higher age and smoking exposure, and presence of chronic bronchitis. While positive associations with the first two factors were consistently described in previous studies on airway obstruction, the evidence for chronic bronchitis was less clear. Using FEV_1/FVC < lower limit of normal (LLN) to define obstruction, the incidence was 7.2 cases/1000 PY, and female sex emerged as additional risk factor, possibly indicating a higher susceptibility of females to develop obstruction. The positive linear association with age expectedly disappeared (LLN is dependent on age). We also observed that 20.9% of obstructive cases at baseline did not persist. Most of them showed mild obstruction severity at baseline.

When investigating how longitudinal change in severity of obstruction over 11 years was correlated with health service use for respiratory problems or shortness of breath while walking, we found that progression from mild (stage I) to moderate to severe obstruction (stage II and more), and persistence in moderate to severe stages was most strongly associated with more frequent health service use and dyspnea at follow-up. Importantly, the association was robust to adjusting for smoking exposure and pre-existing symptoms or health service use. Persisting in a state of mild obstruction was not associated with higher service use or symptoms, but reversing from an obstructive state at baseline to normal at follow-up was marginally associated with higher health service use, despite adjustment for presence of asthma or wheezing without a cold at either examination. Interestingly, lung function values for both FEV_1 and FVC were on average above the predicted values in this category, indicating more a mismatch between FEV_1 and FVC than abnormally low lung function. These observations could possibly be due to cases of hidden asthma.

The main limitation of this study was an observed selection of participants in better health condition during follow-up time. To assess the possible impact of these selection processes on associations with health service use and dyspnea at exercise, we reran the analysis giving more weight to participants with underrepresented characteristics in the sample. The results were robust.

We concluded that pre-bronchodilation spirometry as often available in epidemiological studies has prognostic value in predicting future adverse health events and health service use, although it might misclassify hidden asthma for COPD. The observation that mild obstruction stages were only associated with adverse events if they further progressed suggested the need for bronchodilation. Also, further clinical characteristics such as medication intake and symptoms for asthma, or functional scores such as the BODE index in the case of COPD could help correctly identifying participants at risk of progression in an early disease stage.

7.2. Paper 2: HMOX1 and GST variants modify attenuation of FEF25-75% decline due to PM10 reduction

We investigated whether variants in genes *heme-oxygenase 1* (*HMOX-1*) and the *glutathione S-tranferase* (*GST*) gene superfamily, which comprises several isoforms including the M, P, and T gene classes, modified the effect of a change in PM_{10} exposure between surveys on lung function decline.

The studied variants included one single nucleotide polymorphism (SNP) for genes *GSTP1*, three SNPs for gene *HMOX-1*, a well studied functionally relevant DNA repeat in the promoter region of *HMOX-1*, and gene deletions in genes *GSTM1* and *GSTT1*, which are highly prevalent in European populations. These variants were selected because of previously published associations with pulmonary function or respiratory disease, except for the three *HMOX-1* SNPs which were carefully chosen to maximally capture observed DNA variability for *HMOX-1* in European populations. Both *HMOX-1* and the *GST* genes belong to the lung's first line defense against oxidative stress, and were thus very good candidates for interaction effects.

The analysis model we used was strongly based on previously published work [69], which showed that a reduction in air pollution exposure was on average associated with attenuation in the natural, age-related lung function decline. The observed effects were thereby particularly strong in the smaller airways fraction, captured by the decline in the mid-flow measure FEF_{25-75}.

In our analysis, we observed that variants in *GSTP1* and *HMOX-1*, significantly modified the effect of a reduction in PM_{10} exposure. Participants having mutant alleles in the *HMOX-1* SNPs, long repeat alleles in the *HMOX-1* promoter region, or the mutant alleles in *GSTP1* profited most from the air quality improvement. In analogy to the previous work, strongest interaction effects were seen for decline in FEF_{25-75}, and those for FEV_1 and FVC were weaker and sometimes inconsistent compared to FEF_{25-75}.

The conclusion for this analysis was that variation in genes involved in the body's defense against oxidative stress modifies the beneficial effects of a reduction in air pollution. Individuals with a differing endogenous capacity to cope with oxidative stress profited to a different extent from the improvement in air quality. Once the molecular mechanisms are better understood, this finding potentially has policy implications for the definition of exposure limits, as those should protect the weakest and most susceptible members of society.

7.3. Paper 3: Different genes interact with particulate matter and tobacco smoke exposure in affecting lung function decline in the general population.

In the third paper, we assessed whether SNPs in a broad set of pathways and genes involved in oxidative stress defense or its endogenous production interact with exposure to air pollution on lung function decline, and compared the impact of such interactions to those observed with tobacco smoke exposure. As tobacco smoke exposure is measured using a cumulative measure of pack years, cumulative PM_{10} exposure estimates were used in this analysis. No analysis of interaction between oxidative stress genes and air pollution on respiratory outcomes in such a broad scale was published so far.

The study was based on a subset of 669 non-asthmatic adults with available covariate information and genome-wide data that would cover the genes and pathways of interest. Genes and pathways had been identified by searches in functional gene databases and by feeding the resulting gene list into a protein interaction database. The analysis comprised 152 genes, 14 pathways, and 12679 SNPs.

Pathway analysis was applied using the ARTP-method published by Yu and colleagues[155] to maximally increase statistical power, but also to obtain interaction p-values on the higher biological levels of genes and pathways.

After correcting for multiple testing, we found that genes *CRISP2* significantly, and *SNCA* marginally interacted with PM_{10} on the decline in FEV_1/FVC. In comparison, a pure SNP-level analysis yielded one SNP in gene *PARK2* significantly interacting with PM_{10} on FEV_1-decline, besides the already observed interaction in *SNCA*. We attempted to replicate the interactions with *CRISP2* and *PARK2* in the remainder of the SAPALDIA population (without genome-wide data), but without success (the SNP in SNCA was not chosen for replication due its low MAF of 7%). When looking at the interaction pattern in an exploratory manner, i.e. inspecting nominally significant genes ($p_{interaction}<0.05$) across outcomes and exposures, we observed that different genes and pathways were interacting with PM_{10} and tobacco smoke exposure. Based on the interaction effect estimates for the strongest SNP of each nominally significant gene, the impact of tobacco smoke exposure on lung function decline was larger than that of PM_{10}. Also, besides significant interaction, tobacco smoke exposure often presented a significant main effect per se. On the other side, the percent variability in lung function decline explained was very similar between the exposures, ranging up to 28.5% for PM_{10} and 26.1% for tobacco smoke, and to 33.3% for both. However, this finding was based on adjusted R square values from models including all strongest interacting SNPs from each nominally interacting gene simultaneously. According to the winners curse phenomenon[156], these percentages are likely to be overestimated given the small size of our study sample and the applied filtering for strongest interaction signals. Further, they are questioned by the non-replication of the significant interaction signals in *CRISP2* and *PARK2*. As a consequence, estimated percentages of explained variability have not been included into the manuscript, but they are presented as supplemental information in Appendix 4 of the present thesis only.

In conclusion, we were not able to detect consistent, significant interactions using either SNP-level or gene-level analysis after strict correction for multiple testing. However, by evaluating nominally

significant interactions, we observed suggestive evidence that different genes and pathways are involved in handling oxidative stress from PM_{10} and tobacco smoke. This could possibly be due to the different levels of exposure imposed by PM_{10} and tobacco smoke, or to a lesser extent also by differences in their composition. Such effects would be missed in a gene main effect analysis, as it is most often done in genetic studies including GWAS today.

8. Discussion

8.1. General discussion

The goals of this thesis were: I) to estimate the incidence of COPD in the Swiss general population and identify its main determinants, II) to investigate whether ambient air pollution contributes to COPD development, and III) to compare the possible impact of ambient air pollution on COPD development to that of the most important known risk factor, tobacco smoke exposure.

Assessing the importance of environmental effects thereby focused on the role of oxidative stress, which is thought to be a major etiological pathway in COPD, and whose role was presently investigated by studying gene environment interactions. The scientific work of this thesis was based on available data from the population-based SAPALDIA cohort study.

In the first study we estimated the incidence of airway obstruction in SAPALDIA primarily using the COPD definition proposed by the GOLD initiative, but with pre-instead of post-bronchodilation spirometry. We assessed the prognostic value of longitudinal changes in obstruction severity regarding future adverse health events and health service use. Our incidence estimates are at the higher range of comparable values from other population based studies, the reasons for which cannot be clearly established. Estimates diminished considerably when using the lower limit of normal (LLN) of the FEV_1/FVC ratio to define disease. This decline is expected given the high proportion of participants aged 50 years and over, for which the GOLD criteria are known to lead to substantial over-diagnosis. In this respect it is also important to consider the accumulating evidence that COPD is not just a smoking related disease[1,61], and hence comparing age-structure and smoking burden across studies might miss other important determinants. Analogously to other studies, we identified age, and smoking as determinants of incidence of obstruction, and found additional evidence for a possible role of chronic bronchitis and sex (the latter when using the LLN criterion). We observed that the transition to or persistence in more severe stages during follow-up was correlated with higher health service use and more dyspnea at exercise. This suggests that repeated pre-bronchodilation spirometry still has value in epidemiological research to predict health outcomes. At the same time, our observation that 20% of our obstructive cases at baseline did not persist indicates that pre-bronchodilation spirometry is misclassifying asthma for COPD. This is also supported by our observations that the category of non-persistent obstruction on average presented normal lung function values, but was marginally associated with later health service use for respiratory problems, which might be due to presence of hidden asthma. Our results thus support the use of bronchodilation in lung function testing to reduce the potentially large misclassification with hidden asthma in studies of COPD. Accordingly, bronchodilation was carried out in the second follow-up examination of SAPALDIA taking place in 2010/2011. The observation that mild obstruction is only associated with later adverse outcomes if it didn't persist or increased suggests that instead of basing COPD severity classification solely on spirometry, further characteristics might be useful to define who is at risk for progression among mildly obstructive cases. A possible example could be the multidimensional BODE index, which is used in clinical COPD management today and comprises spirometry, body mass index, exercise capacity, and grade of dyspnea.

Consequently, in the ensuing studies we focused on lung function decline as an intermediate phenotype and proxy measure of COPD. This decision was also influenced by the fact that statistical power had to be maximized to assess interaction effects, which are more demanding in terms of sample size requirements than main effect studies.

In the second study, we investigated whether variation in two genes belonging to the lung's first line of defense against oxidative stress could alter the beneficial effects of reducing ambient PM_{10} exposure on the natural, age-related lung function decline reported previously[69]. We could show that a part of the investigated genetic variants modified the response to air quality improvement, which implicated that beneficial effects of a reduction in air pollution are not equally distributed across the population, but are influenced by the individual capacity to cope with oxidative stress. Similar to the paper upon which the analysis was based[69], we found the strongest interaction effects in the mid-flow measure FEF_{25-75}, but also weaker effects on FEV_1. This is probably due to the size-specific deposition profile of particulate matter. The fine, biologically more active $PM_{2.5}$ portion[157,158] of PM_{10} is preferentially retained in the small airways whose patency is approximated by FEF_{25-75}. These results were robust to adjusting for the FVC in different ways (the FEF_{25-75} parameter is known to be strongly dependent on the measured FVC). The interaction effects on FEV_1 were mostly consistent with those in FEF_{25-75}, but did not reach statistical significance. The evidence for effects on FEV_1 was thus suggestive.

The observed interactions are important from a Public Health perspective as they potentially have implications for the setting of air pollution limits. Clean air legislation should aim to protect the weakest members of society[159]. However, a limitation of the study was that functional alterations on the enzyme level could not be defined for most variants with which interactions were observed. Likewise, no data on functional changes was available for our studied *heme oxygenase 1* (*HMOX1*) SNPs, as they were not selected on the basis of known functional deficits, but because they tag a large part of the common variability in the *HMOX1* gene. The chosen Ile105Val polymorphism in GSTP1 is well studied and known to alter the detoxification rate of diolepoxides[160]. However, the direction of allele-specific effects was dependent on the studied outcome and exposure in previous epidemiological studies[102]. It is therefore very difficult to make inferences about underlying molecular mechanisms in the case of complex exposure mixtures such as PM_{10}. As a consequence, it was not possible to make a strong case about the Public Health importance of our findings by relating them to the underlying biology.

The observed interactions between ambient air pollution and genetic variation in enzymes belonging to the first line of defense from oxidative stress, encouraged us to assess the impact of PM_{10} exposure on lung function decline in a broad set of oxidative stress related genes and pathways in the third paper. Effects were also compared to those of tobacco smoking, the classical COPD risk factor. The broad set of candidate genes and pathways necessitated focusing on the subset of SAPALDIA participants with available genome-wide data, at the cost of sample size and hence statistical power. To improve power and maximally exploit the available data, we reduced the complexity of our analysis models and applied modern, pathway or gene set analysis methods. These reduce the dimensionality of the data by making use of pre-existing biological knowledge and integrate the lower level SNP-association signals onto the upper biological levels of genes and pathways[161]. The incorporation of biological knowledge can also eventually help to interpret the results. The analytical strength of the methods is that they combine multiple moderate to weak association signals from SNPs belonging to the same gene or pathway, which can yield significant

associations overall. Several statistical methods are thereby available to derive the overall signal across multiple SNPs, and we used the one published by Yu and colleagues[155]. As a comparison, we also performed a classic SNP-level analysis with correction for multiple testing. We detected one significant interaction with PM_{10} exposure on FEV_1/FVC decline at the gene (*CRISP2*) and one on FEV_1 decline at the SNP-level (SNP rs12190800 in gene *PARK2*), and found suggestive evidence that the effects of PM_{10} exposure and tobacco smoking were mediated through different genes and pathways with only a minimal overlap. Importantly, interaction effect sizes conveyed by tobacco smoking were roughly twice as large as those from PM_{10} (except for decline in FEV_1/FVC where comparable effects were observed), indicating a higher effect on lung function decline in smoking individuals. On the other side, first estimates of the percent explained variability in lung function decline were very similar for both exposures, and although these estimates were questioned by methodological issues, they nevertheless suggest that the impact of PM_{10} exposure on lung function decline at the population level might be comparable to tobacco smoke. This underlines the importance of Public Health efforts to establish a clean air environment, and makes particulate matter pollution an important target for preventive action.

From a methodological point of view, our analysis was the first to assess interactions between a very broad set of oxidative stress genes and environmental risk factors via pathway analysis methods in respiratory epidemiology. We could not overcome the negative impact of a restricted sample size, but the work showed the practical feasibility and promising first results of applying modern analysis methods to assess the role of entire genes and pathways in mediating environmental effects.

8.2. Limitations

Our studies faced several limitations. First, the lack of post-bronchodilation spirometry hampered our ability to calculate a robust estimate for the burden of COPD in Switzerland. The estimated high incidence rates of airway obstruction might thus be an overestimation due to undiagnosed asthma cases. Our observations related to the cases of non-persistent obstruction would support such an explanation. We could however show the value of available pre-bronchodilation measurements regarding prediction of future health events. We thereby controlled for misclassification by hidden asthma as far as possible. Post-bronchodilation lung function testing has been applied at the second SAPALDIA follow-up examination (SAPALDIA 3), and as soon as these data will be available for analysis, they will enable us to shed light on the question on how much the observed incidence rate was overestimated by reversible airway obstruction.

Our observations during the assessment of the COPD burden suggested focusing on lung function decline for further analyses. This was also beneficial to increase statistical power for studying gene-environment interactions. The main limitation in this respect was that we had to base our estimates of lung function decline on only two measurements 11 years apart, which introduces a certain degree of misclassification by measurement error. But we did not expect the degree of measurement error to be influenced by the genetic background of the participant or the level of ambient air pollution. Thus errors in lung function measurement would primarily induce non-differential misclassification, reduce statistical power for our analyses, and possibly lead to an underestimation of the true air pollution effects.

A limitation of our genetic data is that for many of the investigated genetic variants the functional impact on the enzyme level is not known. This makes it often impossible to give an explanation of the molecular mechanisms underlying observed genetic main effects or interactions with environmental exposures. This is particularly true for genome-wide association studies, which were originally conceived as hypothesis-free approaches to uncover new genes and pathways in disease causation. The limitation is also due to the technical advances that genotyping procedures have undergone in the last two decades, which enable fast and large-scale measurement of DNA-variation while functional studies involving in vitro and in vivo experiments still require more time. It can thus be expected, that the molecular basis of many consistently observed genetic associations and interactions today will be uncovered in the next years.

An additional limitation comes to effect in our last analysis involving genome wide data. The employed genotyping platform (Illumina Human 6010quad BeadChip) was designed to cover known, commonly occurring genetic variation in the human genome as much as possible. It might thus not optimally cover all the genes and pathways that were of interest for our analysis. However, our measured and imputed genome wide data was based on the HapMap project[154], which represents the current state of knowledge on human genetic variation and serves as reference standard for analysis. Consecutive releases of the HapMap data progressively increased the resolution of the genetic variation map, and hence it is likely that the coverage of our genes of interest will be improved in the future.

Finally, a limitation observed in all of our studies is that participants at the follow-up examination were on average in better health than those who participated only at baseline. We regularly took account of this selection process by comparing our analysis sample with participants who took part only in the baseline assessment. If this comparison indicated a potential for bias in the results and the chosen analysis technique allowed weighing, we applied a sensitivity analysis giving more weight to underrepresented participants to increase their influence on the analysis results. Sensitivity analysis did not alter our results in these cases. Otherwise, the descriptive comparison of study and baseline sample served as a basis to think about the expected direction and range of biases. Based on the known biological effects of air pollution and previous literature showing increased susceptibility of persons with pre-existing chronic diseases[122,123,139], we concluded that the selection process taking place during follow-up would more likely lead to an underestimation of the true health effects of air pollution than inducing spurious associations.

8.3. Strengths

One of the strengths of our studies is SAPALDIA's population-based design, which allows investigating the research questions in a Swiss general population sample of adults with very wide age range. This offers the possibility to investigate the whole severity spectrum of COPD disease including the early, pre-clinical stages, which can yield important insights for the development of preventive measures. In this respect, focusing a part of the analyses on lung function decline and its acceleration could capture potentially important risk factors in the general population, which could ultimately contribute to later development of COPD.

A further excellent characteristic of the SAPALDIA study from which our work took benefit is the availability of high quality and validated estimates for exposure to particulate matter air pollution (PM_{10}) on the individual level. Moreover, the PM_{10} estimates covered the whole range of air pollution exposure throughout Switzerland thanks to the careful selection of study areas in the original design phase. In this respect, the available air pollution data in SAPALDIA offers an almost unique opportunity to investigate the health effects of air pollution on the population level.

Third, despite the limits imposed by missing bronchodilation, the SAPALDIA study disposes of high quality spirometry data obtained from standardized lung function testing according to international quality standards. Further, spirometry devices were regularly checked and calibrated and their comparability was assessed before each assessment. These rigorous quality control measures ensured to keep measurement error as small as possible and maximize the data quality in the two examinations. This also benefited the analyses related to this thesis.

Our genetic data offered the strengths of a large sample size of over 6000 samples, on which robust candidate genes studies could be carried out. Further, the selection of candidate SNPs for genotyping within study projects was guided by published results from respiratory and cardiovascular genetic epidemiology studies. This yielded very good candidate SNPs for the investigation of oxidative-stress related health effects of air pollution, such as those we successfully studied in the second paper.

Finally, a further strength of our work was the application of modern analysis techniques to uncover significant interactions of ambient air pollution with a broad set of oxidative stress genes and pathways in the third study, although its return was reduced by the limited number of non-asthmatic participants with genome-wide data. The study has nevertheless shown that the approach is feasible, shows promising results, and will likely yield important insights into disease-related molecular pathways with a larger data basis.

8.4.Conclusions

From a Public Health perspective, our study results have potential implications with regard to the burden of COPD and associated risk factors in Switzerland.

First, according to our estimates the incidence of COPD in Switzerland is at the higher end in comparison with data from other countries, with the reservation that our results are based on pre-bronchodilation spirometry. But misclassified asthma cases appeared to impose additional burden on the health system as well, insofar the overestimated portion of COPD-incidence is also reflected in terms of its consequences on individual health and health system resources. The burden of airway obstruction thus warrants further investigation and monitoring at the population level, but with more accurate methods including bronchodilation to enable a robust separation of the burden imposed by COPD opposed to asthma.

Second, although effects of an improvement in air pollution can be measured in terms of a change in the normal-age related lung function decline at the population level, the benefits from cleaner air are not equally distributed among all members of society, but are influenced by the individual genetic profile determining the ability to cope with oxidative stress. Thus current estimates of the projected

benefits from reductions in air pollution[69,162,163] are on average correct at a population level, but individual benefits might be much larger or smaller depending on individual susceptibility. Lack of functional data on the molecular level currently precludes defining robust genetic high- or low-risk profiles in this respect.

Third and last, although a small sample size hampered a statistical proof, results from our pathway interaction analyses suggest that the contribution of ambient air pollution to lung function decline and probably also COPD is mediated by oxidative stress related pathways, which potentially differ from those involved in tobacco smoke effects. Genetic variations within these pathways might constitute individual susceptibility. Moreover, our tentative estimates of the percentage of explained variability in lung function decline suggest the share of air pollution in the burden of COPD on the population level might be equal to that of tobacco smoke. This is of high Public Health importance as air pollution is omnipresent in densely populated areas, and hence exposure is involuntary and not avoidable. According to the well known "prevention paradox" present in programs targeting the whole population[164], small attenuations in average lung function decline of the population could lead to a substantial reduction of disease cases at the higher end of the distribution. In this respect, a reassessment of our results by further studies with larger sample sizes is clearly warranted. Meanwhile, the role of air pollution in respiratory disease at the population level needs reconsidering and deserves greater attention in Public Health policy.

8.5. Outlook

The work in this thesis has yielded evidence that COPD-related respiratory disease constitutes an important Public Health problem in Switzerland and that ambient air pollution, as represented by exposure to PM_{10} in our study, might be an important contributor to this burden. Moreover, the results of our studies underline the important role of oxidative stress in shaping the velocity of lung function decline and mediating the effects of ambient air pollution and tobacco smoking on respiratory health.

A future identification of key enzymes and pathways defining an individual's susceptibility to oxidative stress would be an important milestone for Public Health, as it would allow defining high-risk subgroups within the population. Administration of antioxidant substances would naturally lend itself as preventive measure to counteract high oxidative stress levels. A Mexican study on genetic susceptibility to air pollution in children provided first evidence for effectiveness in this respect[165]. Given that it would be comparably easy to implement by enriching the daily diet with fruits and vegetables or possibly by dietary supplements, running campaigns for a healthy diet in periods of high air pollution might be worth considering on the population level or targeting children, the elderly or persons with cardiovascular disease. Tailoring these measures to high-risk individuals would increase efficacy, but requires the previous definition of a detailed individual risk profile.

As molecular pathways in our body are highly complex networks whose parts can interact but also compensate for each other, their characterization will require large efforts and resources. In this context, studying the effects of variants in the genome might be advantageous, as they are resistant to environmental exposure effects and hence do not underlie reverse causality. Thus, while genetic variants rarely constitute major determinants of complex chronic diseases, they can still pinpoint

underlying pathways which might substantially contribute to disease causation via environmental exposures[166,167]. Moreover, identification of important pathways opens the potential for effective prevention even before the underlying biology is completely understood and detailed risk profiles can be defined. The current scientific evidence supports an important role of oxidative stress reactions in COPD causation, and it is therefore crucial to continue studying the respective genes and pathways to boost developments of preventive measures and therapeutic applications.

The studies in this thesis have also uncovered several limitations. These impede drawing firm conclusions regarding the true disease burden and extent of excess risk conferred by air pollution, which would provide the required scientific foundation on which to base Public Health action.

The discovered limitations clearly show the need for a large population based cohort with detailed characterization of phenotypes, comprising post-bronchodilation spirometry in the case of respiratory disease, but also including other important features such as respiratory symptoms, medication, extent of disability in daily life, as well as data on hospitalizations and treatment, which would help to measure the disease-related burden on the level of individuals but also health systems.

A more detailed characterization of the disease phenotype would also help to identify etiological factors, as different subtypes of disease might arise on different risk factors. For the same purpose, such a cohort would also require a large biobank comprising serum and DNA samples of each participant. Serum measurements would allow to identify important biomarkers as measures of biological exposure doses or early, intermediate stages of disease. Genetic information, preferentially in the form of genome-wide data would offer the possibility to identify molecular pathways involved in disease causation and ultimately, when the main molecular processes on which a disease arises are identified, also to determine individual susceptibility.

In this respect, it is important to realize that most population based genetic epidemiological studies investigate associations between disease phenotypes and variation in the deoxyribonucleic acid (DNA), thus inherited susceptibility. But environmental exposures as well as the normal aging process also alter the expression pattern of genomic DNA, i.e. the level of gene and consequently enzyme activity in target tissues. It is likely that such processes are very important for disease causation in tissues directly exposed to the outer environment such as the lungs. And while technologies to measure the activity of genes at a genome-wide scale have been developed and used on white blood cell DNA[168], it will in future also be mandatory to collect direct samples of the studied target tissue from participants of large epidemiological studies. Because tissue sampling is normally not possible without imposing health risks to participants, close collaboration with clinical institutions to gain samples from diagnostic or treatment procedures is necessary.

The establishment of such a large-scale cohort with detailed phenotype and genotype data including tissue samples thus requires considerable investments in terms of long-term interdisciplinary collaboration and dedicated resources, not to forget careful preparation to comply with ethical requirements and required broad informed consent of participants when bringing disease- and treatment-related data into epidemiological studies.

From a scientific perspective, such an investment would be clearly worth doing, as it would also offer the possibility to investigate the etiology of other chronic diseases, which are expected to increase in prevalence due to population aging. This can probably not be achieved by a loose collaboration of existing Swiss cohort studies on a national level, as they were designed with different goals in mind,

and thus dispose of different data collections from differing populations. The establishment of a new, national large scale cohort instead necessitates close and multidisciplinary collaborations from the side of Public Health experts on the national level with maximal exploitation of preexisting expertise.

In this respect, the SAPALDIA study can contribute its expertise in international collaborations in the field of genetic epidemiology, and particularly in genome-wide association studies in respiratory but also other chronic diseases, which has been initiated by the availability of genome-wide data through the GABRIEL project. These collaborations have led to a buildup of expertise in handling and analyzing large scale genomic data, including gene-environment interaction in chronic disease. In this context, it's long standing and profound experience with biobanking is a further important asset. Of equal importance, substantial expertise has been built up in the assessment of respiratory and cardiovascular phenotypes as well as measurement of air pollution exposure in course of the SAPALDIA examinations.

The future establishment of a national, large scale cohort aiming to uncover susceptibility to chronic, non-communicable disease by concerted and dedicated action from the Swiss Public Health community will likely enable us to determine to which extent air pollution contributes to COPD development, thus providing urgently needed answers on the role of air pollution in COPD causation, expected benefits from further air quality improvements, and the best preventive strategy with respect to the distribution of susceptibility in the population.

References

1. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease (updated 2010). (Accessed September 27th, 2011, at http://www.goldcopd.org/uploads/users/files/GOLDReport_April112011.pdf.)
2. Pauwels RA, Buist AS, Calverley PM, et al. 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;163:1256-76.
3. Celli BR, MacNee W, Agusti A, et al. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. European Respiratory Journal 2004;23:932-46.
4. Pauwels RA, Rabe KF. Burden and clinical features of chronic obstructive pulmonary disease (COPD). Lancet 2004;364:613-20.
5. World Health Organization. Global status report on noncommunicable diseases 2010. Geneva: World Health Organization; 2011.
6. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. The Lancet 1997;349:1498-504.
7. Sullivan SD, Ramsey SD, Lee TA. The economic burden of COPD. Chest 2000;117:5S-9S.
8. Halbert RJ, Natoli JL, Gano A, et al. Global burden of COPD: systematic review and meta-analysis. European Respiratory Journal 2006;28:523-32.
9. Buist AS, McBurnie MA, Vollmer WM, et al. International variation in the prevalence of COPD (The BOLD Study): a population-based prevalence study. The Lancet 2007;370:741-50.
10. Menezes AMB, Perez-Padilla R, Jardim JB, et al. Chronic obstructive pulmonary disease in five Latin American cities (the PLATINO study): a prevalence study. The Lancet 2005;366:1875-81.
11. Atsou K, Chouaid C, Hejblum G. Variability of the chronic obstructive pulmonary disease key epidemiological data in Europe: systematic review. BMC Medicine 2011;9:7. epub ahead of print, doi: 10.1186/741-7015-9-7.
12. De Marco R AS, Cerveri I, Corsico A, Anto JM. Incidence of chronic obstructive pulmonary disease in a cohort of young adults according to the presence of chronic cough and phlegm. American Journal of Respiratory and Critical Care Medicine 2007:32-9.
13. Lindberg A, Eriksson B, Larsson LG, et al. Seven-year cumulative incidence of COPD in an age-stratified general population sample. Chest 2006;129:879-85.
14. Lokke A, Lange P, Scharling H, et al. Developing COPD: a 25 year follow up study of the general population. Thorax 2006;61:935-9.
15. Johannessen A, Omenaas ER, Bakke PS, et al. Implications of reversibility testing on prevalence and risk factors for chronic obstructive pulmonary disease: a community study. Thorax 2005;60:842-7.
16. Ackermann-Liebrich U, Kuna-Dibbert B, Probst-Hensch NM, et al. Follow-up of the Swiss Cohort Study on Air Pollution and Lung Diseases in Adults (SAPALDIA 2) 1991-2003: methods and characterization of participants. Sozial- und Präventivmedizin 2005;50:245-63.
17. Martin BW, Ackermann-Liebrich U, Leuenberger P, et al. SAPALDIA: methods and participation in the cross-sectional part of the Swiss Study on Air Pollution and Lung Diseases in Adults. Sozial- und Präventivmedizin 1997;42:67-84.
18. Bridevaux PO, Probst-Hensch NM, Schindler C, et al. Prevalence of airflow obstruction in smokers and never-smokers in Switzerland. European Respiratory Journal 2010;36:1259-69.
19. Bridevaux PO, Gerbase MW, Probst-Hensch NM, et al. Long-term decline in lung function, utilisation of care and quality of life in modified GOLD stage 1 COPD. Thorax 2008;63:768-74.
20. Mannino DM, Braman S. The Epidemiology and Economics of Chronic Obstructive Pulmonary Disease. Proceedings of the American Thoracic Society 2007;4:502-6.

21. Han MK. Update in chronic obstructive pulmonary disease in 2010. American Journal of Respiratory and Critical Care Medicine 2011;183:1311-5.
22. Fletcher C, Peto R. The natural history of chronic airflow obstruction. British Medical Journal 1977;1:1645-8.
23. Bednarek M, Maciejewski J, Wozniak M, et al. Prevalence, severity and underdiagnosis of COPD in the primary care setting. Thorax 2008;63:402-7.
24. Soriano JB, Zielinski J, Price D. Screening for and early detection of chronic obstructive pulmonary disease. The Lancet 2009;374:721-32.
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;63:1046-51.
26. Celli BR, Halbert RJ, Isonaka S, et al. Population impact of different definitions of airway obstruction. European Respiratory Journal 2003;22:268-73.
27. 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 2007;176:532-55.
28. Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. European Respiratory Journal 2005;26:948-68.
29. Quanjer PH, Tammeling GJ, Cotes JE, et al. 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. European Respiratory Journal Suppl 1993;16:5-40.
30. Quanjer PH. Standardized lung function testing. Bulletin Europeen de Physiopathologie Respiratoire 1983;19:1-95.
31. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. American Journal of Respiratory and Critical Care Medicine 1999;159:179-87.
32. Kuster SP, Kuster D, Schindler C, et al. Reference equations for lung function screening of healthy never-smoking adults aged 18-80 years. The European Respiratory Journal : official journal of the European Society for Clinical Respiratory Physiology 2008;31:860-8.
33. Langhammer A, Johnsen R, Gulsvik A, et al. Forced spirometry reference values for Norwegian adults: the Bronchial Obstruction in Nord-Trondelag Study. European Respiratory Journal 2001;18:770-9.
34. Roca J, Burgos F, Sunyer J, et al. References values for forced spirometry. Group of the European Community Respiratory Health Survey. European Respiratory Journal 1998;11:1354-62.
35. Pistelli F, Bottai M, Carrozzi L, et al. Reference equations for spirometry from a general population sample in central Italy. Respiratory Medicine 2007;101:814-25.
36. Brandli O, Schindler C, Kunzli N, et al. Lung function in healthy never smoking adults: reference values and lower limits of normal of a Swiss population. Thorax 1996;51:277-83.
37. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. European Respiratory Journal 2005;26:319-38.
38. Mannino DM. Defining chronic obstructive pulmonary disease… and the elephant in the room. European Respiratory Journal 2007;30:189-90.
39. Miller MR, Pedersen OF, Pellegrino R, et al. Debating the definition of airflow obstruction: time to move on? European Respiratory Journal 2009;34:527-8.
40. Quanjer PH, Enright PL, Miller MR, et al. The need to change the method for defining mild airway obstruction. European Respiratory Journal 2011;37:720-2.
41. Medbø A, Melbye H. Lung function testing in the elderly--Can we still use FEV1/FVC<70% as a criterion of COPD? Respiratory Medicine 2007;101:1097-105.
42. Hardie JA, Buist AS, Vollmer WM, et al. Risk of over-diagnosis of COPD in asymptomatic elderly never-smokers. European Respiratory Journal 2002;20:1117-22.

43. Hnizdo E, Glindmeyer HW, Petsonk EL, et al. Case Definitions for Chronic Obstructive Pulmonary Disease. Journal of Chronic Obstructive Pulmonary Disease 2006;3:95-100.
44. Roberts SD, Farber MO, Knox KS, et al. FEV1/FVC Ratio of 70% Misclassifies Patients With Obstruction at the Extremes of Age. Chest 2006;130:200-6.
45. Brandli O, Schindler C, Leuenberger PH, et al. Re-estimated equations for 5th percentiles of lung function variables. Thorax 2000;55:173-4.
46. Huang Z, Willett WC, Colditz GA, et al. Waist circumference, waist:hip ratio, and risk of breast cancer in the Nurses' Health Study. American Journal of Epidemiology 1999;150:1316-24.
47. Shirtcliffe P, Weatherall M, Marsh S, et al. COPD prevalence in a random population survey: a matter of definition. European Respiratory Journal 2007;30:232-9.
48. Vollmer WM, Gíslason, Burney P, et al. Comparison of spirometry criteria for the diagnosis of COPD: results from the BOLD study. European Respiratory Journal 2009;34:588-97.
49. Mannino DM, Sonia Buist A, Vollmer WM. Chronic obstructive pulmonary disease in the older adult: what defines abnormal lung function? Thorax 2007;62:237-41.
50. Cerveri I, Corsico AG, Accordini S, et al. Underestimation of airflow obstruction among young adults using FEV1/FVC <70% as a fixed cut-off: a longitudinal evaluation of clinical and functional outcomes. Thorax 2008;63:1040-5.
51. de Marco R, Accordini S, Anto JM, et al. Long-term Outcomes in Mild/Moderate COPD in the European Community Respiratory Health Survey. American Journal of Respiratory and Critical Care Medicine 2009;180:956-63.
52. Enright PL. GOLD stage I is not a COPD risk factor. Thorax 2007;62:1107.
53. García-Rio F, Soriano JB, Miravitlles M, et al. Overdiagnosing Subjects With COPD Using the 0.7 Fixed Ratio: Correlation With a Poor Health-Related Quality of Life. Chest 2011;139:1072-80.
54. Mannino DM, Diaz-Guzman E. Interpreting lung function data using 80% predicted and fixed thresholds identifies patients at increased risk of mortality. Chest 2011:epub ahead of print, doi: 10.1378/chest.11-0797.
55. Celli BR, Thomas NE, Anderson JA, et al. Effect of pharmacotherapy on rate of decline of lung function in chronic obstructive pulmonary disease: results from the TORCH study. American Journal of Respiratory and Critical Care Medicine 2008;178:332-8.
56. Troosters T, Celli B, Lystig T, et al. Tiotropium as a first maintenance drug in COPD: secondary analysis of the UPLIFT® trial. European Respiratory Journal 2010;36:65-73.
57. Lapperre TS, Snoeck-Stroband JB, Gosman MME, 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;151:517-27.
58. Decramer M, Cooper CB. Treatment of COPD: the sooner the better? Thorax 2010;65:837-41.
59. Miravitlles M, Anzueto A. Insights into interventions in managing COPD patients: lessons from the TORCH and UPLIFT® studies. International Journal of Chronic Obstructive Pulmonary Disease 2009;4:185-201.
60. Telenga ED, Kerstjens HAM, Postma DS, et al. Inhaled corticosteroids in chronic obstructive pulmonary disease: a review. Expert Opinion on Pharmacotherapy 2010;11:405-21.
61. Salvi SS, Barnes PJ. Chronic obstructive pulmonary disease in non-smokers. The Lancet 2009;374:733-43.
62. Eisner MD, Anthonisen N, Coultas D, et al. An Official American Thoracic Society Public Policy Statement: Novel Risk Factors and the Global Burden of Chronic Obstructive Pulmonary Disease. American Journal of Respiratory and Critical Care Medicine 2010;182:693-718.
63. Trupin L, Earnest G, San Pedro M, et al. The occupational burden of chronic obstructive pulmonary disease. European Respiratory Journal 2003;22:462-9.
64. 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;60:645-51.
65. Hnizdo E, Sullivan PA, Bang KM, et al. Association between chronic obstructive pulmonary disease and employment by industry and occupation in the US population: a study of data

from the Third National Health and Nutrition Examination Survey. American Journal of Epidemiology 2002;156:738-46.
66. Eisner MD, Balmes J, Katz PP, et al. Lifetime environmental tobacco smoke exposure and the risk of chronic obstructive pulmonary disease. Environmental Health 2005;4:7.
67. The Health Consequences of Involuntary Exposure to Tobacco Smoke: A Report of the Surgeon General. Washington, DC; 2006.
68. Leuenberger P, Schwartz J, Ackermann-Liebrich U, et al. Passive smoking exposure in adults and chronic respiratory symptoms (SAPALDIA Study). Swiss Study on Air Pollution and Lung Diseases in Adults, SAPALDIA Team. American Journal of Respiratory and Critical Care Medicine 1994;150:1222-8.
69. Downs SH, Schindler C, Liu LJ, et al. Reduced exposure to PM10 and attenuated age-related decline in lung function. New England Journal of Medicine 2007;357:2338-47.
70. Schikowski T, Sugiri D, Ranft U, et al. Long-term air pollution exposure and living close to busy roads are associated with COPD in women. Respiratory Research 2005;6:152.
71. Hu G, Zhou Y, Tian J, et al. Risk of COPD from exposure to biomass smoke: a metaanalysis. Chest 2010;138:20-31.
72. Viegi G, Maio S, Pistelli F, et al. Epidemiology of chronic obstructive pulmonary disease: health effects of air pollution. Respirology 2006;11:523-32.
73. World Health Organization. Occupational and Environmental Health Team. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide : global update 2005 : summary of risk assessment. Geneva: World Health Organization; 2006.
74. Witte JS, Hoffmann TJ. Polygenic Modeling of Genome-Wide Association Studies: An Application to Prostate and Breast Cancer. OMICS: A Journal of Integrative Biology 2011;15:393-8.
75. Molfino NA, Coyle AJ. Gene-environment interactions in chronic obstructive pulmonary disease. International Journal of Chronic Obstructive Pulmonary Disease 2008;3:491-7.
76. Senn O, Russi EW, Imboden M, et al. alpha1-Antitrypsin deficiency and lung disease: risk modification by occupational and environmental inhalants. European Respiratory Journal 2005;26:909-17.
77. Janciauskiene SM, Bals R, Koczulla R, et al. The discovery of alpha1-antitrypsin and its role in health and disease. Respiratory Medicine 2011;105:1129-39.
78. Silverman EK. Genetic Epidemiology of COPD. Chest 2002;121:1S-6S.
79. Abboud RT, Vimalanathan S. Pathogenesis of COPD. Part I. The role of protease-antiprotease imbalance in emphysema [State of the Art Series. Chronic obstructive pulmonary disease in high- and low-income countries. Edited by G. Marks and M. Chan-Yeung. Number 3 in the series]. The International Journal of Tuberculosis and Lung Disease 2008;12:361-7.
80. Stockley RA. Neutrophils and protease/antiprotease imbalance. American Journal of Respiratory and Critical Care Medicine 1999;160:S49-52.
81. MacNee W. Oxidative stress and lung inflammation in airways disease. European Journal of Pharmacology 2001;429:195-207.
82. Wood AM, Stockley RA. The genetics of chronic obstructive pulmonary disease. Respiratory Research 2006;7:130.
83. Rahman I, Adcock IM. Oxidative stress and redox regulation of lung inflammation in COPD. European Respiratory Journal 2006;28:219-42.
84. MacNee W. Pulmonary and systemic oxidant/antioxidant imbalance in chronic obstructive pulmonary disease. Proceedings of the American Thoracic Society 2005;2:50-60.
85. Hall IP, Lomas DA. The genetics of obstructive lung disease: big is beautiful. Thorax 2010;65:760-1.
86. Silverman EK, Spira A, Pare PD. Genetics and Genomics of Chronic Obstructive Pulmonary Disease. Proceedings of the American Thoracic Society 2009;6:539-42.

87. Smolonska J, Wijmenga C, Postma DS, et al. Meta-analyses on suspected chronic obstructive pulmonary disease genes: a summary of 20 years' research. American Journal of Respiratory and Critical Care Medicine 2009;180:618-31.
88. Hersh CP, DeMeo DL, Lange C, et al. Attempted Replication of Reported Chronic Obstructive Pulmonary Disease Candidate Gene Associations. American Journal of Respiratory Cell and Molecular Biology 2005;33:71-8.
89. Smolonska J, Wijmenga C, Postma DS, et al. Erratum: Meta-analyses on suspected chronic obstructive pulmonary disease genes: a summary of 20 years' research. American Journal of Respiratory and Critical Care Medicine 2010;181:765.
90. Cho MH, Boutaoui N, Klanderman BJ, et al. Variants in FAM13A are associated with chronic obstructive pulmonary disease. Nature Genetics 2010;42:200-2.
91. Pillai SG, Ge D, Zhu G, et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. PLoS Genetics 2009;5:e1000421.
92. Hancock DB, Eijgelsheim M, Wilk JB, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. Nature Genetics 2010;42:45-52.
93. Repapi E, Sayers I, Wain LV, et al. Genome-wide association study identifies five loci associated with lung function. Nature Genetics 2010;42:36-44.
94. Soler Artigas M, Wain LV, Repapi E, et al. Effect of 5 Genetic Variants Associated with Lung Function on the Risk of COPD, and their Joint Effects on Lung Function. American Journal of Respiratory and Critical Care Medicine 2011;184:786-95.
95. Obeidat Me, Wain LV, Shrine N, et al. A Comprehensive Evaluation of Potential Lung Function Associated Genes in the SpiroMeta General Population Sample. PloS One 2011;6:e19382.
96. Pillai SG, Kong X, Edwards LD, et al. Loci identified by genome-wide association studies influence different disease-related phenotypes in chronic obstructive pulmonary disease. American Journal of Respiratory and Critical Care Medicine 2010;182:1498-505.
97. Caramori G, Adcock I. Gene-environment interactions in the development of chronic obstructive pulmonary disease. Current Opinion in Allergy and Clinical Immunology 2006;6:323-8.
98. Walter R, Gottlieb DJ, O'Connor GT. Environmental and genetic risk factors and gene-environment interactions in the pathogenesis of chronic obstructive lung disease. Environmental Health Perspectives 2000;108 Suppl 4:733-42.
99. Cookson WO. State of the art. Genetics and genomics of chronic obstructive pulmonary disease. Proceedings of the American Thoracic Society 2006;3:473-5.
100. Yang IA, Fong KM, Zimmerman PV, et al. Genetic susceptibility to the respiratory effects of air pollution. Thorax 2008;63:555-63.
101. Minelli C, Wei I, Sagoo G, et al. Interactive effects of antioxidant genes and air pollution on respiratory function and airway disease: a HuGE review. American Journal of Epidemiology 2011;173:603-20.
102. Romieu I, Moreno-Macias H, London SJ. Gene by environment interaction and ambient air pollution. Proceedings of the American Thoracic Society 2010;7:116-22.
103. Guénégou A, Leynaert B, Bénessiano J, et al. Association of lung function decline with the heme oxygenase-1 gene promoter microsatellite polymorphism in a general population sample. Results from the European Community Respiratory Health Survey (ECRHS), France. Journal of Medical Genetics 2006;43:e43.
104. Siedlinski M, van Diemen CC, Postma DS, et al. Heme oxygenase 1 variations and lung function decline in smokers: proof of replication. Journal of Medical Genetics 2008;45:400.
105. Hackett TL, Stefanowicz D, Aminuddin F, et al. Effect of gene environment interactions on lung function and cardiovascular disease in COPD. International Journal of Chronic Obstructive Pulmonary Disease 2011;6:277-87.

106. He JQ, Connett JE, Anthonisen NR, et al. Glutathione S-transferase variants and their interaction with smoking on lung function. American Journal of Respiratory and Critical Care Medicine 2004;170:388-94.
107. Ege MJ, Strachan DP, Cookson WO, et al. Gene-environment interaction for childhood asthma and exposure to farming in Central Europe. The Journal of Allergy and Clinical Immunology 2011;127:138-44, 44 e1-4.
108. Thomas D. Methods for Investigating Gene-Environment Interactions in Candidate Pathway and Genome-Wide Association Studies. Annual Review of Public Health 2010;31:21-36.
109. Le Marchand L, Wilkens LR. Design Considerations for Genomic Association Studies: Importance of Gene-Environment Interactions. Cancer Epidemiology Biomarkers & Prevention 2008;17:263-7.
110. de Marco R, Accordini S, Antonicelli L, et al. The Gene-Environment Interactions in Respiratory Diseases (GEIRD) Project. International Archives of Allergy and Immunology 2010;152:255-63.
111. Brunekreef B, Holgate ST. Air pollution and health. The Lancet 2002;360:1233-42.
112. Dockery DW, Pope CA, 3rd, Xu X, et al. An association between air pollution and mortality in six U.S. cities. The New England Journal of Medicine 1993;329:1753-9.
113. Pope CA, 3rd, Thun MJ, Namboodiri MM, et al. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. American Journal of Respiratory and Critical Care Medicine 1995;151:669-74.
114. Abbey DE, Nishino N, McDonnell WF, et al. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. American Journal of Respiratory and Critical Care Medicine 1999;159:373-82.
115. Krewski D, Burnett RT, Goldberg M, et al. Reanalysis of the Harvard Six Cities Study, part I: validation and replication. Inhalation Toxicology 2005;17:335-42.
116. Krewski D BR, MS Goldberg MS, et al. Reanalysis of the Harvard six cities study and the American Cancer Society study of particulate air pollution and mortality. Boston: Health Effects Institute; 2000.
117. Krewski D, Jerrett M, Burnett RT, et al. Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality. Research Report of the Health Effects Institute 2009:5-136.
118. Brook Rd FB. Air Pollution and Cardiovascular Disease: A Statement for Healthcare Professionals From the Expert Panel on Population and Prevention Science of the American Heart Association. Circulation 2004:2655-71.
119. Oberdorster G, Finkelstein JN, Johnston C, et al. Acute pulmonary effects of ultrafine particles in rats and mice. Research Report of the Health Effects Institute 2000:5-74; disc 5-86.
120. Oberdorster G, Sharp Z, Atudorei V, et al. Extrapulmonary translocation of ultrafine carbon particles following whole-body inhalation exposure of rats. Journal of Toxicology and Environmental Health Part A 2002;65:1531-43.
121. Oberdorster G, Sharp Z, Atudorei V, et al. Translocation of inhaled ultrafine particles to the brain. Inhalation Toxicology 2004;16:437-45.
122. Brook RD, Rajagopalan S, Pope CA, 3rd, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation 2010;121:2331-78.
123. Ruckerl R, Schneider A, Breitner S, et al. Health effects of particulate air pollution: A review of epidemiological evidence. Inhalation Toxicology 2011;23:555-92.
124. Atkinson RW, Anderson HR, Sunyer J, et al. Acute effects of particulate air pollution on respiratory admissions: results from APHEA 2 project. Air Pollution and Health: a European Approach. American journal of respiratory and critical care medicine 2001;164:1860-6.
125. Katsouyanni K, Touloumi G, Samoli E, et al. Confounding and effect modification in the short-term effects of ambient particles on total mortality: results from 29 European cities within the APHEA2 project. Epidemiology 2001;12:521-31.

126. Samet JM, Dominici F, Zeger SL, et al. The National Morbidity, Mortality, and Air Pollution Study. Part I: Methods and methodologic issues. Research Report of the Health Effects Institute 2000:5-84.
127. Zanobetti A, Schwartz J, Dockery DW. Airborne particles are a risk factor for hospital admissions for heart and lung disease. Environmental Health Perspectives 2000;108:1071-7.
128. Peacock JL, Anderson HR, Bremner SA, et al. Outdoor air pollution and respiratory health in patients with COPD. Thorax 2011;66:591-6.
129. Hoek G, Brunekreef B, Goldbohm S, et al. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. The Lancet 2002;360:1203-9.
130. Ackermann-Liebrich U, Leuenberger P, Schwartz J, et al. Lung function and long term exposure to air pollutants in Switzerland. Study on Air Pollution and Lung Diseases in Adults (SAPALDIA) Team. American Journal of Respiratory and Critical Care Medicine 1997;155:122-9.
131. Zemp E, Elsasser S, Schindler C, et al. Long-term ambient air pollution and respiratory symptoms in adults (SAPALDIA study). The SAPALDIA Team. American Journal of Respiratory and Critical Care Medicine 1999;159:1257-66.
132. Kelly FJ, Fussell JC. Air pollution and airway disease. Clinical and Experimental Allergy 2011;41:1059-71.
133. Perez L, Rapp R, Kunzli N. The Year of the Lung: outdoor air pollution and lung health. Swiss Medical Weekly 2010;140:w13129.
134. Dockery DW, Cunningham J, Damokosh AI, et al. Health effects of acid aerosols on North American children: respiratory symptoms. Environmental Health Perspectives 1996;104:500-5.
135. Raizenne M, Neas LM, Damokosh AI, et al. Health effects of acid aerosols on North American children: pulmonary function. Environmental Health Perspectives 1996;104:506-14.
136. Gauderman WJ, Avol E, Gilliland F, et al. The effect of air pollution on lung development from 10 to 18 years of age. The New England Journal of Medicine 2004;351:1057-67.
137. Jerrett M, Shankardass K, Berhane K, et al. Traffic-related air pollution and asthma onset in children: a prospective cohort study with individual exposure measurement. Environmental Health Perspectives 2008;116:1433-8.
138. McConnell R, Islam T, Shankardass K, et al. Childhood incident asthma and traffic-related air pollution at home and school. Environmental Health Perspectives 2010;118:1021-6.
139. Sacks JD, Stanek LW, Luben TJ, et al. Particulate matter-induced health effects: who is susceptible? Environmental Health Perspectives 2011;119:446-54.
140. Lodovici M, Bigagli E. Oxidative stress and air pollution exposure. Journal of Toxicology 2011;2011:487074.
141. Araujo JA, Nel AE. Particulate matter and atherosclerosis: role of particle size, composition and oxidative stress. Particle and Fibre Toxicology 2009;6:24.
142. Epidemiology Standardization Project. e.c. American Thoracic Society. Recommended respiratory disease questionnaires for use with adults and children in epidemiological research. The American Review of Respiratory Disease 1978;118:7-52.
143. Burney PG, Luczynska C, Chinn S, et al. The European Community Respiratory Health Survey. European Respiratory Journal 1994;7:954-60.
144. Standardization of Spirometry, 1994 Update. American Thoracic Society. American Journal of Respiratory and Critical Care Medicine 1995;152:1107-36.
145. Kunzli N, Ackermann-Liebrich U, Keller R, et al. Variability of FVC and FEV1 due to technician, team, device and subject in an eight centre study: three quality control studies in SAPALDIA. Swiss Study on Air Pollution and Lung Disease in Adults. European Respiratory Journal 1995;8:371-6.
146. Kunzli N, Kuna-Dibbert B, Keidel D, et al. Longitudinal validity of spirometers--a challenge in longitudinal studies. Swiss Medical Weekly 2005;135:503-8.

147. Monn C, Brandli O, Schindler C, et al. Personal exposure to nitrogen dioxide in Switzerland. SAPALDIA team. Swiss Study on Air Pollution and Lung Diseases in Adults. Science of the Total Environment 1998;215:243-51.
148. Liu LJ, Curjuric I, Keidel D, et al. Characterization of source-specific air pollution exposure for a large population-based Swiss cohort (SAPALDIA). Environmental Health Perspectives 2007;115:1638-45.
149. SAEFL. Modelling of PM10 and PM2.5 ambient concentrations in Switzerland 2000 and 2010. Bern: Swiss Agency for the Environment, Forests, and Landscape (SAEFL); 2003.
150. SAEFL. Modelling of NO2 and benzene ambient concentrations in Switzerland 2000 to 2020. Bern: Swiss Agency for the Environment, Forests, and Landscape (SAEFL); 2004.
151. Imboden M, Downs SH, Senn O, et al. Glutathione S-transferase genotypes modify lung function decline in the general population: SAPALDIA cohort study. Respiratory Research 2007;8:2.
152. Moffatt MF, Gut IG, Demenais F, et al. A large-scale, consortium-based genomewide association study of asthma. The New England Journal of Medicine 2010;363:1211-21.
153. Li Y, Willer CJ, Ding J, et al. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genetic Epidemiology 2010;34:816-34.
154. Consortium IH. A haplotype map of the human genome. Nature 2005;437:1299-320.
155. Yu K, Li Q, Bergen AW, et al. Pathway analysis by adaptive combination of P-values. Genetic Epidemiology 2009;33:700-9.
156. Kraft P. Curses--winner's and otherwise--in genetic epidemiology. Epidemiology 2008;19:649-51; discussion 57-8.
157. Kampa M, Castanas E. Human health effects of air pollution. Environmental Pollution 2008;151:362-7.
158. Val S, Martinon L, Cachier H, et al. Role of size and composition of traffic and agricultural aerosols in the molecular responses triggered in airway epithelial cells. Inhalation Toxicology 2011;23:627-40.
159. Kramer CB, Cullen AC, Faustman EM. Policy implications of genetic information on regulation under the Clean Air Act: the case of particulate matter and asthmatics. Environmental Health Perspectives 2006;114:313-9.
160. Sundberg K, Johansson AS, Stenberg G, et al. Differences in the catalytic efficiencies of allelic variants of glutathione transferase P1-1 towards carcinogenic diol epoxides of polycyclic aromatic hydrocarbons. Carcinogenesis 1998;19:433-6.
161. Fridley BL, Biernacka JM. Gene set analysis of SNP data: benefits, challenges, and future directions. European Journal of Human Genetics 2011;19:837-43.
162. Kunzli N, Kaiser R, Medina S, et al. Public-health impact of outdoor and traffic-related air pollution: a European assessment. The Lancet 2000;356:795-801.
163. Schindler C, Keidel D, Gerbase MW, et al. Improvements in PM10 exposure and reduced rates of respiratory symptoms in a cohort of Swiss adults (SAPALDIA). American Journal of Respiratory and Critical Care Medicine 2009;179:579-87.
164. Rose G. Rose's strategy of preventive medicine. K.T. Khaw, M. Marmot, editors. Oxford: Oxford University Press; 2008.
165. Romieu I, Sienra-Monge JJ, Ramirez-Aguilar M, et al. Genetic polymorphism of GSTM1 and antioxidant supplementation influence lung function in relation to ozone exposure in asthmatic children in Mexico City. Thorax 2004;59:8-10.
166. Probst-Hensch NM. Chronic age-related diseases share risk factors: do they share pathophysiological mechanisms and why does that matter? Swiss Medical Weekly 2010;140:w13072.
167. Janssens AC, van Duijn CM. Genome-based prediction of common diseases: advances and prospects. Human Molecular Genetics 2008;17:R166-73.
168. Laird PW. Principles and challenges of genomewide DNA methylation analysis. Nature Reviews Genetics 2010;11:191-203.

Appendices

Appendix 1

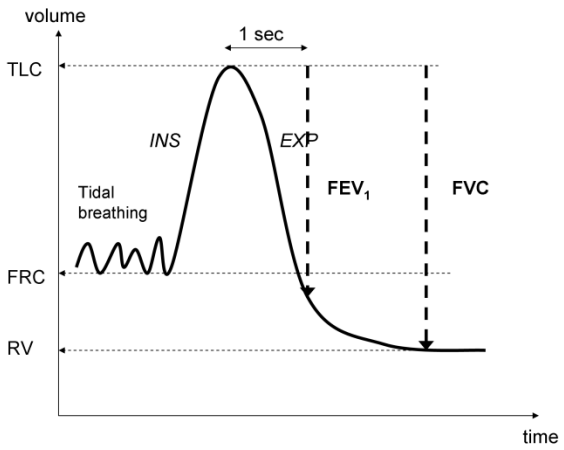
Lung function testing

Lung function testing is usually done with the tested person in an upright sitting position. Upon prompting by the lab technician, the person inspires maximally from normal tidal breathing, stays at full inspiration for 1-2 seconds, and then exhales as fast as possible (in a blast) until complete exhalation, where he/she needs to persist until the technician announces the end of the test. Two criteria are recommended to define the end of the test[1]. Formally, the end of test is reached if no changes in volume occur (or changes <25ml) in the volume–time curve for ≥1 second, whereby the person has made an exhalation effort for ≥6 seconds. Alternatively, the test is stopped if the person cannot or should not (for medical reasons) exhale further to prevent side-effects from the manoeuvre.

Figure A1 provides a graphical representation of a spirometry manoeuvre in form of a time-volume curve. The air volume exhaled during the 1st second of forced expiration is called forced expiratory volume in the first second (FEV_1), and the maximally exhaled volume forced vital capacity (FVC). Thereby the ratio of FEV_1/FVC is a measure of airway obstruction, as tissue remodelling and narrowing of the larger bronchi affects primarily FEV_1, and to a lesser extent and more indirectly FVC (the residual volume remaining in the lung after full exhalation can be increased due to trapped air and indirectly decrease FVC).

Lung function testing must fulfil quality standards, such as those set out by the ATS and ERS[1]: Spirometry devices need to be regularly maintained and calibrated. Probands must be instructed by

Figure A1 Volume-time curve of spirometric FEV_1 and FVC manoeuvres



INS: inspiration; EXP: expiration; FEV_1: expiratory volume in 1st second of forced expiration; FVC: forced vital capacity, the maximally exhalable volume; FRC: functional residual volume, the air volume remaining in the lung after normal exhalation during tidal breathing; RV: residual volume, air volume remaining in the lung after maximal forced exhalation; TLC: total lung capacity, air volume in the lung after maximal inhalation

the lab technicians on how to perform the testing correctly, and must be continuously motivated by them to invest the largest possible effort in the manoeuvres. A detailed list of criteria for acceptable test quality was established to ensure valid and reliable lung function measurements (table A1)

Applying bronchodilation prior to spirometric testing is important to distinguish fixed, anatomical narrowing of the airways as present in COPD from a reversible airway obstruction induced by hyper-reactive smooth airway musculature, as typically present in asthmatic disease. The bronchodilating agent thereby relaxes the airway muscle layer. It has been shown that irreversible airway obstruction rates can be overestimated by up to 50% when relying on only pre-bronchodilation spirometry[2-5]. Based on data from a Norwegian population-based study, it seems that mild obstruction stages seem to be most prone to this kind of misclassification[6]. There is no detailed consensus about how to perform bronchodilation, but the ATS/ERS taskforce for standardization of spirometry recommends applying four separate doses of 100µg of a short-acting bronchodilator such as salbutamol[7]. The doses should be administered by a metered dose inhaler with a spacer. Lung function testing should then be repeated 15 minutes after application.

Table A1 Criteria for acceptability of spirometric testing according to the ATS/ERS guidelines, from [1]

Within-manoeuvre criteria

- Individual spirograms are ''acceptable'' if
  - They are free from artefacts
    - Cough during the first second of exhalation
    - Glottis closure that influences the measurement
    - Early termination or cut-off
    - Effort that is not maximal throughout
    - Leak
    - Obstructed mouthpiece
  - They have good starts
    - Extrapolated volume <5% of FVC or 0.15 litres, whichever is greater
  - They show satisfactory exhalation
    - Duration of ≥6 seconds (3 seconds for children) or a plateau in the volume–time curve or
    - If the subject cannot or should not continue to exhale

Between-manoeuvre criteria

- After three acceptable spirograms have been obtained, apply the following tests
  - The two largest values of FVC must be within 0.150 litres of each other
  - The two largest values of FEV_1 must be within 0.150 litres of each other

If both of these criteria are met, the test session may be concluded

If both of these criteria are not met, continue testing until

- Both of the criteria are met with analysis of additional acceptable spirograms or
- A total of eight tests have been performed (optional) or
- The patient/subject cannot or should not continue

Save, as a minimum, the three satisfactory manoeuvres

References:

1. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. European Respiratory Journal 2005;26:319-38.
2. Johannessen A, Omenaas E, Bakke P, Gulsvik A. Incidence of GOLD-defined chronic obstructive pulmonary disease in a general adult population. International Journal of Tuberculosis and Lung Disease 2005;9:926-32.
3. Kim SJ, Suk MH, Choi HM, et al. The local prevalence of COPD by post-bronchodilator GOLD criteria in Korea. Int J Tuberc Lung Dis 2006;10:1393-8.
4. Pérez-Padilla R, Hallal PC, Vázquez-García JC, et al. Impact of Bronchodilator Use on the Prevalence of COPD in Population-Based Samples. COPD: Journal of Chronic Obstructive Pulmonary Disease 2007;4:113-20.
5. Vollmer WM, Gíslason, Burney P, et al. Comparison of spirometry criteria for the diagnosis of COPD: results from the BOLD study. European Respiratory Journal 2009;34:588-97.
6. Johannessen A, Omenaas ER, Bakke PS, Gulsvik A. Implications of reversibility testing on prevalence and risk factors for chronic obstructive pulmonary disease: a community study. Thorax 2005;60:842-7.
7. Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. European Respiratory Journal 2005;26:948-68.

Appendix 2

Measurement of air pollution and individual exposure assessment

Different devices and methods are available for measuring gaseous and particulate matter pollutants. Pollutant gases can be assessed using methods based on chemoluminesence where the substances or their chemical predecessors (after prior conversion) emit light signals proportional to their concentration in the air, or by passive diffusion sampling tubes which absorb the substances over several days and are then analyzed in a laboratory. The gold standard for measuring particulate matter exposure consists of using high volume samplers. These devices are run over 24h or more during which they suck a known amount of air through a filter. The filter is changed regularly and the amount of deposited particles is determined by weighting under standardized conditions in a lab. Knowing the sampling volume allows to calculate the respective mean concentration per cubic meter. Newer sampling devices measure particulate matter concentrations continuously using oscillating microbalances. Air is sucked through a thin, oscillating glass tube with a filter on its tip. The tube oscillation frequency, which depends on the particle mass impacted on the filter, allows calculating the concentration.

To obtain estimates on personal exposure from measured air pollution concentrations in the field, several physicochemical characteristics of the pollutants must be considered. Due to suspension and transport in the air as well as relative chemical inertness compared to gaseous pollutants, particulate matter pollution is more homogeneously distributed at the local level. From the methodological side of view, the reduced spatial and temporal variability of particulate matter concentration makes it easier to obtain relatively accurate exposure measurements in epidemiological studies.

An important aspect is thereby how well ambient measurements of particulate matter exposure correlate with personal exposures. In urban areas, people spend much time indoors, and most of it at home, which is the reason why home outdoor exposures are frequently used as proxy for individual exposure .It has been shown that outdoor fine particulate matter measurements ($PM_{2.5}$) correlate well with individual exposure[1], and spatiotemporal variability of individual exposure is well approximated by measured outdoor variability[2]. On the other side, obtaining accurate exposure estimates for gaseous pollutants is more difficult, as pollutants are more volatile and undergo chemical reactions with atmospheric components.

As a consequence methods to improve exposure assessment and accurately estimate individual exposures based on environmental measurements have been more developed for PM exposure. PM exposure was initially estimated by using measurements from monitoring devices placed at key locations in a city or region of interest. Often, studies just used data from air pollution monitors already operated by environmental agencies. Participants were then assigned the values measured at the closest monitoring station to their residence address as exposure estimate.

Estimation methods were subsequently refined by more and more sophisticated prediction models, which are usually calibrated on a set of measurements. Techniques include use of different statistical interpolation techniques, geographic information system (GIS) or land use data around air pollution monitors, modeling of pollution levels based on meteorological data and national inventories on pollution emissions[3,4], or newly also by using data on atmospheric composition from satellites, which

can also estimate gaseous pollution[5]. Nowadays, hybrid models combining these techniques with personal or environmental measurements are frequently used[3].

References:

1. Janssen NA, de Hartog JJ, Hoek G, et al. Personal exposure to fine particulate matter in elderly subjects: relation between personal, indoor, and outdoor concentrations. Journal of the Air and Waste Management Association 2000;50:1133-43.
2. Oglesby L, Kunzli N, Roosli M, et al. Validity of ambient levels of fine particles as surrogate for personal exposure to outdoor air pollution--results of the European EXPOLIS-EAS Study (Swiss Center Basel). Journal of the Air and Waste Management Association 2000;50:1251-61.
3. Jerrett M, Arain A, Kanaroglou P, et al. A review and evaluation of intraurban air pollution exposure models. Journal of Exposure Analysis and Environmental Epidemiology 2005;15:185-204.
4. Zou B, Wilson JG, Zhan FB, Zeng Y. Air pollution exposure assessment methods utilized in epidemiological studies. Journal of Environmental Monitoring 2009;11:475-90.
5. van Donkelaar A, Martin RV, Brauer M, et al. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. Environmental Health Perspectives 2010;118:847-55.

Appendix 3

Short Version of Health questionnaire used in SAPALDIA 2 (German)

Question Number	*Question* *☐ Answers*
T_H00010	Haben Sie in den letzten 12 Monaten irgendwann ein pfeifendes Atemgeräusch in der Brust gehabt? ☐ nein → *gehen Sie bitte zu Frage T_H00040, S. 1* ☐ ja ☐ weiss nicht ☐ Weigerung
T_H00020	Haben Sie in den letzten 12 Monaten Mühe gehabt mit Atmen, wenn Sie dieses pfeifende Atemgeräusch in der Brust gehabt haben? ☐ nein ☐ ja ☐ weiss nicht ☐ Weigerung
T_H00030	Haben Sie in den letzten 12 Monaten dieses pfeifende Atemgeräusch gehabt, ohne dass Sie gleichzeitig erkältet waren? ☐ nein ☐ ja ☐ weiss nicht ☐ Weigerung
T_H00040	Sind Sie in den letzten 12 Monaten irgendwann aufgewacht mit einem Druckgefühl oder Engegefühl in der Brust? ☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00050 Haben Sie in den letzten 12 Monaten tagsüber einen Anfall von Atemnot gehabt, wenn Sie ruhig waren? (gemeint ist "in Ruhe")

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00060 Haben Sie in den letzten 12 Monaten einen Anfall von Atemnot nach körperlicher Anstrengung gehabt?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00070 Sind Sie in den letzten 12 Monaten jemals aufgewacht, weil sie einen Anfall von Atemnot gehabt haben?

☐ nein ⟶ *gehen Sie bitte zu Frage T_H00100, S. 2*

☐ ja
☐ weiss nicht
☐ Weigerung

T_H00080 Sind Sie in den letzten 3 Monaten durchschnittlich mindestens einmal in der Woche mit einem Anfall von Atemnot aufgewacht?

☐ nein
☐ ja

☐ weiss nicht

☐ Weigerung

T_H00100 Sind Sie in den letzten 12 Monaten jemals wegen eines Hustenanfalles aufgewacht?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00110 Husten Sie normalerweise morgens nach dem Aufstehen?

☐ nein ——— *gehen Sie bitte zu Frage T_H00130, S. 2*

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00120 In welchen Jahreszeiten husten Sie normalerweise morgens nach dem Aufstehen?

☐ unabhängig von der Jahreszeit

☐ nur im Winter

☐ nur im Frühling, Sommer oder Herbst

☐ weiss nicht

☐ Weigerung

T_H00130 Husten Sie normalerweise tagsüber oder nachts?

☐ nein ⟶ *wenn T_H00110 „nein" **und** T_H00130 „nein"*

gehen Sie bitte zu Frage T_H00170, S. 3

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00140 Husten Sie so an den meisten Tagen während mindestens 3 Monaten im Jahr?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00150 In welchen Jahreszeiten husten Sie normalerweise tagsüber oder in der Nacht?

☐ unabhängig von der Jahreszeit

☐ nur im Winter

☐ nur im Frühling, Sommer oder Herbst

☐ weiss nicht

☐ Weigerung

T_H00160 Seit wie vielen Jahren?

_____(Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H00170 Haben Sie normalerweise Auswurf morgens nach dem Aufstehen?

☐ nein ⟶ *gehen Sie bitte zu Frage T_H00190, S. 3*

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00180 In welchen Jahreszeiten haben Sie normalerweise Auswurf morgens nach dem Aufstehen?

☐ unabhängig von der Jahreszeit

☐ nur im Winter

☐ nur im Frühling, Sommer oder Herbst

☐ weiss nicht

☐ Weigerung

T_H00190 Haben Sie normalerweise tagsüber oder nachts Auswurf?

☐ nein ⟶ *wenn T_H00170 „nein" **und** T_H00190 „nein"*

gehen Sie bitte zu Frage T_H00310, S. 4

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00200 Haben Sie normalerweise an den meisten Tagen während mindestens 3 Monaten pro Jahr solchen Auswurf?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00210 In welchen Jahreszeiten haben Sie normalerweise tagsüber oder nachts Auswurf?

☐ unabhängig von der Jahreszeit
☐ nur im Winter

☐ nur im Frühling, Sommer oder Herbst
☐ weiss nicht

☐ Weigerung

T_H00220 Seit wie vielen Jahren?

_____(Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H00310 Haben Sie jemals Asthma gehabt?

☐ nein ⟶ *gehen Sie bitte zu Frage T_H00500, S. 5*

☐ ja
☐ weiss nicht

☐ Weigerung

T_H00320 Wurde dies von einem Arzt bestätigt?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00370 Haben Sie in den letzten 12 Monaten einen Asthmaanfall gehabt?

☐ nein ⟶ *gehen Sie bitte zu Frage T_H00430, S. 4*

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00380 Wie viele Asthmaanfälle haben Sie in den letzten 12 Monaten gehabt?

_____(Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H00390 Wie viele Asthmaanfälle haben Sie in den letzten 3 Monaten gehabt?

_____(Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H00430 Nehmen Sie zur Zeit irgendwelche Medikamente gegen Asthma (auch Inhalationsmittel, Aerosole oder Tabletten)?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00500 Haben Sie allergischen Schnupfen oder Heuschnupfen?

☐ nein ⟶ *gehen Sie bitte zu Frage T_H00520, S. 5*

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00640 Haben Sie in diesem Jahr schon Heuschnupfen gehabt?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00520 Hatten Sie jemals Probleme mit Niesen oder mit einer laufenden oder verstopften Nase, ohne erkältet zu sein oder eine Grippe zu haben?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H00730 Haben Sie eine chronische Erkrankung, die Sie in irgendeiner Weise einschränkt?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

Haben Sie etwas von dem Folgenden?

T_H00740 Arthritis

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00741 Hoher Blutdruck

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00742 Schwerhörigkeit

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00743 Krampfadern

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00744 Grauer Star (Linsentrübung)

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00745 Herzkrankheiten

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00746 Depression

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00747 Diabetes/Zuckerkrankheit

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00748 Migräne/oft auftretende

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00749 Krebs (Stellen Sie die Frage so: Haben Sie Krebs gehabt?)

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00750 Schlaganfall

☐ nein

☐ ja, aber nicht vom Arzt diagnostiziert

☐ ja, vom Arzt diagnostiziert

☐ weiss nicht

☐ Weigerung

T_H00880 Mit welchem Alter haben Sie Ihre vollzeitliche Ausbildung abgeschlossen?

(0 entspricht hauptberuflich Student)

_____ (Zahl einfüllen)

☐ weiss nicht

☐ Weigerung

Was machen Sie zur Zeit?

T_H00890 voll erwerbstätig

☐ Nein

☐ Ja

☐ weiss nicht

☐ Weigerung

T_H00891 teilweise erwerbstätig

☐ Nein

☐ Ja

☐ weiss nicht

☐ Weigerung

T_H00892 Hausfrau/Hausmann

☐ Nein

☐ Ja

☐ weiss nicht

☐ Weigerung

T_H00893 in Ausbildung

☐ Nein

☐ Ja

☐ weiss nicht

☐ Weigerung

T_H00894 pensioniert/Rentner

☐ Nein

☐ Ja

☐ weiss nicht

☐ Weigerung

T_H00895 arbeitslos

☐ Nein

☐ Ja

☐ weiss nicht

☐ Weigerung

T_H00896 längerer Militärdienst (z.B. RS), längere Ferien (z.B. nach Schulabschluss oder zwischen zwei Stellen)

☐ Nein

☐ Ja

☐ weiss nicht

☐ Weigerung

T_H00897 krank oder invalid

☐ Nein

☐ Ja

☐ weiss nicht

☐ Weigerung

T_H00898 mache etwas anderes

☐ Nein

☐ Ja

☐ weiss nicht

☐ Weigerung

T_H01000 Haben Sie jemals in einem Beruf gearbeitet, bei dem Sie Dampf, Gas, taub oder Rauch ausgesetzt waren?

☐ Nein

☐ Ja

☐ weiss nicht

☐ Weigerung

T_H01340 Leben Sie in derselben Wohnung/Haus wie in der letzten Untersuchung?

- ☐ Nein
- ☐ Ja
- ☐ weiss nicht
- ☐ Weigerung

T_H01720 Welche Aussage beschreibt Ihre Wohnsituation am besten? Ich wohne ...

- ☐ im Stadt/Dorfzentrum an stark befahrener Strasse
- ☐ im Stadt/Dorfzentrum an wenig bis mässig befahrener Strasse
- ☐ im Aussenquartier/am Dorfrand an mässig bis stark befahrener Strasse
- ☐ im Aussenquartier/am Dorfrand an wenig befahrener Strasse
- ☐ in alleinstehenden Haus auf dem Land
- ☐ weiss nicht
- ☐ Weigerung

T_H01730 Wie gross ist werktags das Verkehrsaufkommen auf der Strasse, an welcher Sie wohnen?

- ☐ Stark befahrene Strasse/ununterbrochener Verkehrsfluss
- ☐ Mässig befahrene Strasse/viele Autos fahren vorbei
- ☐ Wenig befahrene Strasse/nur ab und zu ein paar Autos
- ☐ weiss nicht
- ☐ Weigerung

T_H01740 Wie oft fahren an Wochentagen Lastwagen durch die Strasse, an welcher Sie wohnen?

- ☐ nie
- ☐ selten
- ☐ öfter am Tag

☐ fast den ganzen Tag

☐ weiss nicht

☐ Weigerung

T_H02040 Haben Sie schon einmal mindestens ein Jahr lang geraucht?

(„Ja" heisst mindestens 20 Zigarettenpackungen oder 360g Tabak im ganzen Leben ODER: mindestens 1 Zigarette pro Tag, oder eine Zigarre pro Woche für ein Jahr).

☐ nein ⟶ *gehen Sie bitte zu Frage T_H02150, S. 10*

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02050 In welchem Alter haben Sie angefangen, regelmässig zu rauchen?

_____ (Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H02060 Rauchen Sie zur Zeit (im letzten Monat)?

☐ nein ⟶ *gehen Sie bitte zu Frage T_H02105, S. 9*

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02070 Wie viel rauchen Sie jetzt im Durchschnitt?

Anzahl Zigaretten pro Tag

_____ (Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H02072 Anzahl Zigarren pro Woche

_____ (Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H02073 Pfeifentabak in Gramm pro Woche

_____ (Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H02105 In welchem Alter haben Sie aufgehört zu rauchen?

_____ (Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H02110 In der gesamten Zeit, in der Sie rauchten, haben Sie durchschnittlich wie viel geraucht?

Anzahl Zigaretten pro Tag

_____ (Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H02150 Sind Sie in den letzten 12 Monaten regelmässig Tabakrauch ausgesetzt gewesen? (regelmässig heisst, an den meisten Tagen oder Nächten)

☐ nein ⟶ *gehen Sie bitte zu Frage T_H02280, S. 10*

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02160 Sie selber nicht mitgezählt, wie viele Personen rauchen in Ihrem Haushalt?

_____ (Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H02170 Rauchen an Ihrem Arbeitsplatz andere Personen regelmässig?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02190 Wie viele Stunden sind Sie täglich dem Tabakrauch von anderen Leuten ausgesetzt?

_____ (Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H02280 Haben Sie seit der letzten Untersuchung jemals inhalierbare Glucocorticoide (Kortison) benutzt? (Liste zeigen)

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02390 Haben Sie in den letzten 12 Monaten die Notfallstation eines Spitals aufgesucht wegen Atemproblemen?

☐ nein ⟶ *gehen Sie bitte zu Frage T_H02420, S. 11*

☐ ja
☐ weiss nicht

☐ Weigerung

T_H02400 War dies wegen Asthma, Mühe mit der Atmung oder wegen des pfeifenden Atemgeräusches?

☐ nein

☐ ja

☐ weiss nicht
☐ Weigerung

T_H02410 Wie oft in den letzten 12 Monaten?

_____ (Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

T_H02420 Haben Sie in den letzten 12 Monaten die Notfallstation eines Spitals aufgesucht wegen Herz-Kreislaufproblemen?

☐ nein ⟶ *gehen Sie bitte zu Frage T_H02480, S. 11*

☐ ja

☐ weiss nicht

☐ Weigerung

War dies wegen

T_H02430 Angina pectoris

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02431 Herzinfarkt

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02432 Herzrhythmusstörungen

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02480 Haben Sie in den letzten 12 Monaten eine Nacht in einem Spital verbracht wegen Atemproblemen?

☐ nein ⟶ *gehen Sie bitte zu Frage T_H02520, S. 12*

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02490 War dies wegen Asthma, Mühe mit der Atmung oder pfeifender Atmung?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02520 Haben Sie in den letzten 12 Monaten eine Nacht in einem Spital verbracht wegen Herz-/Kreislaufproblemen?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02590 Sind Sie in den letzten 12 Monaten von einem Arzt untersucht worden wegen Atembeschwerden oder wegen Mühe mit der Atmung?

☐ nein ⟶ *gehen Sie bitte zu Frage T_H02650, S. 12*

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02600 War dies wegen Asthma, wegen Mühe mit der Atmung oder wegen eines pfeifenden Atemgeräusches?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H02650 Sind Sie in den letzten 12 Monaten von einem Arzt untersucht worden wegen Herz-/Kreislaufbeschwerden?

☐ nein

☐ ja

☐ weiss nicht

☐ Weigerung

T_H03030 Wie viele Tage konnten Sie in den letzten 12 Monaten wegen Asthma, wegen Mühe mit der Atmung oder wegen pfeifender Atemgeräusche nicht zur Arbeit gehen?

_____ (Zahl einfüllen)
☐ weiss nicht

☐ Weigerung

Appendix 4

Percent explained variability in lung function decline by interactions of oxidative stress genes with PM_{10} and tobacco smoke exposure.

exposure / model	decline in FEV_1			decline in FEV_1/FVC			decline in FEF_{25-75}		
	G*E	G, E only	E only	G*E	G, E only	E only	G*E	G, E only	E only
interval PM_{10}	28.5	20.4		18.9	10.4		13.5	7.9	
pack years	26.1	19.2		15.6	10.1		13.2	8.6	
both	33.3	19.7	19.9	22.5	9.7	10.8	18.0	8.0	8.4

Values are adjusted R2-values from linear regression models including the strongest interacting SNP from each nominally significant gene (p-value for interaction<0.05, see table 4 in the manuscript for the full list of genes) simultaneously. All models controlled for covariates sex, age and height at follow-up, principal components of population stratification, study area, pack years smoked at baseline and during follow-up and PM_{10}-exposure during follow-up. Environmental exposure means interval PM_{10}-exposure or pack years smoked during follow-up.

G*E: full gene-environment interaction model including a multiplicative interaction term
G, E only: model specifying only SNP- and environmental main effects, but no interaction term
E only: model specifying only environmental main effects, without SNPs

Curriculum Vitae

Personal Information

Name:	Ivan Curjuric, MD, MSc Epidemiology
Function:	Research Assistant, Biostatistician, PhD student
Business address:	Swiss Tropical and Public Health Institute SwissTPH Socinstrasse 57, P.O. Box, 4002 Basel, Switzerland
Business phone:	061 /284 86 12
FAX:	061 /284 81 05
e-mail:	ivan.curjuric @unibas.ch
Home address:	Bernhofstrasse 51, 8134 Adliswil
Birth date:	29.08.1975
Marital status:	married, 2 children

Education

1994	Matura Typus B, Kantonsschule Freudenberg, Zürich, Switzerland
2000	Medical Diploma, University of Zürich, School of Medicine, Zürich, Switzerland
2003	Medical dissertation, University of Zürich (promotion date Sept 30th 2003), Zürich, Switzerland
2009	Master in Epidemiology, University of London, London School of Hygiene and Tropical Medicine, London, United Kingdom
2007-2012	PhD studies in Epidemiology, University of Basel, Basel, Switzerland

Positions and Employment

2000-2002	Assistant medical doctor, Swiss Poison Control Centre, Zürich, Switzerland
2002-2003	Assistant medical doctor, Department of Surgery, Regional Hospital of Lachen, Lachen, Switzerland
2004	Research Assistant, Institute of Social and Preventive Medicine Basel, Basel, Switzerland
2004-2005	Assistant medical doctor, Department of Geriatric Medicine, City Hospital Waid, Zürich, Switzerland
2005-2009	Research Assistant/Biostatistician, Institute of Social and Preventive Medicine, Basel, Switzerland
2010-2012	Research Assistant/Biostatistician, Swiss Tropical and Public Health Institute SwissTPH, Basel, Switzerland

Other Experience

2003	Medical Dissertation: “Fehler in der Medikamententherapie im Fallkollektiv des Schweizerischen *Toxikologischen* Informationszentrums“. Descriptive analysis of cases of drug treatment errors occurring in health institutions or at home, and recorded in the databases of the Swiss Poison Control Centre.
2004	Elaboration of the report „Krebsmonitoring in der Schweiz und Europa“ by Curjuric I, Quinto C and Ackermann-Liebrich U, Basel, Institute of Social- and Preventive Medicine, 2004. The report, written in preparation of an expert meeting on cancer monitoring (Bern, March 2005), comprised the collection and description of indicators for cancer monitoring available at the national level, taking colon cancer as an example.
2009	Master-Thesis in Epidemiology: “Developmental origins of asthma and atopic disease, the role of maternal nutrition and epigenetic effects: a birth cohort study in Switzerland”. Elaboration of a study design as final master thesis for the MSc course in Epidemiology at London School of Hygiene and Tropical Medicine.

Publication list (as of October 2011):

1. Artigas MS, Loth DW, Wain LV, et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. Nat Genet 2011. Epub Sept 25. doi: 10.1038/ng.941
2. Curjuric I, Zemp E, Dratva J, et al. Determinants of change in airway reactivity over 11 years in the SAPALDIA population study. Eur Respir J 2011;37:492-500.
3. Moffatt MF, Gut IG, Demenais F, et al. A large-scale, consortium-based genomewide association study of asthma. N Engl J Med 2010;363:1211-21.
4. Probst-Hensch NM, Curjuric I, Pierre-Olivier B, et al. Longitudinal change of prebronchodilator spirometric obstruction and health outcomes: results from the SAPALDIA cohort. Thorax 2010;65:150-6.
5. Bridevaux PO, Probst-Hensch NM, Schindler C, et al. Prevalence of airflow obstruction in smokers and never-smokers in Switzerland. Eur Respir J 2010;36:1259-69.
6. Dratva J, Schindler C, Curjuric I, et al. Perimenstrual increase in bronchial hyperreactivity in premenopausal women: results from the population-based SAPALDIA 2 cohort. J Allergy Clin Immunol 2010;125:823-9.
7. Curjuric I, Imboden M, Schindler C, et al. HMOX1 and GST variants modify attenuation of FEF25-75-decline due to PM10 reduction. Eur Respir J 2009.
8. Imboden M, Schwartz J, Schindler C, et al. Decreased PM10 exposure attenuates age-related lung function decline: genetic variants in p53, p21, and CCND1 modify this effect. Environ Health Perspect 2009;117:1420-7.
9. Bridevaux PO, Gerbase MW, Schindler C, et al. Sex-specific effect of body weight gain on systemic inflammation in subjects with COPD: results from the SAPALDIA cohort study 2. Eur Respir J 2009;34:332-9.
10. Liu LJ, Curjuric I, Keidel D, et al. Characterization of source-specific air pollution exposure for a large population-based Swiss cohort (SAPALDIA). Environ Health Perspect 2007;115:1638-45.

Accomplished doctoral training

Course	*ECTS Points*	*Certificate[1]*
Advanced Methods in Epidemiology: Analysis of Clustered Data and Multilevel Modeling. Part of MPH (master of public health) program Location: ISPM Bern Date: Mai 2007 Guidance: Dr. Martin Röösli; several lecturers	1.5	CP
Kurs Literaturrecherche University of Basel Date: 24. & 26. Nov 2007 Guidance: Dr. Philipp Mayer	0.5	CP
Practical Bayesian Methods for the Health Sciences SSPH+ PhD-course Location: ISPM Bern Date: March 20 2008 Guidance: Prof. Sander Greenland	0.5	CP
Umweltepidemiologie Part of MPH (master of public health) program Location: Basel Date: August 2008 Guidance: Prof. Dr. Charlotte Braun-Fahrländer	3.5	E
2nd Paris Workshop on Molecular and Statistical Genomic Epidemiology 3-day training course within the GABRIEL project Location: Paris Date: 13.-16. Mai 2007 Guidance: Prof. Ivo Gut	0.75*	CP
Working with the HapMap Wellcome Trust 4-day training course Location: Wellcome Trust Genome Campus, Hinxton, Cambridge Date: November 2007 Guidance: Prof. Manolis Dermitzakis	1*	CP
Advances in Population-based Studies of Complex Genetic Disorders Netherlands Institute of Health Sciences course Location: Erasmus MC, Rotterdam Date: Spring 2008 Guidance: Prof. Dr. Cornelia van Duijn	1.4	CP
Bioinformatics 101 Introductory E-learning course of the University of Bern Date: Spring 2008 Guidance: Dr. Marc Solioz	2.5	CP
GABRIEL Workshop on Genome Wide Analysis Methods University Hospital Groningen, The Netherlands Date: 25. & 26. May 2009 Guidance: Prof. Dirkje S Postma Prof. M. Boezen	0.75	P

[1] E=Examination, W=Written assignment CP = confirmation of participation, P = participation

* Estimated ECTS credit points based on working hours (30h/point)

Congress attendance

• European Respiratory Society ERS Annual Congress, Stockholm Sept2007	Oral presentation
• Swiss Public Health Congress, Zürich July 2009	Oral presentation at workshop on Genomic Literacy
• Swiss Public Health Congress, Basel, August 2011	Oral Presentation
• International Society for Environmental Epidemiology ISEE Annual Congress, Barcelona Sept 2011	Oral Presentation at symposium on gene air pollution interactions
• International Genetic Epidemiology Society IGES Annual Congress, Heidelberg Sept 2011	Poster Presentation

Given seminars

• Student Seminar, SwissTPH Aug 2010	Oral Presentation of PhD-work
• Workshop on Pathway Analysis Methods, Sept 2010 Paris, France	Oral Presentation of pathway analysis methods for studying gene-environment interactions
• Student Seminar, SwissTPH May 2011	Oral Presentation of work in progress (pathway analysis of gene-air pollution interactions)
• Omics Group Meeting, Swiss TPH May 2011	Oral Presentation of work in progress (pathway analysis of gene-air pollution interactions)

Teaching activities

- Introductory Lecture on Genetic Epidemiology in MPH Genetic Epidemiology course Oct 2008, University of Zürich, guidance: Prof. N. M Probst-Hensch), SwissTPH, Basel
- Tutoring in Basic of Medical Statistics for medical students, University of Basel, yearly collaboration from 2007 to 2010, guidance: Prof. Ch. Schindler, SwissTPH, Basel