General introduction and outline of thesis
The lungs have an important function to provide the body with oxygen and eliminate carbon dioxide from the blood. Breathing keeps us alive, and an impaired lung function adversely affects this process. Research into the pathophysiology of lung diseases aims to determine the factors which contribute to an impaired lung function, in order to reach the ultimate goal of preventing and treating respiratory diseases. Lung function tests and the genotype arrays are important tools that help us to understand lung health and to investigate pathways that lead to an impaired lung function and respiratory disease.

1. CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD)

COPD is the third leading cause of death worldwide (1). According to the Global Initiative for Chronic Obstructive Lung Disease (GOLD), “COPD is a common, preventable and treatable disease that is characterized by persistent respiratory symptoms and airflow limitation that is due to airway and/or alveolar abnormalities usually caused by significant exposure to noxious particles or gases”(2). COPD has two features; small airways disease (chronic bronchitis), and/or parenchymal destruction (emphysema), of which the relative contributions might vary from one individual to another.

1.1. Risk factors

Tobacco smoking is considered as the most important risk factor, in addition to occupational, outdoor and indoor air pollution (e.g. biomass fuel, dust, fumes and agricultural pesticides (3, 4)). Other factors also contribute to the development of COPD, including genetics, chronic bronchitis, asthma and airway hyper-responsiveness, poor lung growth during childhood, but also aging, female sex and low socioeconomic status (5-8).

1.2. Comorbidities and consequences

Comorbidities are abundant in COPD (9-11), as COPD often coexists with pulmonary diseases such as asthma (12), pulmonary fibrosis (13) and lung cancer (14, 15) and with extrapulmonary disease conditions (16) such as cerebrovascular disease (including stroke (17-19)), cognitive impairments (20, 21), obstructive sleep apnoea, cardiovascular disease (e.g. sudden cardiac death (22), peripheral artery disease (23, 24)), osteoporosis (25-29), metabolic disorders, depression and anxiety, skeletal muscle dysfunction (30, 31) and gastrointestinal reflux (32, 33). It is unclear whether those co-morbidities are the consequence of COPD or simply co-occur with COPD resulting from shared risk factors (10).
Individuals with COPD have distinct features which define the disease progression (34, 35). Late stage COPD can lead to serious complications, including COPD exacerbation with hospitalization, pulmonary hypertension and an increased risk of mortality.

2. LUNG FUNCTION

Lung function tests are important to detect and diagnose chronic lung diseases, assess disease severity, and monitor disease progression and treatment effects. The two most widely used lung function tests are spirometry and diffusing capacity of carbon dioxide, each of them assessing a distinct physiological function of the airways and lungs, respectively. Below, these tests are described in short.

2.1. Spirometry

Spirometry is a technique that measures the ability of the lungs to transport the air in and out of the lungs. The two most important measures in spirometry are the Forced Expiratory Volume in one second (FEV₁) and the Forced Vital Capacity (FVC). FEV₁ provides information about the expiratory airflow, and the FVC about the maximal amount of air that is exhaled. The ratio of these two measurements FEV₁/FVC, the so-called Tiffeneau index, is used for the diagnosis of chronic obstructive lung diseases. A reduced FEV₁/FVC ratio is indeed indicative of the presence of airflow limitation, which is a key diagnostic criterion for the diagnosis of COPD, but can also be reduced in patients with uncontrolled or severe asthma.

Lung function parameters decline with age, and the rate of decline depends on the exposure to environmental factors, such as smoking and air pollution, and genetic predisposition (36). Impaired lung function can be determined by many factors, including impaired lung growth during foetal life, childhood and adolescence. Importantly, impaired lung function also serves as an independent determinant for adverse health outcomes, including subsequent COPD (37, 38), cardiovascular disease (39), hospitalizations (40) and all-cause mortality (38, 40-42).

2.2. Diffusing capacity

The diffusing capacity of the lung for carbon monoxide (DLCO; also called transfer factor, TLCO) measures the total capacity of the body to exchange gases by the single-breath technique (SBT). Herein, small concentrations of carbon monoxide (CO) and a tracer gas (helium or another inert gas) are added to the inspiratory air in order to determine the levels of gas exchange. Patients are asked to inhale small amounts of CO, hold their breath for 10 seconds and exhale. Exhaled air is then sampled and the difference
between the added and the exhaled CO concentration yields the diffusing capacity (Figure 1) (43-46). DLCO is the product of two components that can be obtained by the SBT; firstly, the alveolar volume (VA) and secondly, the rate constant of CO removal from alveolar gas, the so called KCO, after taking the barometric pressure into account. Although it is generally accepted that DLCO/VA and KCO are physiologically equal, both are not interchangeable according to Hughes et al. (47). It is important to interpret DLCO/VA as the DLCO expressed per unit alveolar volume and not as the DLCO corrected for the alveolar volume, since DLCO/VA is not constant when VA changes. Therefore, DLCO/VA reflects the physiology more accurately than DLCO (47), since the very same DLCO measure can reflect different combinations of KCO and VA, each of these combinations reflect different pathophysiology (47-49).

Diffusing capacity is decreased by diseases such as emphysema, pulmonary fibrosis or pulmonary hypertension, and can be slightly increased in asthma or obesity (50). Since DLCO is also decreased in patients with anaemia, it is important to correct the measured DLCO values for blood haemoglobin levels.

Like impaired lung function measured by spirometry, a decreased diffusing capacity measured by the SBT has been associated with adverse health outcomes (51). In patients with severe emphysema, low diffusing capacity was associated with poor long-term survival (52) and it may even be a better predictor of mortality in COPD outpatients than spirometry measures (53).

**Figure 1 schematic representation of the single-breath DLCO breathing manoeuvre.**
DLCO: Diffusing capacity of the Lung for Carbon Monoxide (CO)
3. DETERMINANTS OF IMPAIRED LUNG FUNCTION

There are many determinants of impaired lung function which can be distinguished into environmental, life style and genetic factors.

3.1. Environmental factors
Tobacco smoke is considered as the strongest risk factor for impaired lung function and COPD. However, a significant proportion of individuals who never smoked develop COPD, suggesting that other factors than smoking should also be considered.

Half of worlds households and the majority of the rural households use coal and biomass fuel as their primary source of domestic energy (54). It is suggested that exposure to biomass smoke might be the most important risk factor for COPD, since more people are exposed to biomass smoke worldwide than to tobacco smoke. In addition, occupational exposures to noxious gases, mineral dusts, biological dusts and fumes are also associated with airway obstruction (55, 56). Finally, raised concentrations of outdoor air pollutants, especially in fast growing industrial countries, are strongly associated with COPD exacerbations and worsening of pre-existing COPD (57).

3.2. Smoking
Cigarette smoking is considered as the major preventable risk factor of respiratory morbidity and mortality in developed countries (58). The effect of smoking on lung development and function is pronounced in early life and during adulthood. Maternal smoking during pregnancy, for example, is associated with increased risk of wheezing and asthma in the offspring at adolescence (59).

It is believed that smoking exaggerates chronic inflammation in diseased lungs. Oxidative stress, genetic susceptibility and epigenetic modifications play an important role in this, by amplifying the inflammation that is initiated by smoking (60).

Smoking has a modifying role in lung function. Hunninghake and colleagues demonstrated that genetic variation in MMP12 (in smoking individuals) is associated with lung function. This effect was not found in non-smokers (61, 62). Another study by Xu et. al., also showed effect modification by smoking in the association between serum vitamin D and lung function (63). Furthermore, many interesting genetic and epigenetic loci were found for nicotine dependence (smoking behaviour), these pathways were also found in genome-wide and epigenome-wide association studies of lung function (64, 65). Whether there is genetic predisposition of nicotine dependence that influences lung function or whether we are dealing with pleiotropic effects remains to be investigated.
4. GENETIC EPIDEMIOLOGY

In genetic epidemiology, we study genetic and epigenetic determinants of health and disease in populations of unrelated individuals or in families. In the following paragraphs, we will summarize some important techniques used in genetic epidemiology.

4.1. Genome-wide association studies

A Genome-Wide Association Study (GWAS) is a hypothesis generating epidemiological study that aims to observe associations between Single Nucleotide Polymorphisms (SNPs) and traits of interest to understand the underlying biology and to ultimately improve treatment and prevention strategies (66). In the last decade, genetic studies such as GWAS and exome array studies (67), have been instrumental to identify genetic determinants of impaired lung function and COPD (68, 69). Those studies have also been successful in unravelling pathways and genes through which disease evolves in the body, such as the already consistently replicated loci hedgehog-interacting protein (HHIP), Family with sequence similarity 13 member A (FAM13A) and the smoking related nicotinic acetylcholine receptor CHRNA5 (67, 70-72) (Figure 2).

![Manhattan plot](image)

**Figure 2** Manhattan plot from the publication of Hobbs and colleagues (70) Gene names in gray are previously known COPD or lung function (FEV1 or FEV1/FVC) loci; novel gene loci are indicated in black.

4.2. Epigenome-wide association studies

Similar to GWAS, an epigenome-wide association study (EWAS) is a hypothesis generating study investigating how epigenetic variations such as DNA methylation, histone modifications and non-coding RNA activities can influence phenotypes and...
Among all forms of DNA methylation, methylation of cytosines in the context of cytosine–guanine dinucleotides (CpG) is the most prominent one. DNA methylation of cytosine is mapped through treatment of the genomic DNA with sodium bisulfite (74). The unmethylated cytosines will then convert to uracils, followed by thymidines that can be easily detected in a PCR, while the methylated cytosines remain unchanged (Figure 3). DNA methylation plays a key role in normal cellular function and homeostasis, whereas abnormal DNA methylation is associated with disease. So far, the role of epigenetic variation, particularly DNA methylation, has been mainly studied in the context of cancer (73, 75-77). EWAS on other complex diseases are also emerging and correlation between genotype and epigenotype (methQTL) favours the integrated approach between GWAS and EWAS in the future (73).

Figure 3 Detection of methylated regions after treatment with sodium bisulfite. DNA methylation of cytosine is mapped through treatment of the genomic DNA with sodium bisulfite (74). The unmethylated cytosines will then convert to uracils (UG), followed by thymidines (TG) that can be easily detected in a PCR, while the methylated cytosines remain unchanged (CG).

4.3. Genome-wide linkage analysis

GWAS studies have been functional to identify common genetic variants of disease. However, an emerging viewpoint suggests that rare variants -which are not well studied using GWAS- could explain a substantial part of the missing heritability (78, 79). Before, linkage analysis was the main tool to investigate rare genetic variance in complex traits with familial aggregation, until the focus was shifted to GWAS. Genome-wide linkage analysis is a powerful tool to localize disease genes on specific chromosomes. It relies on the theory that genes that are nearby each other are more linked and segregated together during meiosis compared to genes further away, and thereby links chromosomal fragments together among related individuals with trait similarity. In their most typical form, linkage studies have three steps. The first step is the identification of
the disease locus, next, sequence analysis is performed to define the rare deleterious variant, and finally, functional analysis follows to unravel the pathophysiology (80). With the increased use of whole genome sequencing (WGS), linkage studies have now more potential again to be used in identifying rare variants that cannot be captured in GWAS due to power issues (78-81).

5. AIMS AND OUTLINE OF THIS THESIS

The general aim of this thesis was to study the epidemiology of impaired lung function and COPD, to unravel its genetic and non-genetic determinants, and finally to study the impact of impaired lung function on adverse health outcomes, such as mortality.

Chapter 2: Epidemiology of COPD

In the first chapter of this thesis we studied the epidemiology of COPD. Here we focused on the prevalence and incidence of COPD in middle-aged and older subjects in the Rotterdam Study, with approximately 15,000 participants and a follow-up period of up to 25 years.

Chapter 3: Genetic epidemiology

In the chapters 3.1 and 3.2, we studied the heritability and the genetic determinants of pulmonary function tests; spirometry and diffusing capacity, respectively. In chapter 3.3, we investigated the association between occupational exposures on DNA methylation, and assessed the effect of occupational exposures on lung function in never smoking individuals via epigenetic mechanisms. In chapter 3.4, we also studied the epigenetic signature of diffusing capacity. Finally, in chapter 3.5, we studied rare variants that possibly affect COPD through linkage and exome sequence analyses. For the research work in this chapter, we collaborated with other large cohorts from The Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium (chapter 3.1), with colleagues from the Framingham Heart Study (chapter 3.2 and 3.4), Lifelines (chapter 3.3) and the Erasmus Rucphen Family (ERF) study (chapter 3.5).

Chapter 4: Clinical relevance

In the third chapter, we aimed to investigate the adverse effects of an impaired lung function on health outcomes. In chapter 4.1, we investigated the effect of an increased pulmonary artery to aorta ratio, as measured on CT scans of the thorax, on mortality in individuals with COPD, and explored the modifying effect of diffusing capacity in this association. Finally, in chapter 4.2, we investigated the association between COPD and the risk to develop peripheral artery disease (PAD) and mortality.
REFERENCES


42. Schunemann HJ, Dorn J, Grant BJB, Winkelstein W, Trevisan M. Pulmonary function is a long-term predictor of mortality in the general population - 29-year follow-up of the Buffalo Health Study. Chest. 2000;118(3):656-64.


69. Melen E, Guerra S. Recent advances in understanding lung function development. F1000Res. 2017;6:726.