

Emphysema on low-dose CT

assessment, implications,
and technical considerations

H.J. Wisselink

Emphysema on low-dose CT
assessment, implications, and technical considerations

H.J. Wisselink

Ik heb met heel mijn hart elke vorm van wijsheid onderzocht,
want ik wilde alles wat onder de hemel gebeurt doorgronden.
— Prediker 1:13, NBV

I sought out every kind of wisdom with all my heart,
for I wanted to understand all that happens under the heavens.
— Ecclesiastes 1:13

Initial markup and typesetting: L^AT_EX
Final markup and printing: Gildeprint
Cover design (draft): Lydia Bijker-Wisselink
Cover design (finalisation): Gildeprint

The printing of this thesis was financially supported by MathWorks[®].

© Copyright 2023, H.J. Wisselink

All rights reserved. The copyright of the articles that have been published has been transferred to the publishers of the respective journal. No part of this thesis may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without prior written permission by the author.



university of
 groningen

Emphysema on low-dose CT

assessment, implications, and technical considerations

PhD thesis

to obtain the degree of PhD at the
 University of Groningen
 on the authority of the
 Rector Magnificus Prof. J.M.A. Scherpen
 and in accordance with
 the decision by the College of Deans.

This thesis will be defended in public on

Wednesday 18 October 2023 at 11.00 hours

by

Hendrik Joost Wisselink

born on 18 June 1993

Supervisors

Prof. R. Vliegenthart
Prof. G.H. de Bock

Co-supervisors

Dr. M. Rook
Dr. M.A. Heuvelmans

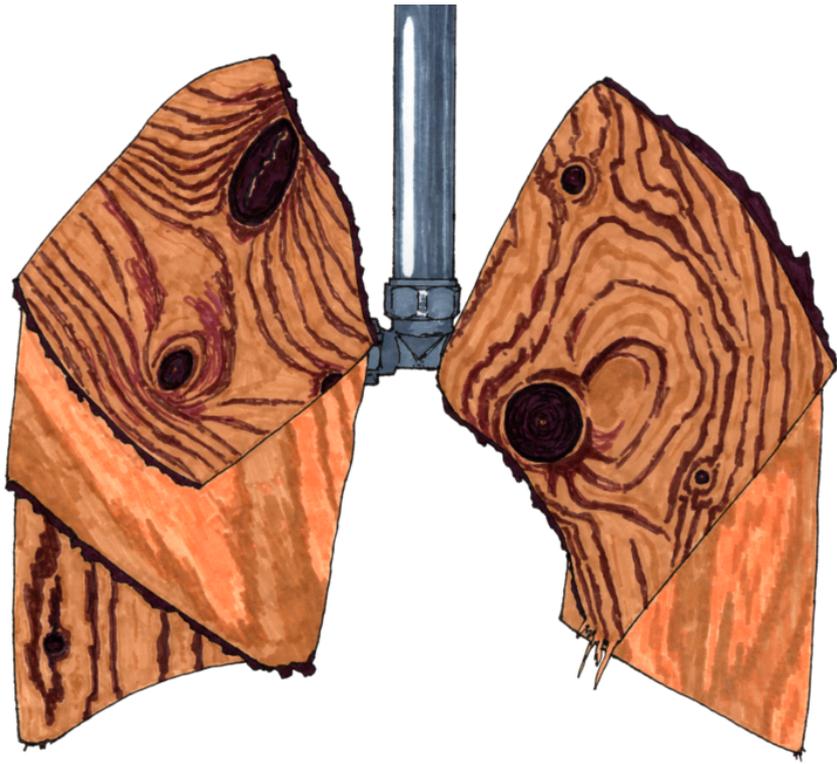
Assessment Committee

Prof. D.J. Slebos
Prof. P.A. de Jong
Prof. T. Flohr

Paranymphs

Ivan Dudurych

Daan L.B. van den Oever



Contents

Table of contents

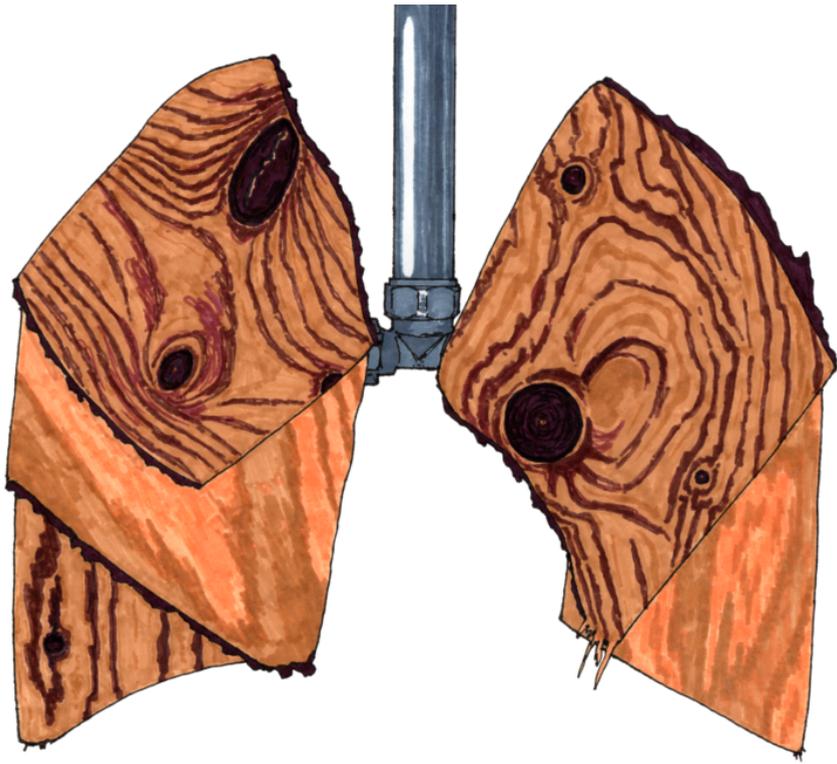
1	General introduction and outline	1
2	CT-defined emphysema prevalence in a Chinese and Dutch general population	11
3	Association between Chest CT–defined Emphysema and Lung Cancer: A Systematic Review and Meta-Analysis	29
4	Predicted versus CT-derived total lung volume in a general population: the ImaLife study	49
5	Potential for dose reduction in CT emphysema densitometry with post-scan noise reduction: a phantom study	65
6	Ultra-low-dose CT combined with noise reduction techniques for quantification of emphysema in COPD patients: An intra-individual comparison study with standard-dose CT	83
7	Improved precision of noise estimation in CT with a volume-based approach	99
8	CT-based emphysema characterisation per lobe: a proof of concept	111
9	General discussion	129
10	Bibliography	139
	Supplementary materials	182
	Summary (English)	233
	Samenvatting (Nederlands)	239
	Acknowledgements	247



A digital version of this thesis is available on
<https://hjwisselink.nl/PhD/thesis.pdf>.



The propositions are available on
<https://hjwisselink.nl/PhD/propositions.html>.



Chapter 1

General introduction and outline

1.1 General introduction

Big Three diseases Chronic obstructive pulmonary disease (COPD), lung cancer, and cardiovascular disease (CVD) are a major cause of mortality and morbidity worldwide. They are sometimes collectively referred to as the Big Three diseases or B3 [12]. Back in 2000, they were responsible for 36 % of the global mortality (COPD: 5.8 %, lung cancer: 2.4 %, CVD: 27.9 %), and the incidence of the B3 is increasing [13]. Two decades later, in 2019, the B3 diseases were responsible for 41 % of the global mortality, and this proportion is expected to increase further in the coming decades [13].

The B3 share risk factors, the most important of which are advanced age and smoking [14–16]. There is also mounting evidence that the presence of each of the B3 diseases may be an independent risk factor for one of the others [4, 17, 18]. It is however possible this association is (at least partly) due to insufficient correction for shared risk factors. In addition to often occurring together, the B3 mostly exist in the same anatomical region. COPD and lung cancer occur only in the lungs and the majority of CVD-caused mortality is due to thoracic disease [13]. A single computer tomography (CT) scan makes it possible to identify them and estimate their severity [12].

Early detection and screening Potentially, the B3 diseases are logical candidates to consider for screening programmes. For screening, several criteria must be met. For instance, the disease must pose a large health issue, there has to be an early stage or precursor that can be detected by screening, and there has to be a treatment for the disease once diagnosed. Furthermore, the costs of screening and treatment must be in balance with the benefits [19].

In the case of lung cancer, an early and treatable stage exist: in several screening studies, for instance, 59–73 % of screen-detected lung cancers were stage I, compared to 13 % of lung cancer detected in the non-screening group. CT screening is expected to substantially reduce lung cancer-specific mortality [20–22].

For CVD, studies are ongoing to determine the effectiveness (and cost-effectiveness) of screening [23, 24]. Preliminary results from the DANCAVAS trial shows screening for CVD is cost-effective [25]. The results from the ongoing ROBINSKA trial may be used to assess what the value of screening for CVD with CT scans (instead of a questionnaire) will be [26].

Whereas lung cancer and CVD may be potential candidate diseases for screening programmes, this is not (yet) the case for COPD alone. For COPD there currently exists no curative treatment, although removing underlying causes (e.g. smoking) can slow progression [27, 28]. Because smoking is the major modifiable risk factor for the B3, smoking cessation merits special attention. A study from the UK showed that participants in a lung cancer screening trial are more likely to stop smoking, which could also apply to collective B3 screening [29].

Lung cancer screening is currently being implemented in several countries, screen-

ing for CVD is under consideration, while screening for COPD alone is not expected in the foreseeable future [25, 30]. Possibly, there may be a future role for combined B3 screening, even if the main benefit of COPD screening would be to improve risk assessment for lung cancer and CVD. This calls for reliable detection of COPD.

Because the B3 diseases are interconnected, implementation of lung cancer screening or screening for CVD can be used to implement a multi-disease screening (i.e. screening for all three), increasing cost-effectiveness [31].

In addition to these screening CTs, thoracic CT scans are also often acquired in regular clinical care. In 2019, approximately 600 000 thoracic CT scans were performed in the Netherlands, accounting for 30 % of CT scans [32]. These CT scans may have indications other than B3 diseases, providing opportunistic screening (i.e., early disease might be detected as a side-benefit). This high volume of thoracic CT scans demands effective and efficient ways to assess the presence and severity of each of the B3 diseases.

COPD This thesis will primarily focus on parameters measurable on CT as related to COPD. COPD is a compound disease, consisting of bronchitis (and/or bronchiolitis) and pulmonary emphysema [33]. A schematic overview of the relevant anatomy is shown in Figure 1. As shown in this figure, emphysema is the destruction of the alveoli (i.e. air sacs), which causes the gas-exchanging surface area to decrease. Bronchitis and bronchiolitis, by contrast, do not affect the alveoli but causes the walls of the airways to become thicker, which impedes good airflow. Bronchitis is the inflammation of the larger airways (i.e. the bronchi), while bronchiolitis is the inflammation of the smaller airways (i.e. the bronchioles). The relative severities of bronchitis and emphysema can be different for each patient, and may have different progression rates [33]. The specific COPD phenotype has an impact on the disease burden, symptoms, treatment options, and prognosis [36–38].

Emphysema quantification While the gold standard to diagnose COPD currently is a pulmonary function test (PFT), the role for CT in diagnosing COPD is ever-increasing [39, 40]. This is because the effects of airway wall changes and alveolar destruction are readily visible on CT scans [40, 41]. Density on CT is measured in Hounsfield Units, where -1000 HU represents air or vacuum and 0 HU represents water. Since the intra-alveolar septa are destroyed in patients with emphysema, affected tissue has a lower density than healthy tissue, which is visible on CT. Because of this lower density, emphysema can be quantified with densitometry by counting the number of voxels with a density below a certain threshold [42]. The COPD-specific imaging biomarkers derived from CT (e.g. the percentage of voxels with a density below -950 HU) correlate well with PFT

results and with outcomes like mortality [42, 43]. It is also possible to assess the presence and severity of CT-based emphysema visually, which provides additional prognostic value [40, 43].

Emphysema qualification Visual CT review allows the determination of the emphysema subtype. There are currently three recognised subtypes of emphysema [40, 44]. The most common subtype is centrilobular emphysema. This subtype is characterised by loss of the tissue centrally in the acini. While lesions can coalesce into larger areas, in early disease, lesions are generally more or less scattered throughout the lobes, with an apical predominance. Paraseptal emphysema is characterised by lesions along the septa and fissures, and tends to be upper-lobe predominant. While it has less impact on pulmonary function than centrilobular emphysema, it may negatively impact the eligibility for treatments like endobronchial lung volume reduction [37]. The third subtype is panlobular emphysema, which is characterised by generalised destruction of lung parenchyma classically due to a genetic defect called alpha-1 antitrypsin deficiency (A1AD), although other causes like Ritalin abuse do exist [45]. Panlobular emphysema due to A1AD is usually lower lobe predominant.

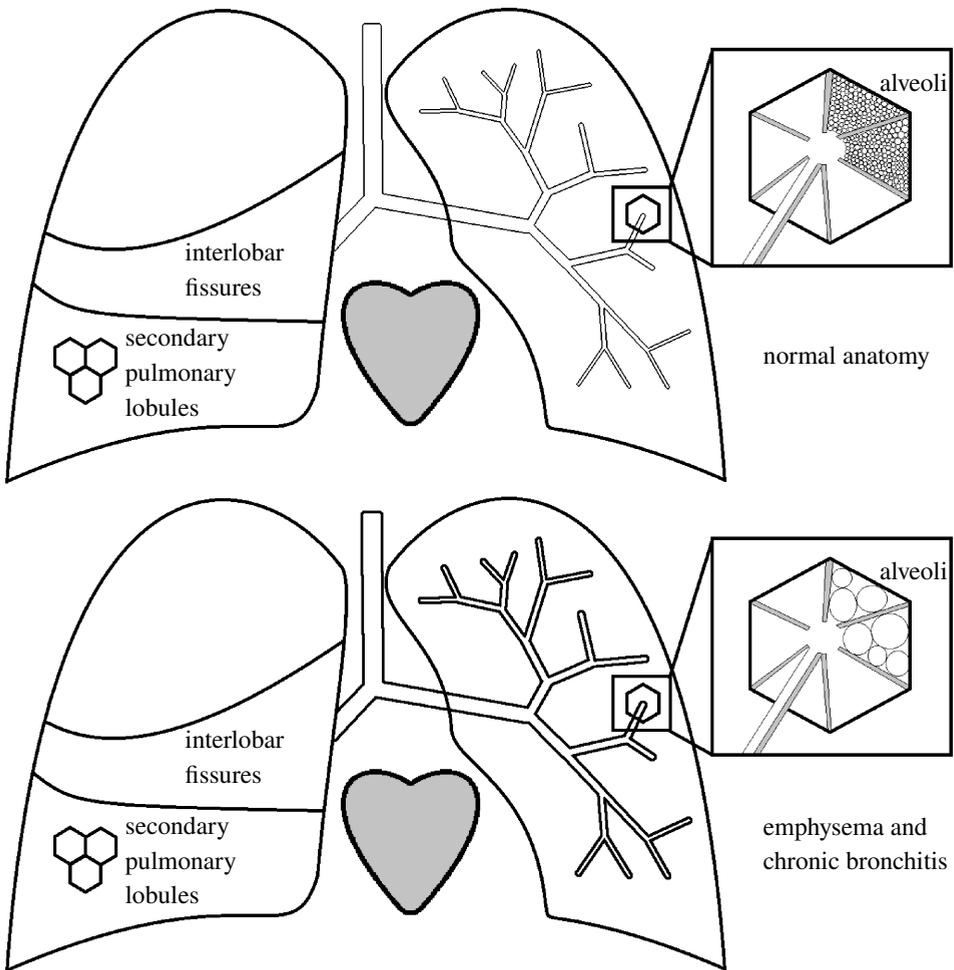


Figure 1: schematic anatomical overview

The lungs are divided into separate lobes by the interlobar fissures. Each lobe is further subdivided into segments, primary pulmonary lobules, secondary pulmonary lobules, and acini, until the airways reach the smallest pulmonary units: the alveoli. The alveoli provide most of the surface area where air and blood can exchange gases: 50 – 100 m². In the actual pulmonary anatomy, there are approximately 16–23 generations (splits) between trachea and alveoli. In COPD the airway walls thicken and the inter-alveolar walls are destroyed. [34, 35]

1.2 Outline

The three subtypes can be visually distinguished. The different aetiology and clinical outcomes of each emphysema subtype are an active topic of research [46]. The Dutch-Chinese NELCIN B3 project includes general population-based cohort studies, in which participants underwent non-contrast low-dose chest CT imaging [10, 47]. In **chapter 2**, CT scans from 2343 participants (1200 participants from ImaLife, and 1143 participants from the Chinese NELCIN B3 cohort) are visually assessed. The visually assessed emphysema prevalence, subtype, and severity are correlated to demographic data and environmental exposures including smoking, aiming to explore the risk factors.

As mentioned previously, cigarette smoking is a risk factor for many diseases, including COPD, lung cancer, and CVD. In addition to this, there are indications that the presence of one disease may be a risk factor for other diseases apart from shared risk factors [17, 18]. In **chapter 3**, a meta-analysis is presented, in which the association between emphysema (visual and quantitative density-based analysis) and lung cancer is explored.

In patient care for COPD, as well as other pulmonary diseases, it is important to establish the pulmonary function [41, 48]. It is important to express absolute measurements as a percentage of predicted, because an 80-year-old healthy woman of 1.60 m is expected to have a much smaller lung volume than a 2.10 m tall 50-year-old male COPD patient with hyperinflation. The expected PFT-result for a comparable healthy person is derived from models that are generally based on age, sex, and height [48]. This presumes the expected value can reliably be predicted. In **chapter 4**, the lung volume is measured on CT scans from a general population-based sample and the measured volumes are compared to the predicted volumes. The aim of this chapter is to determine how well the predicted lung volume matches the measured lung volume.

Because CT scanners use potentially harmful X-rays, there are ongoing efforts to reduce the radiation dose. However, since electrical noise in the detector is approximately constant, reducing the radiation causes a reduction in the signal to noise ratio (SNR), resulting in a ‘snowy’ image appearance which impedes accurate interpretation of the images. To accurately assess CT scans, it is imperative to maintain sufficient image quality. Therefore, it is important to optimise the acquisition and reconstruction parameters to minimise required radiation, while limiting the amount of image noise and maximising the useful information. Because noise changes the apparent density of the tissue, this potentially changes the results of density-based quantitative analysis, as used for the quantification of pulmonary emphysema. The aim of **chapter 5** is to determine which combination of acquisition parameters, reconstruction settings, and noise reduction options would result in a scan that has an acceptable quality to perform density-based quantification of

emphysema. To do so, a total of 384 parameter combinations are used to acquire CT scans of the COPDGene phantom. From this scan data, the effect of each parameter on quantitative accuracy is assessed.

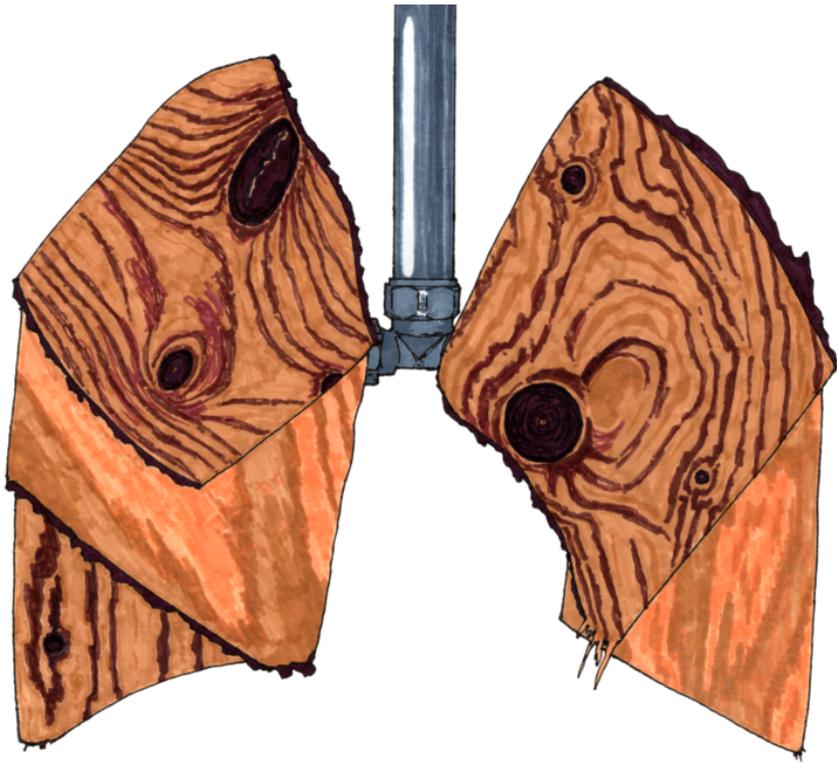
Because the results in chapter 5 are based on a phantom study, further validation is required to confirm the findings in CT scans of humans. For **chapter 6**, we use data from forty-nine COPD patients who underwent CT scans with the clinical standard protocol (SDCT), as well as an ultra-low-dose CT (ULDCT). The same noise reduction methods from chapter 5 are applied to the ULDCT scans, resulting in seven denoised CT scans in addition to the unprocessed ULDCT and SDCT. Emphysema is quantified for each patient on all nine reconstructions. The results from the SDCT are considered the reference standard. The aim is to determine which noise reduction setting best matches the SDCT result.

For many quantitative analyses, including those in chapters 5 and 6, it is important to determine the magnitude of noise in a medical image. The most common method to measure noise magnitude on CT scans is to measure the standard deviation of the density of pixels or voxels in a circular region of interest (ROI). This is generally performed in either air, blood, or a region of tissue with a reasonably homogeneous density [49]. While there are very sophisticated mathematically complex methods to reliably estimate image noise, these are generally cumbersome to apply, making them unsuitable for clinical practice and undesirable for research. Using a spherical volume of interest (VOI) instead of a circular ROI drastically increases the number of included voxels, which should reduce inter-measurement variability. Such a change could be implemented in clinical practice and in research, without increasing work or requiring specialised separate software. In **chapter 7** the effect of measuring a spherical VOI instead of a circular ROI is explored. In this chapter the noise is measured in the air in the trachea and main bronchi. The results of the ROI and VOI-based measurements from SDCT and ULDCT are compared to a segmentation-based ground truth.

Another strategy to improve the assessment of emphysema on CT is presented in a proof-of-concept study in **chapter 8**. The Fleischner criteria are the current standard for visual classification of emphysema on CT [40]. For centrilobular emphysema it defines five categories, for paraseptal emphysema it defines two categories, and panlobular is only defined as a yes/no classification. Because of this low granularity, there is a high degree of variability in emphysema severity and distribution within the lungs in groups with the same classification. In this chapter the Fleischner criteria are expanded by adding more categories for paraseptal and panlobular emphysema to homogenise the number of categories. Additionally, the extended classification is applied to each lobe separately, allowing the computation of the emphysema sum score. This sum score is intended to provide a more granular description of the overall emphysema severity. Such a granular description allows distinguishing cases with potentially clinically relevant differences in

severity, which might have the same severity classification.

In summary, this thesis will explore CT-defined emphysema. First the disease burden will be assessed, then some technical considerations will be discussed, and finally strategies for improvement of measurements will be presented.



Chapter 2

CT-defined emphysema prevalence in a Chinese and Dutch general population

Submitted.

*Xiaofei Yang, Yihui Du, Hendrik Joost Wisselink,
Yingru Zhao, Marjolein A. Heuvelmans,
Harry J.M. Groen, Monique D. Dorrius,
Marleen Vonder, Zhaoxiang Ye,
Rozemarijn Vliegenthart, Geertruida H. de Bock*

Abstract

Background The prevalence of and risk factors for CT-defined emphysema are poorly defined and may vary among populations. This study determines and compares the prevalence, subtypes, severity, and risk factors for emphysema assessed by low-dose CT (LDCT) in Chinese and Dutch general populations.

Methods We included LDCT scans of 1143 participants from a Chinese lung cancer screening study and 1200 participants from a Dutch population-based study. An experienced radiologist visually assessed the scans for emphysema presence (\geq trace), subtype, and severity. Logistic regression analyses, overall and stratified by smoking status, were performed and adjusted for fume exposure, demographic and smoking data.

Results The Chinese population had a comparable proportion of women to the Dutch population (54.9 % vs 58.9 %), was slightly older (61.7 ± 6.3 vs 59.8 ± 8.1), included more never-smokers (66.4 % vs 38.3 %), had a higher emphysema prevalence ([58.8 % vs 39.7 %], adjusted odds ratio [aOR] 2.06, 95 % CI 1.68 – 2.53), and more often had centrilobular emphysema (54.8 % vs 32.8 %, $p < 0.0001$), but no differences in emphysema severity. After stratification, only in never-smokers an increased odds of emphysema was observed in the Chinese compared to the Dutch (aOR 2.62, 95 % CI 1.99 – 3.45). Never-smokers in both populations shared older age (aOR 1.59, 95 % CI 1.25 – 2.02 vs 1.26, 95 % CI 0.97 – 1.64) and male sex (aOR 1.50, 95 % CI 1.02 – 2.22 vs 1.93, 95 % CI 1.26 – 2.96) as risk factors for emphysema.

Conclusions Never-smokers had a higher prevalence of mainly centrilobular emphysema in the Chinese general population compared to the Dutch after adjusting for confounders, indicating that factors other than smoking, age and sex contribute to presence of emphysema.

2.1 Introduction

Chronic obstructive pulmonary disease (COPD) is a common and progressive respiratory disorder that places an immense burden on health care systems [50]. Emphysema is an important phenotype that manifests as parenchymal destruction [51, 52]. A study by Steiger et al. showed that 76.5 % of people with CT-defined emphysema had no prior diagnosis of COPD despite 23.6 % having moderate or severe disease [53]. Visual emphysema on CT independently increases the risk of lung cancer and all-cause mortality [4, 43]. If we are to develop effective national health policies for the evidence-based deployment of finite health care resources to correct both the under-diagnosis of emphysema and associated risk of lung cancer, we urgently need to clarify the epidemiology and causes of CT-defined emphysema [54].

Several studies have reported the CT-defined prevalence of emphysema, with variations from 38.0 % in Poland to 60.1 % in the United States that result from differences in diagnostic strategies (e.g., scanning protocol and evaluation guideline) and risk exposures [43, 55]. To evaluate and compare the prevalence of emphysema between areas, we therefore need studies that use the same diagnostic strategies and assess risk factors in a similar way. However, potential risk factors differ between countries, with notable differences in smoking rates, outdoor air pollution, and cooking-related household air pollution between Asia and western countries [54, 56]. Much is known about the prevalence and risk factors for lung function-defined COPD [15, 57, 58]; however, little is known about the prevalence of, and the factors that contribute to, CT-defined emphysema in general populations (e.g., the similarities and differences between Asian and western populations). As part of the Netherlands and China Big 3 diseases (NELCIN B3) project, which was initiated for early detection of lung cancer, COPD and cardiovascular disease, international comparison of the epidemiological features of emphysema and associated risk factors will help to inform strategies for disease prevention and therapy development [47].

The aim is to determine and compare the prevalence, subtypes, and severity of emphysema assessed by low-dose CT (LDCT) between Chinese and Dutch general populations and to explore the related risk factors.

2.2 Methods

Study design, study population, and eligibility

This study included a sample of Chinese participants from the NELCIN B3 study and Dutch participants from the Imaging in Lifelines (ImaLife) study [10, 47]. These prospective studies were designed to find early imaging biomarkers for the “big three” thoracic diseases (i.e., COPD, coronary artery disease, and lung can-

cer).

As part of the NELCIN B3 study, 4000 participants were invited from the general population and underwent LDCT lung cancer screening at Tianjin Medical University Cancer Institute and Hospital, China if they met the following inclusion criteria: any smoking status, age 40 – 74 years, resident in Tianjin city for at least 3 years, and no self-reported history of any malignant tumour [47]. The ImaLife study comprised a subset from the Lifelines study, a cohort study in the northern Netherlands [59]. The ImaLife study included 12 000 participants with an age ≥ 45 years, and after completing lung function tests in the second assessment, having an LDCT scan at University Medical Center Groningen (UMCG). The Ethics Committee of Biomedicine Research of the Second Military Medical University and of the UMCG approved the NELCIN B3 study (registration number: NCT03992833) and the ImaLife study (registration number: NL58592.042.16), respectively. Participants in both cohorts provided written informed consent.

The current analysis included a consecutive series of participants aged 45 – 74 years who underwent LDCT between May and October 2017 in the NELCIN B3 study ($n=1143$). An approximately matched number of participants with the same age range who underwent the LDCT between June and October 2019 in the ImaLife study were also included ($n=1200$, Figure 1). We excluded participants if they had interstitial fibrosis, pneumothorax, and/or incomplete data (i.e., missing demographic data or CT scans). For this study, the outcome of interest was visually assessed emphysema on LDCT scan. Participants were classified as having either no emphysema or at least trace emphysema.

Data collection and definitions

In the two prospective cohorts, trained interviewers conducted structured face-to-face interviews using questionnaires. They gathered information about exposure to smoking (i.e., smoking status and passive smoking), demographics (i.e., age, sex, body mass index [BMI], and educational level), and exposure to either cooking fumes or fireplace fumes (see Table S2.1 [p. 183] for definitions). The educational level was categorised into low, moderate, and high [60, 61]. BMI was categorised into < 25 and ≥ 25 kg/cm². The cohorts differed slightly in the definitions of smoking status, passive smoking, and cooking/fireplace fume exposure [47, 62]. Smoking status collected in the second-round assessment (2014 – 2016) in the Dutch cohort was used, since the data collection was close to the CT scan acquisition (2019), and was supplemented with data from the nearest previous round of assessment in case of missing smoking data. Educational level, passive smoking, and fireplace fume exposure were based on the baseline data collection (2007 – 2013). We used the age and BMI recorded at the CT scan acquisition (2019). The Chinese cohort had no interval between collecting the demographic data and acquiring the CT scan, whereas the demographic data in the Dutch cohort were collected before

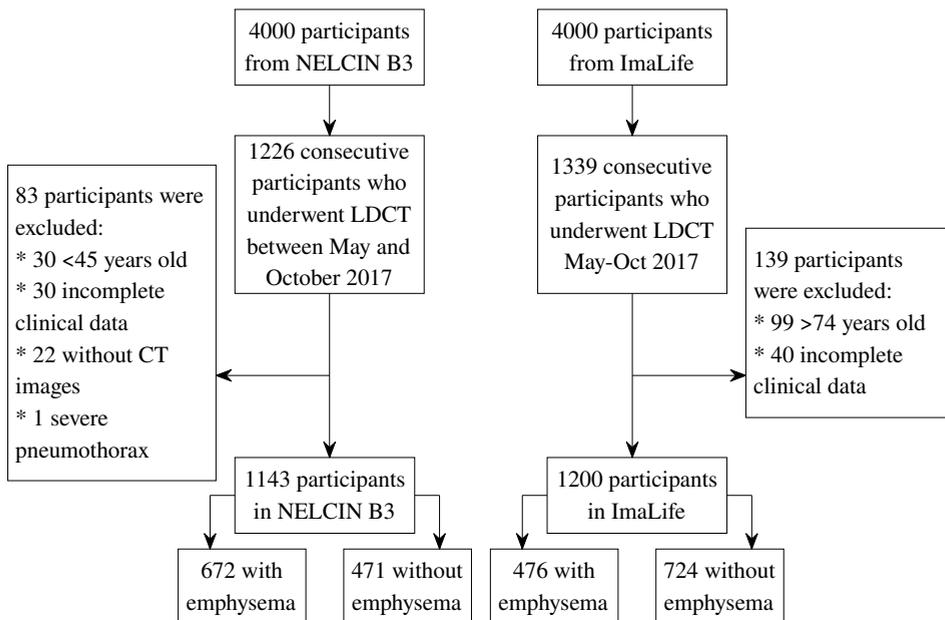


Figure 1: Flowchart of study design
LDCT: low-dose CT.

the CT scan acquisition with a gap of several years (10 ± 10 years for passive smoking and fireplace fume exposure; 5 ± 1 years for smoking status).

CT scan acquisition

The Chinese study used a 64-detector row CT system (SOMATOM Definition AS 64, Siemens Healthineers, Germany) for the non-contrast LDCT chest examinations, with the following parameters: 120 kVp, 35 mAs (reference), and pitch 1.0. Reconstruction kernel D45f was applied to reconstruct the images at 1.0 mm thickness and 0.7 mm increment. All participants were scanned head first in the supine position during an inspiratory breath hold.

The Dutch study used a third-generation dual-source CT (SOMATOM Force, Siemens Healthineers, Germany) for the non-contrast LDCT chest examinations, with the following parameters: 120 kVp, 20 mAs (reference), and pitch 2.5. Reconstruction kernel Br40 was applied to reconstruct the images at 1.0 mm thickness and 0.7 mm increment. All participants were scanned head first in the supine position during an inspiratory breath hold.

CT image quality for the Chinese and Dutch study was assessed and compared based on 50 randomly selected cases. Any systematic bias in depicting air was quantified by measuring the mean HU in the trachea. Noise levels were quantified by measuring the HU standard deviation of regions of interest (ROI) on 1.0 mm

and 10.0 mm thickness. The ROI with a fixed area (50 mm²) was placed in the tracheal lumen at the level of 1 cm and 2 cm above carina of the trachea.

Visual emphysema assessment

One radiologist (XY) with 6 years' experience visually assessed emphysema on Chest CT for all Chinese and Dutch participants, using a standard protocol created by the Fleischner society [40]. Interobserver agreement was determined based on 100 randomly selected cases in each cohort by a second radiologist (ZN) with 3 years' experience for the Chinese participants and a clinical physician (HJW) with 4 years' experience for the Dutch participants. All readers received training before the assessment and used a standardised protocol. All readers performed the visual emphysema assessments using the minimum intensity projection in version VB30A of the Syngo.via software suite (Siemens Healthineers, Germany). They used a 10 mm thickness (WC:−850 HU, WW:400 HU) and multiplanar reconstruction with 1 mm thickness (WC:−750 HU, WW:700 HU) based on the D45f kernel or the Br40 kernel CT images [43, 63].

Emphysema (low attenuation areas or lucencies) was scored according to the Fleischner criteria [40]. If present (\geq trace), emphysema was further categorised as one of the three predominant subtypes, centrilobular (CLE), paraseptal (PSE), and panlobular (PLE). The predominant subtype was noted by the most severe one in cases of mixed emphysema. CLE was classified as trace ($<0.5\%$), mild ($0.5 - 5\%$), moderate ($>5\%$), confluent and advanced destructive. PSE was classified as mild (< 1 cm lucencies) or substantial (mainly > 1 cm lucencies).

Statistical analysis

We described continuous variables as means and standard deviations and categorical variables as frequencies and percentages. Kappa statistics for emphysema and weighted kappa coefficients for CLE and PSE severity were calculated to assess interobserver agreement. To compare emphysema prevalence between the two populations, we performed univariate and multivariable logistic regression analyses to estimate the odds ratios (ORs) and 95 % confidence intervals (95 % CIs). In the multivariable analysis, we adjusted for age (per 10-year increase), sex, smoking status, passive smoking, BMI, educational level, and cooking/fireplace fume exposure. In addition, we performed analyses stratified by smoking status and by cohorts. Chi-squared tests were conducted to analyse differences in emphysema subtype and severity between the two populations with emphysema. To assess the robustness, we performed a sensitivity analysis by repeating the main analysis for the main subtype of emphysema (CLE) by limiting the emphysema threshold to 'at least trace' or 'at least mild'. All analyses were conducted using the SPSS Version 28.0 (IBM, Armonk, NY, USA) with an extension of "STATS_WEIGHTED_KAPPA", treating $p < 0.05$ as statistically significant. As

there are limited data available regarding the prevalence of CT-defined emphysema in literature, we could not perform a prior sample size estimation. As an alternative, a post hoc power calculation was performed using G power Version 3.1.9 (Heinrich Heine University Düsseldorf, Germany).

2.3 Results

Population characteristics

We included 2343 participants in this analysis (Figure 1), comprising 1143 Chinese participants and 1200 Dutch participants, with comparable proportions of women (627 [54.9 %] vs 707 [58.9 %], respectively; $p = 0.0472$). Compared with the Dutch, the Chinese population was slightly older (61.7 ± 6.3 vs 59.8 ± 8.1 , $p < 0.0001$) and included more never-smokers (759 [66.4 %] vs 459 [38.3 %], $p < 0.0001$) (Table 1). As shown in Table S2.2 [p. 184], among the never-smokers, the Chinese participants were also older (61.1 ± 6.5 vs 58.1 ± 8.5 , $p < 0.0001$) and had lower BMI ($< 25 \text{ kg/m}^2$, 57.7 % vs 44.2 %, $p < 0.0001$). The prevalence of passive smoking exposure was higher in the overall Chinese participants (44.0 % vs 22.6 %, $p < 0.0001$) and never-smokers (35.4 % vs 15.5 %, $p < 0.0001$, Table S2.2 [p. 184]) than in the Dutch. No difference was observed between Chinese and Dutch participants in cooking/fireplace fume exposure (6.7 % vs 6.1 %, $p = 0.5184$).

CT image quality and interobserver agreement

Regarding systematic bias, the mean density of air in the trachea was -987 ± 7 HU (13 HU higher than the theoretical density) and -970 ± 5 HU (30 HU higher than the theoretical density) for the Chinese and Dutch cohort respectively based on 1.0 mm slice thickness; this was -1015 ± 2 HU (a difference of 15 HU) and -1009 ± 8 HU (a difference of 9 HU) for 10 mm. Regarding image noise, the mean standard deviation of the ROI for air in the trachea was 31.8 ± 6.1 HU and 23.4 ± 4.6 HU for respectively the Chinese and Dutch cohorts based on 1.0 mm slice thickness. The SD was 10.9 ± 1.9 and 7.0 ± 1.3 for 10 mm.

Agreement between readers when assessing emphysema was good in both the Chinese participants (κ 0.76, 95 % CI 0.63 – 0.89) and the Dutch participants (κ 0.87, 95 % CI 0.76 – 0.97). Similarly, the agreement was good for the severity of CLE (κ_w 0.77, 95 % CI 0.67 – 0.88) and PSE (κ_w 0.77, 95 % CI 0.58 – 0.96) in the Chinese participants, and was comparable for the severity of CLE (κ_w 0.87, 95 % CI 0.78 – 0.96) and PSE (κ_w 0.84, 95 % CI 0.66 – 1.00) in the Dutch participants.

Table 1: Characteristics of participants (overall and those with emphysema \geq trace) in the Chinese and Dutch cohorts.
Values are either N (%), or (mean \pm SD)

Characteristics	Chinese cohort		Dutch cohort	
	Total	With emphysema	Total	With emphysema
Participants	1143 (48.8)	672 (58.8)	1200 (51.2)	476 (39.7)
Age	61.7 \pm 6.3	62.7 \pm 6.1	59.8 \pm 8.1	61.0 \pm 7.7
Sex				
Women	627 (54.9)	302 (44.9)	707 (58.9)	239 (50.2)
Men	516 (45.1)	370 (55.1)	493 (41.1)	237 (49.8)
Smoking status				
Never	759 (66.4)	383 (57.0)	459 (38.3)	127 (26.7)
Former	115 (10.1)	84 (12.5)	571 (47.6)	245 (51.5)
Quit years	11.9 \pm 10.8	12.4 \pm 11.6	20.6 \pm 12.2	20.0 \pm 12.2
Pack-years	22.5 \pm 19.2	23.6 \pm 19.9	10.3 \pm 9.8	12.9 \pm 11.5
Current	269 (23.5)	205 (30.5)	170 (14.2)	104 (21.8)
Pack-years	25.2 \pm 17.7	27.2 \pm 18.5	19.9 \pm 12.3	22.0 \pm 12.8
Passive Smoking				
No	640 (56.0)	361 (53.7)	929 (77.4)	343 (72.1)
Yes	503 (44.0)	311 (46.3)	271 (22.6)	133 (27.9)
BMI (kg/m ²)				
<25	643 (56.3)	398 (59.2)	473 (39.4)	195 (41.0)
\geq 25	500 (43.7)	274 (40.8)	727 (60.6)	281 (59.0)
Educational level				
Low	431 (37.7)	278 (41.4)	242 (20.2)	111 (23.3)
Moderate	418 (36.6)	224 (33.3)	615 (51.2)	232 (48.7)
High	294 (25.7)	170 (25.3)	343 (28.6)	133 (27.9)
Cooking or fireplace fume				
No	1066 (93.3)	208 (88.5)	1127 (93.9)	179 (95.2)
Yes	77 (6.7)	27 (11.5)	73 (6.1)	9 (4.8)

BMI: body mass index; SD: standard deviation.

Prevalence, subtype, and severity of emphysema

Emphysema (at least trace) was present in 672 (58.8 %) Chinese and in 476 (39.7 %) Dutch participants. The prevalence of trace, mild, and moderate, confluent-advanced CLE in the Chinese population was 38.2 %, 11.5 %, 2.8 %, and 2.2 %, respectively; by contrast, the prevalence was lower in the Dutch population for the severity levels (24.0 %, 6.3 %, 1.9 % and 1.0 %, respectively; overall $p < 0.0001$). The prevalence of emphysema (trace or above) in Chinese current, former and never smokers was 76.2 %, 73.0 % and 50.5 %, respectively; the corresponding prevalence in Dutch participants was 61.2 %, 42.9 % and 27.7 %, respectively. CLE was the most common subtype in participants with emphysema in each cohort (93.2 % and 82.6 %, respectively), followed by PSE (6.8 % and 17.4 %, respectively), and none were classified with PLE. Among those with emphysema, the proportion of CLE was higher in the Chinese than in the Dutch participants (93.2 % vs 82.6 %, $p < 0.0001$) and the severities of CLE or PSE were comparable (Table 2). When limiting the emphysema threshold to at least mild, emphysema prevalence (20.6 % vs 15.7 %, $p < 0.0021$), and the proportion of CLE (80.4 % vs 55.9 %, $p < 0.0001$) in the Chinese was still significantly higher than in the Dutch but no difference was observed for the distribution of severity of CLE or PSE (Table S2.3 [p. 185]).

Risk factors for CT-defined emphysema

Chinese versus Dutch cohort

Participants in the Chinese cohort had two-fold increased odds of emphysema after adjusting for covariates, with an adjusted OR of 2.06 (95 % CI 1.68 – 2.53) compared to the Dutch cohort (Table 3). After stratification by smoking status, this was only observed in never-smokers (2.62, 95 % CI 1.99 – 3.45; $p < 0.0001$), and not in current smokers (aOR 1.10, 95 % CI 0.63 – 1.90; $p = 0.7459$) or former smokers (aOR 1.58, 95 % CI 0.94 – 2.67; $p = 0.0858$) (Table 4). Meanwhile, the Chinese participants also had higher odds for CLE (Table S2.4 [p. 186] and Table S2.5 [p. 187]) than the Dutch, and after stratification by smoking status, still only Chinese never-smokers had the increased odds (Table S2.6 [p. 188] and Table S2.7 [p. 189]) regardless of using the threshold “at least trace” or “at least mild” for emphysema.

Chinese and Dutch cohort

Overall, when combining participants from both cohorts, participants with emphysema were typically older (aOR 1.46 per 10 years of age increase, 95 % CI 1.29 – 1.66), male (aOR 1.59, 95 % CI 1.32 – 1.93), and current smokers (aOR 2.78, 95 % CI 2.13 – 3.64) or former smokers (aOR 1.58, 95 % CI 1.26 – 1.99) compared to participants without emphysema; they also had lower BMI (aOR 0.73, 95 % CI 0.61 – 0.87 for BMI ≥ 25 kg/m²). We found no evidence for an associa-

Table 2: Distribution of subtype and severity of emphysema (\geq trace) in participants with emphysema in the Chinese and Dutch cohorts.

	Chinese Cohort n=672, n (%)	Dutch Cohort n=476, n (%)	p-value
Predominant subtype of emphysema			<0.0001 [#]
CLE	626 (93.2)	393 (82.6)	
PSE	46 (6.8)	83 (17.4)	
Severity of CLE			0.4666 [#]
Trace	437 (69.8)	288 (73.3)	
Mild	132 (21.1)	70 (17.8)	
Moderate	32 (5.1)	23 (5.9)	
Confl-Adv	25 (4.0)	12 (3.1)	
Severity of PSE			1.000 [§]
Mild	44 (95.7)	79 (95.2)	
Substantial	2 (4.3)	4 (4.8)	

CLE: centrilobular emphysema; PSE: paraseptal emphysema;

Confl-Adv: confluent or advanced destructive emphysema.

[#] Based on Chi-square testing;

[§] Based on Fisher's Exact Testing.

tion with emphysema for cooking/fireplace fume (aOR 1.31, 95 % CI 0.91 – 1.89, $p = 0.1539$), passive smoking (aOR 1.18, 95 % CI 0.97 – 1.44, $p = 0.0966$) or educational level (Overall $p = 0.1328$; Table 3). When limiting the emphysema threshold to mild or above (Table S2.5 [p. 187]), the risk factors associated with CLE remained the same.

Never-smokers by cohort

After stratifying never-smokers by national cohort, increasing age (aOR 1.59, 95 % CI 1.25 – 2.02 vs 1.26, 95 % CI 0.97 – 1.64 [$p = 0.0812$, per 10 year increase]) and male sex (aOR 1.50, 95 % CI 1.02 – 2.22 vs 1.93, 95 % CI 1.26 – 2.96) were associated with increased odds of emphysema with comparable magnitudes in the Chinese and Dutch participants (Figure 2). The aOR was increased for cooking/fireplace fumes exposure in both cohorts in never-smokers, but this was not significant. Likewise, passive smoking was not associated with emphysema in never-smokers in any of the two populations despite the high passive smoking prevalence in the Chinese (35.4 % vs 15.5 %).

We included 1143 Chinese participants and 1200 Dutch participants. In a post hoc power analysis, the power to detect a difference in emphysema prevalence (at least mild) between the two countries is 0.88 when two tails and an alpha value of 0.05 were applied.

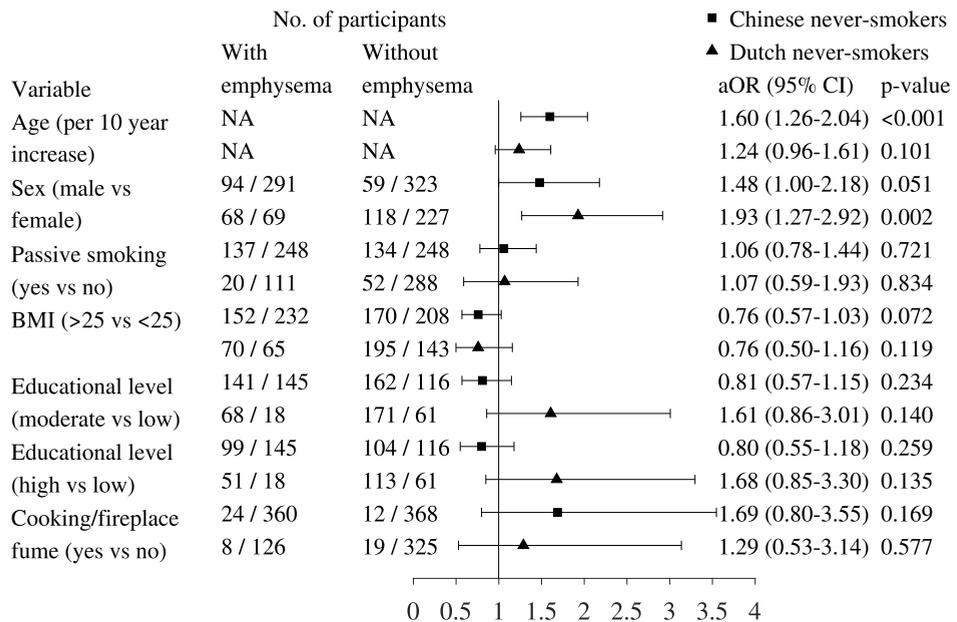


Figure 2: Multivariable logistic regression analysis of risk factors for emphysema (\geq trace) in never-smokers, stratified by national cohort.

BMI: body mass index; NA: not applicable; aOR: adjusted odds ratio; 95% CI: 95 % confidence interval. * $p < 0.05$.

Table 3: Associations between participant characteristics and emphysema (\geq trace).

Variables	Univariate logistical regression			Multivariable logistical regression		
	OR	95% CI	p-value	Adjusted OR	95% CI	p-value
Dutch cohort	1			1		
Chinese cohort	2.17	1.84 – 2.56	<0.0001*	2.06	1.68 – 2.53	<0.0001*
Age (per 10 years increase)	1.63	1.45 – 1.83	<0.0001*	1.46	1.29 – 1.66	<0.0001*
Female sex	1			1		
Male sex	2.21	1.87 – 2.62	<0.0001*	1.59	1.32 – 1.93	<0.0001*
Smoking status						
Never	1		<0.0001*	1		<0.0001*
Former	1.28	1.06 – 1.54	0.0103*	1.58	1.26 – 1.99	<0.0001*
Current	3.30	2.61 – 4.17	<0.0001*	2.78	2.13 – 3.64	<0.0001*
Control	1			1		
Passive smoking	1.65	1.39 – 1.97	<0.0001*	1.18	0.97 – 1.44	0.0966
BMI < 25 kg/m ²	1			1		
BMI \geq 25 kg/m ²	0.73	0.62 – 0.86	0.0001*	0.73	0.61 – 0.87	<0.0004*
Educational level						
Low	1		<0.0001*	1		0.1328
Moderate	0.58	0.47 – 0.70	<0.0001*	0.80	0.65 – 1.00	0.0447*
High	0.66	0.53 – 0.82	0.0002*	0.88	0.69 – 1.12	0.2949
Control	1			1		
Cooking/fireplace fume exposure	1.56	1.12 – 2.19	0.0093*	1.31	0.91 – 1.89	0.1539

95% CI: 95% confidence interval; BMI: body mass index; OR: odds ratio. * $p < 0.05$.

Table 4: Multivariable associations between participant characteristics and emphysema (\geq trace), stratified by smoking status.

Variables	Current smokers (n=439)		Former smokers (n=686)		Never smokers (n=1218)	
	Adjusted OR (95% CI)	p-value	Adjusted OR (95% CI)	p-value	Adjusted OR (95% CI)	p-value
Dutch cohort	1		1		1	
Chinese cohort	1.10 (0.63–1.90)	0.7459	1.58 (0.94–2.67)	0.0858	2.62 (1.99–3.45)	<0.0001*
Age (per 10 years increase)	1.70 (1.20–2.39)	0.0025*	1.57 (1.22–2.01)	0.0004*	1.38 (1.16–1.64)	0.0003*
Female sex	1		1		1	
Male sex	1.12 (0.64–1.96)	0.6937	1.96 (1.38–2.79)	0.0002*	1.71 (1.29–2.27)	0.0002*
Control	1		1		1	
Passive smoking	0.83 (0.52–1.31)	0.4242	1.54 (1.05–2.27)	0.0274*	1.09 (0.83–1.42)	0.5487
Quit smoking years	—		0.99 (0.97–1.00)	0.1383	—	
Pack-years	1.04 (1.02–1.06)	0.0002*	—		—	
BMI < 25 kg/m ²	1		1		1	
BMI \geq 25 kg/m ²	0.53 (0.33–0.84)	0.0071*	0.73 (0.52–1.02)	0.0684	0.75 (0.59–0.96)	0.0222*
Educational level		0.1282		0.2128		0.9322
Low	1		1		1	
Moderate	0.65 (0.39–1.10)	0.1115	0.70 (0.47–1.05)	0.0824	0.95 (0.71–1.28)	0.7285
High	1.09 (0.58–2.05)	0.7938	0.75 (0.47–1.18)	0.2076	0.95 (0.69–1.32)	0.7555
Control	1		1		1	
Cooking/fireplace fume exposure	1.06 (0.50–2.25)	0.8775	1.42 (0.73–2.76)	0.3066	1.42 (0.82–2.46)	0.2119

Pack-years or quit smoking years was adjusted among current and former smokers, respectively.

95% CI: 95% confidence interval; BMI: body mass index; OR: odds ratio. * $p < 0.05$.

2.4 Discussion

In this study of general populations, which had a similar socio-demographics and smoking distribution with the population recruited in the respective cohorts, we found that the Chinese had a higher prevalence of visual emphysema on LDCT than the Dutch [64, 65]. However, this was only seen in never-smokers. Among never-smokers, increasing age, and male sex were associated with the presence of emphysema in both cohorts, but fumes exposure and passive smoking were not. Finally, although the CLE subtype was more common in the Chinese than in the Dutch population, the severity of CLE and PSE were comparable.

In total, 12 % of the Chinese never-smokers had at least mild emphysema, which is consistent with the prevalence reported in Canadian never-smokers (11 %) [66]. In Chinese ever-smokers, 39 % had at least mild emphysema, consistent with the rate in the COPD Gene study in a US population (42 %), while 22 % of our Dutch ever-smokers had at least mild emphysema [67, 68]. This lower prevalence could be explained by the lower pack-years, and fewer men in the Dutch population. The prevalence of emphysema and each emphysema severity level in Chinese was higher than in the Dutch due to the older age, more men, greater current smoking rate, pack-years, and lower BMI in the Chinese population.

However, after adjusting for multiple confounders, the Chinese had a two-fold increased odds for emphysema compared with the Dutch. The increased threshold for emphysema definition had only a minor impact on the higher odds for the Chinese population (\geq trace CLE: aOR 2.19; 95 % CI 1.77 – 2.70; \geq mild CLE: aOR 1.58, 95 % CI 1.15 – 2.17). When stratified by smoking status, only Chinese never-smokers had an increased odds compared to the Dutch. We, therefore, hypothesised that other unmeasured risk factors must account for the difference in emphysema prevalence between the two populations. A well-recognised difference is the higher outdoor air pollution in northern China than in the Netherlands (mean particulate matter 2.5: 95 vs 16 – 18 $\mu\text{g}/\text{m}^3$) [69, 70]. Previous studies have shown that air pollution not only contributes to a higher incidence of emphysema but also becomes an increasingly major risk in low-to-middle-income countries [71, 72]. Contrary to never-smokers, we observed no difference in emphysema prevalence in smokers between the two populations. Likely, this is caused by the overwhelming effect of smoking on emphysema prevalence, which covers any effect of other risk factors that could have resulted in a small difference between these two populations.

Our study showed that older age, male sex, smoking status, and low BMI in the overall population were associated with emphysema on LDCT. This is consistent with earlier reports that these are risk factors for emphysema or COPD [73, 74]. We did not detect a significant association between passive smoking exposure and emphysema in either the overall combined cohorts or the stratified cohorts. A previous study reported that passive smoking was associated with increased odds

of COPD (OR 1.18, 95 % CI 1.01 – 1.39) only when the exposure duration is at least 20 hours/week [75]. Passive smoking in our Chinese cohort was defined as positive only when the exposure duration was ≥ 15 mins/week. The lower cut-off applying for exposure may have led to the nonsignificant result. Though insignificant probably due to the lack of power (6 %), we observed a higher odds for cooking/fireplace fume exposure and emphysema in our study (OR 1.31, 95 % CI 0.91 – 1.89 in all participants; OR 1.42, 95 % CI 0.82 – 2.46 in never smokers). Previous findings showed that poor ventilation in the kitchen is associated with COPD (OR 1.28, 95 % CI 1.14 – 1.43) [76]. Furthermore, among the Chinese and Dutch never-smokers, increasing age and male sex were associated with an increased odds of emphysema, consistent with the finding for COPD risk among Korean never-smokers [77]. Meanwhile, the risk factors age and sex were present at similar effect sizes for emphysema in the Chinese and the Dutch population. The ORs of males for emphysema in the Chinese never-smokers was 1.5 (95 % CI 1.0 – 2.2), which is comparable to the odds of 1.40 (95 % CI 1.21 – 1.63) for COPD reported in a Chinese large-scale and population-based study [78].

Importantly, current smokers in our study had 2.5-fold increased odds of emphysema, reminding us of the importance of smoking in emphysema formation and supporting the necessity of smoking cessation. Our findings also remind clinicians of the need to consider screening older, male participants with low BMI, which could decrease the chance of emphysema underdiagnosis.

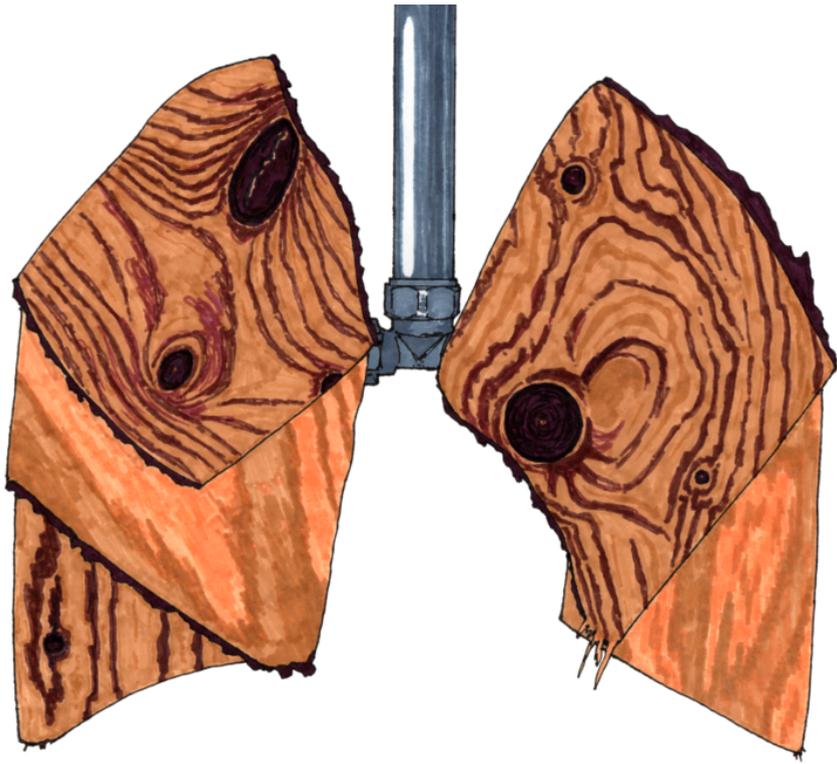
For the CT image quality, the HU deviation and HU standard deviation of air in our two cohorts are slightly higher than the requirements in the phantom (≤ 6 HU for absolute density deviation and ≤ 20 HU for standard deviation) [79]. However, these available requirements for lung density are applicable for quantitative CT assessments of emphysema, and our visual assessment of emphysema is less sensitive to image noise than quantitative assessment [43]. Therefore, we expect limited impact on our results.

Our study has some limitations. First, only one radiologist performed the emphysema assessment; however, the interobserver agreements with two other readers were good to very good, which helps to mitigate this concern.

Second, we might have an unfair comparison between the two cohorts. On the one hand, we collected some characteristics (e.g. smoking, passive smoking, fireplace fume exposure) for Dutch participants several years before the CT scan, making it possible that responses may have changed. On the other hand, the definition of variables like smoking status and passive smoking differed slightly between the Chinese and Dutch cohorts. We expect that a small proportion of misclassification of participants has a limited impact on the effect estimation for emphysema risk.

In conclusion, the Chinese have a higher prevalence of CT-defined emphysema than the Dutch in a sample of a general population, with higher odds of emphysema

among Chinese never-smokers. These findings underscore that factors other than smoking, age, and sex play a key role in emphysema formation, with outdoor air pollution being a hypothetical candidate. Considering the potentially important role of non-smoking factors in emphysema formation, studies should now focus on elucidating other risk factors that contribute to the high prevalence of emphysema in Chinese never-smokers to help to prevent the disease.



Chapter 3

Association between Chest CT–defined Emphysema and Lung Cancer: A Systematic Review and Meta-Analysis

Published in Radiology.

DOI: 10.1148/212904

Xiaofei Yang, Hendrik Joost Wisselink,

Rozemarijn Vliegenthart, Marjolein A. Heuvelmans,

Harry J.M. Groen, Marleen Vonder,

Monique D. Dorrius, Geertruida H. de Bock

Abstract

Background Given the different methods of assessing emphysema, controversy exists as to whether it is associated with lung cancer.

Purpose To perform a systematic review and meta-analysis of the association between chest CT-defined emphysema and the presence of lung cancer.

Materials and Methods The PubMed, Embase, and Cochrane databases were searched up to July 15, 2021, to identify studies on the association between emphysema assessed visually or quantitatively with CT and lung cancer. Associations were determined by emphysema severity (trace, mild, or moderate to severe, assessed visually and quantitatively) and subtype (centrilobular and paraseptal, assessed visually). Overall and stratified pooled odds ratios (ORs) with their 95 % CIs were obtained.

Results Of the 3343 screened studies, 21 studies (107 082 patients) with 26 subsets were included. The overall pooled ORs for lung cancer given the presence of emphysema were 2.3 (95 % CI 2.0 – 2.6; $I^2 = 35\%$; 19 subsets) and 1.02 (95 % CI 1.01 – 1.02; six subsets) per 1 % increase in low attenuation area. Studies with visual (pooled OR, 2.3; 95 % CI 1.9 – 2.6; $I^2 = 48\%$; 12 subsets) and quantitative (pooled OR, 2.2; 95 % CI 1.8 – 2.8; $I^2 = 3.7\%$; eight subsets) assessments yielded comparable results for the dichotomous assessment. Based on six studies (1716 patients), the pooled ORs for lung cancer increased with emphysema severity and were higher for visual assessment (2.5, 3.7, and 4.5 for trace, mild, and moderate to severe, respectively) than for quantitative assessment (1.9, 2.2, and 2.5) based on point estimates. Compared with no emphysema, only centrilobular emphysema (three studies) was associated with lung cancer (pooled OR, 2.2; 95 % CI 1.5 – 3.2; $p < 0.001$).

Conclusion Both visual and quantitative CT assessments of emphysema were associated with a higher odds of lung cancer, which also increased with emphysema severity. Regarding subtype, only centrilobular emphysema was significantly associated with lung cancer.

3.1 Introduction

Lung cancer is the primary cause of cancer-related death worldwide [80], with more than 1 million attributable deaths each year since 2000 [13]. However, lung cancer risk can be reduced by identifying treatable risk factors, such as chronic lung inflammation [81], together with smoking, genetics, diet, and occupational exposure [81]. Emphysema is characterised pathologically by the presence of diffuse chronic inflammation of the lung parenchyma, oxidative stress, and lung destruction [82]. Thus, lung cancer and emphysema are linked by common predisposing risk factors and multiple molecular inflammatory processes [83].

Emphysema can be assessed with the use of chest CT, radiography, or pulmonary function tests, although chest CT has the highest sensitivity and is considered the reference standard for noninvasive assessment [84–86]. Numerous studies have explored the association between the chest CT assessment of emphysema and lung cancer, but they have yielded inconsistent results [87–90]. Associations have been shown between emphysema and lung cancer on chest CT scans for qualitative visual assessment by radiologists [90, 91], but not for automated quantitative assessment [87, 88]. These data were subsequently confirmed by comparing the two methods directly [92], indicating that the method used to assess emphysema may have affected previous outcomes. Consistent with this, a meta-analysis in 2012 showed that visual assessment of emphysema at chest CT was independently associated with lung cancer [93], but no such association was present for quantitative assessment. However, that conclusion was based on data from only two studies. Although systematic reviews in 2020 and 2016 concluded that emphysema assessed with chest CT was associated with an increased risk of lung cancer [17, 94], pooled risk estimates were not provided, nor was data stratified by emphysema assessment method, which may have affected the reported results. Other studies exploring the association of emphysema severity or subtype visible on CT scans with lung cancer have produced mixed results [87, 95–98]. To the best of our knowledge, a pooled analysis of these associations has not been performed.

There is a need to update and synthesise data from existing and new studies, especially those using quantitative emphysema assessment. Our purpose was to perform a systematic review and meta-analysis of the association between emphysema found on chest CT and the presence of lung cancer.

3.2 Materials and Methods

Search strategy and study selection

This systematic review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses reporting guidelines [99] and registered in the international prospective register of systematic reviews, or PROSPERO

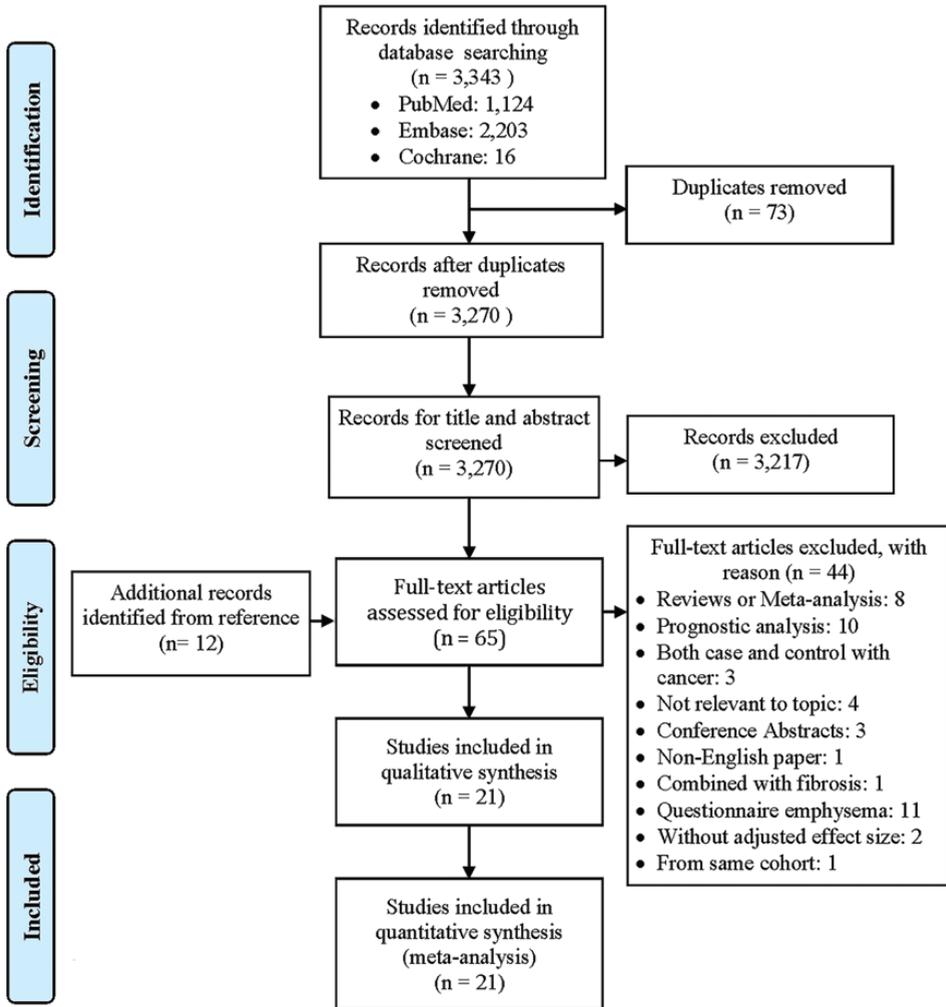


Figure 1: Flowchart of study selection.

(no. CRD42021262163). The published studies were retrieved and screened from the PubMed, Embase, and Cochrane databases from inception to July 15, 2021 (Table S3.1 [p. 191]).

We included studies investigating the association between emphysema and lung cancer if they were original research and published in English, with lung cancer diagnosed with a histopathologic examination (independent of histologic subtype) and emphysema diagnosed with a CT scan. The exclusion criteria of studies are specifically described in Figure 1. For multiple articles concerning the same cohort, we selected the study from which most data could be extracted.

Definitions of emphysema and lung cancer

Visual emphysema was defined as disrupted lung vasculature and parenchyma with low attenuation occupying any lung zone (at least trace) on chest CT, as evaluated by radiologists using the National Emphysema Treatment Trial (NETT) or Fleischner Society guidelines or comparable (Table S3.2 [p. 192]) [40, 100]. Quantitative emphysema was defined by the percentage of total lung volume below a given Hounsfield unit (HU) threshold (-950 HU at full inspiration), reported as the low attenuation area percentage (LAA%). A specific LAA% threshold was defined as “emphysema present”. In the grading of emphysema severity (trace, mild, moderate, and severe), specific percentages of visual (Fleischner Society or NETT) or quantitation were used to assess emphysematous lung tissue destruction at CT (i.e., mild: $0 - 25\%$, moderate: $26 - 50\%$, and severe: $\geq 51\%$). The main emphysema subtypes were paraseptal and centrilobular, which could only be assessed visually on CT. Paraseptal emphysema was defined as the presence of a few well-demarcated, round, juxtapleural lucencies, while centrilobular emphysema was defined as centrilobular distribution of lucencies. Finally, potential cases of lung cancer were confirmed pathologically from surgical, biopsy, or cytologic samples, without specification of the subtype.

Data collection and quality assessment

Two researchers (XY and HJW, 5 and 3 years of experience in radiology, respectively) independently performed all data collection and assessments. Study eligibility was determined by title and abstract screening, followed by full-text evaluation. Disagreements were settled by consensus or referral to a third reviewer (MDD, over 10 years of experience in radiology), and agreement was quantified with use of κ statistics. A standardised table was used to extract data, including first author name; publication year; country; study design; participant source, age, and sex; assessment method; emphysema definition, subtype, and severity; CT scanner, scanning mode, section thickness, reconstruction algorithm, HU threshold, and LAA%; effect sizes, including odds ratios (ORs), risk ratios, and hazard ratios, with 95% CIs; and adjusted or matched factors.

The Newcastle-Ottawa scale was used to assess cohort and case-control study quality by group selection, comparability, and exposure/outcome reliability, with a star-based scale ranging from zero to nine stars [101]. We awarded stars for comparability if there was adjustment for age and sex and additional adjustment for smoking status. Studies were considered to be low, medium, or high quality if they had five or fewer, six to seven, or eight to nine stars, respectively [102]. Any discrepancies were resolved by consensus.

Statistical analysis

We stratified studies by visual or quantitative assessment and set confirmed lung cancer as the main outcome. The adjusted OR given the presence of emphysema was the main outcome, with risk ratios and hazard ratios interpreted as ORs due to the low incidence of lung cancer [103, 104]. When a study reported stratified ORs, an overall OR was estimated by applying a random-effect model. For studies that stratified ORs by severity, we pooled data for moderate and severe emphysema. To estimate the odds of lung cancer developing among patients with and without emphysema, we pooled data under the assumption of homogeneity by applying a random-effect model. Forest plots are presented to illustrate the pooled results and related heterogeneity. Pooled ORs and 95 % CIs are provided for dichotomous or continuous measurements of emphysema. Analyses were repeated for emphysema severity and subtype (visual assessment).

Heterogeneity was estimated with use of the I^2 statistic and quantified as low (0 – 25 %), moderate (26 – 50 %), substantial (51 – 75 %), or considerable (76 – 100 %) [105, 106]. Potential sources of heterogeneity were explored with stratified analysis based on participant sources, study design, effect size study quality, CT section thickness (normal [< 5 mm] vs thin [0.5 – 1.25 mm]), and HU cutoff value. Funnel plots were used to evaluate publication bias. Asymmetry, which is an indication for publication bias, was evaluated visually and with the Egger test. As the next step, the trim-and-fill method was applied to evaluate the stability of our results by correcting for publication bias. The robustness of estimates was evaluated by leave-one-out sensitivity analysis, removing each study sequentially and recalculating the OR.

Statistical analysis was conducted with Stata Standard Edition, version 15.1 (StataCorp); $p < 0.05$ was considered indicative of statistically significant difference.

3.3 Results

Study selection and quality

As shown in Figure 1, 3217 of 3270 studies were excluded after screening abstract and title. Full-text screening resulted in 21 articles that met all criteria for inclusion in the meta-analysis. The κ values of the two screening stages were 0.80 (title and abstract) and 0.62 (full text), respectively. Of the included studies, two featured both visual and quantitative assessment [107, 108], 20 reported emphysema as a dichotomous variable only (visual and quantitative assessment), two as a continuous variable only [109, 110], and four as both variables [87, 88, 96, 107]. This resulted in 26 study subsets for inclusion in the final meta-analysis. Regarding study quality, 15, six, and none were considered high, medium, and low quality, respectively (Table S8.3 [p. 202]).

Study characteristics

Overall, the 21 studies included 3907 patients with lung cancer and 103 175 controls (Table 1 and Table 2), with sample sizes ranging from 120 to 62 124. By study design, cohort studies (52 % [11 of 21 studies]) contributed 1868 cases of lung cancer from 101 679 patients, and case-control studies (48 % [10 of 21]) contributed 2039 cases of lung cancer from 5403 patients. In total, 74 % of the 107 082 patients came from North America (78 874 [11 studies]), 26 % from Europe (27 392 [eight studies]), and 0.8 % from Asia (816 [two studies]).

Visual assessment was used in 12 study subsets with 95 062 patients, while quantitative dichotomous assessment was used in eight study subsets with 4758 patients, identifying emphysema in 25 % (23 742 of 95 062) and 27 % (1079 of 4046), respectively. Moreover, quantitative continuous assessment (i.e., LAA%) was used in six subsets with 10 014 patients. The definitions of emphysema used for visual and quantitative assessment varied across studies (Table S3.2 [p. 192]). The HU threshold for low attenuation area in quantitative assessments varied from -880 to -950 HU, while LAA% cutoffs for the presence of emphysema varied from 1 % to 25 %. This contributed to a wide variation in the incidence of emphysema, from 8 % (44 of 558 patients) to 80 % (195 of 243 patients). Moreover, uniformity was lacking for both HU thresholds and LAA% cutoffs for emphysema severity.

All studies confirmed lung cancer with histologic examination. A total of six studies (three visual, three quantitative; 459 lung cancers among 6242 patients) explored the relationship between emphysema severity and lung cancer, whereas three studies (all visual; 380 lung cancers among 1716 patients) explored the association between emphysema subtype and lung cancer. Participant sources were hospital-based (33 % [seven of 21 studies]) or population-based (67 % [14 of 21]).

Data synthesis and meta-analysis

The overall pooled estimate for the association between emphysema and lung cancer was 2.3 (95 % CI 2.0 – 2.6) (Figure 2), which was robust in the leave-one-out sensitivity analysis (Figure S3.1 [p. 214]). The pooled OR for every 1 % increase in the LAA% was 1.02 (95 % CI 1.01 – 1.02) (Figure S3.2 [p. 214]). Moderate heterogeneity was observed among studies ($I^2 = 34.6$ %; $p = 0.07$), reasonable symmetry was identified at the visual inspection of funnel plot (Figure S3.3 [p. 215]), and the Egger test helped identify evidence of potential publication bias ($p = 0.04$) favouring the existence of unpublished studies. Thus, the trim-and-fill correction for potential publication bias did not alter the association (pooled OR, 2.0; 95 % CI 1.7 – 2.3) (Figure S3.4 [p. 215]).

Table 1: Characteristics of included studies that assessed emphysema visually on chest CT.

Study	Country	Lung Ca yes/no*	age (y) [†]	Source	Study design	Effect size [‡]
de Torres et al. 2007 [91]	Spain	23 1166	Case: 54±8 Control: 54±8	PB	Cohort; prospective	RR: 2.5 (1.0–6.2)
Wilson et al. 2008 [90]	U.S.	99 3539	NS	PB	Cohort; prospective	OR: 3.1 (1.9–5.2)
Li et al. 2011 [111]	U.S.	565 450	Case: 67±8 Control: 66±6	HB	Case-control; retrospective	OR: 2.8 (2.1–3.8)
Maisonneuve et al. 2011 [112]	Italy	85 4511	NS	PB	Cohort; retrospective	HR: 1.8 (1.2–2.6)
Henschke et al. 2015 [113]	U.S.	668 61456	NS	PB	Cohort; prospective	OR: 2.0 (1.4–2.9)
Sanchez-Salcedo et al. 2015 [114]	Spain	53 2936	Case: 60 (55–65) [§] Control: 55 (49–62) [§]	HB	Cohort; prospective	HR: 3.3 (1.8–5.9)
de Torres et al. 2015 [115]	U.S.	134 1419	Overall: 61±7	PB	Cohort; prospective	HR: 2.7 (1.7–4.3)
Liu et al. 2018 [116]	U.S.	73 157	Case: 64 (55–74) [§] Control: 63 (55–74) [§]	PB	Case-control; prospective	OR: 1.8 (1.4–1.9)
González et al. 2019 [98]	Spain	72 215	Case: 64±9 Control: 64±9	PB	Case-control; prospective	OR: 5.4 (2.6–11.4)
Yong et al. 2019 [117]	Norway	367 16257	Case: 62±6 Control: 61±5	PB	Cohort; retrospective	HR: 2.0 (1.6–2.6)

See page 42 for the secondary caption.

Table 2: Characteristics of included studies that assessed emphysema quantitatively on chest CT.

Study	Country	Lung Ca yes/no*	Age (y) [†]	Source	Study design	Effect size [‡]
Kishi et al. 2002 [88]	U.S.	24 96	Case: 64±7	HB	Case-control; retrospective	OR: 1.1 (0.5–2.4)
			Control: 63±6			OR: 1.1 (0.6–1.9) [§]
Maldonado et al. 2010 [87, 118]	U.S.	64 377	Case: 63±7	PB	Case-control; prospective	OR: 1.9 (1.1–3.3)
			Control: 62±6			OR: 1.04 (0.8–1.3) [§]
Gierada et al. 2011 [89]	U.S.	279 279	Case: 63±5	PB	Case-control; retrospective	OR: 2.0 (1.0–3.8)
			Control: 61±5			
Gagnat et al. 2017 [96]	Norway	34 741	Overall: 59±10	PB	Cohort; prospective	HR: 2.4 (0.9–6.2)
						HR: 1.03 (0.7–1.5) [§]
Chubachi et al. 2017 [95]	Japan	21 219	Case: 73±7	HB	Cohort; prospective	OR: 4.2 (1.0–29.0)
			Control: 73±8			
Mouronte-Roibás et al. 2018 [97]	Spain	169 74	Case: 69±9	HB	Cohort; retrospective	OR: 2.2 (1.1–4.3)
			Control: 65±10			
Nishio et al. 2019 [110]	Japan	283 293	Case: 69±10	HB	Cohort; retrospective	OR: 1.01 (1.00, 1.02) [§]
			Control: 65±14			
Husebø et al. 2019 [119]	Norway	31 681	Case: 64±7	HB	Cohort; prospective	HR: 4.4 (1.7–10.8)
			Overall: 58±10			
Labaki et al. 2021 [109]	U.S.	353 6909	Overall: 62±5	PB	Cohort; prospective	HR: 1.02 (1.01–1.03) [§]
Schwartz et al. 2016 [108]	U.S.	341 752	Case: 64±10	PB	Cohort; retrospective	OR (vis.): 1.8 (1.4–2.4)
			Control: 62±9			OR (quant.): 2.7 (1.8–4.0)
Carr et al. 2018 [107]	U.S.	169 671	Case: 66±8	PB	Case-control; prospective	OR (vis.) [§] : 2.3 (1.4–3.8)
			Control: 64±8			OR (quant.) [§] : 1.03 (0.6–1.8)

See page 42 for the secondary caption.

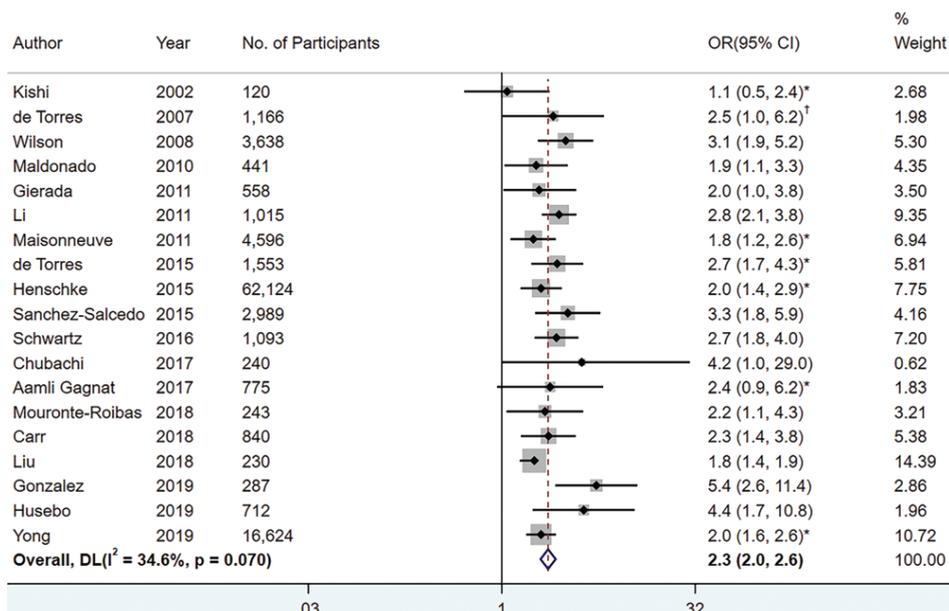


Figure 2: Forest plot of the random-effects meta-analysis for the association between emphysema (dichotomous variable) assessed visually and/or quantitatively with CT and lung cancer in 19 studies. The overall pooled odds ratio (OR) of emphysema for lung cancer was 2.3 (95 % CI 2.0 – 2.6 [$p < 0.001$]). For the studies that assessed emphysema with two methods, only the ORs assessed with the main method were pooled in the overall estimates. Squares and horizontal lines represent estimates and 95% CIs, respectively, for each study part. Diamonds indicate pooled effect sizes with 95% CIs. DL: DerSimonian and Laird; . *: Study reported hazard ratios; †: Study reported risk ratios.

Association between emphysema and lung cancer

The pooled OR for lung cancer given emphysema was 2.3 (95 % CI 1.9 – 2.6) in studies using visual assessment and 2.2 (95 % CI 1.8 – 2.8) in studies using quantitative dichotomous assessment (Figure 3). Low heterogeneity ($I^2 = 3.7\%$; $p = 0.40$) was observed in studies using quantitative assessment, and moderate heterogeneity ($I^2 = 48.4\%$; $p = 0.03$) was observed in studies using visual assessment (Table 3).

Association between emphysema severity and lung cancer

Independent associations existed between different emphysema severities and lung cancer (Figure 4), with the overall pooled ORs for lung cancer gradually increasing (2.2, 3.2, and 3.6) as the emphysema severity increased (trace, mild, and moderate to severe, respectively) (Table 4). Substantial heterogeneity was observed for studies that reported moderate to severe emphysema ($I^2 = 52.6\%$) compared with trace ($I^2 = 0\%$) and mild ($I^2 = 20.7\%$) emphysema. The three studies that

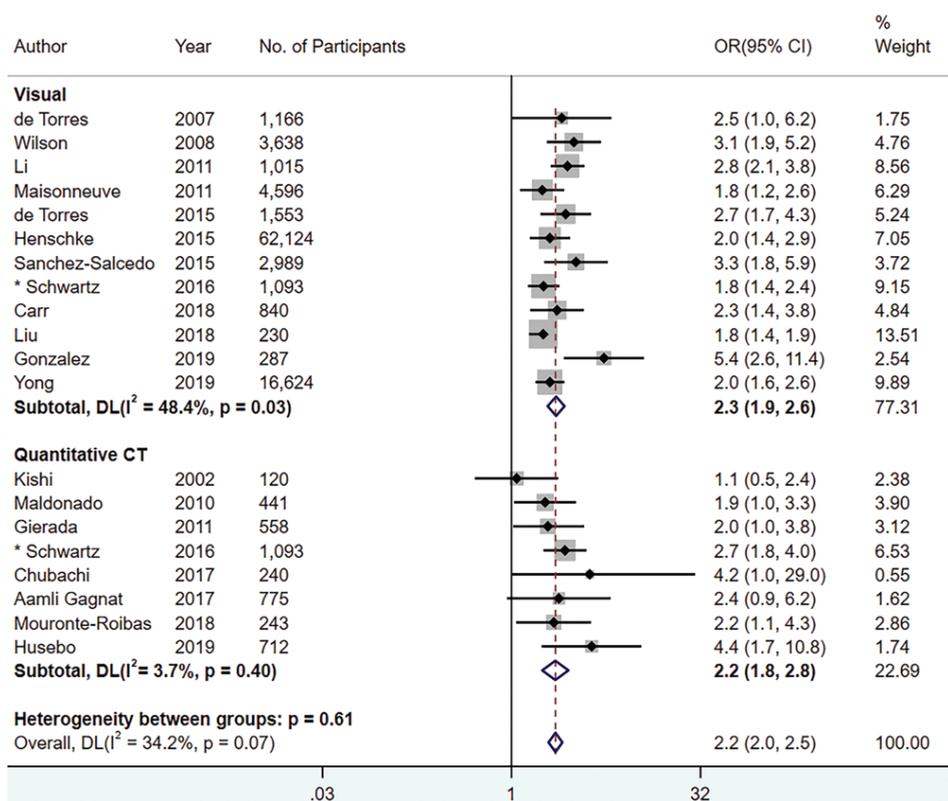


Figure 3: Forest plot of the random-effects meta-analysis for the association between emphysema and lung cancer, stratified by the emphysema assessment method. The pooled odds ratios (ORs) for lung cancer given visual and quantitative dichotomous emphysema assessment were 2.3 (95 % CI 1.9–2.6 [$p < 0.001$]) and 2.2 (95 % CI 1.8 – 2.8 [$p < 0.001$]), respectively. Squares and horizontal lines represent estimates and 95% CIs, respectively, for each study part. Diamonds indicate pooled effect sizes with 95% CIs.

DL: DerSimonian and Laird. *: Study assessed emphysema both visually and quantitatively.

used visual assessment gave pooled ORs of 2.5, 3.7, and 4.5 for trace, mild, and moderate to severe emphysema, respectively; by contrast, the three studies that used quantitative assessment produced corresponding pooled ORs of 1.9, 2.2, and 2.5.

Association between visual emphysema subtypes and lung cancer

The pooled OR for lung cancer odds in the presence of centrilobular emphysema was 2.2 (95 % CI 1.5 – 3.2), with no heterogeneity observed across the three relevant studies ($I^2 = 0\%$). However, we found no evidence of an association between paraseptal emphysema and lung cancer (pooled OR, 1.1; 95 % CI 0.6–2.0) (Table 5), and there was high heterogeneity ($I^2 = 65.6\%$) (Figure 5) in this subset.

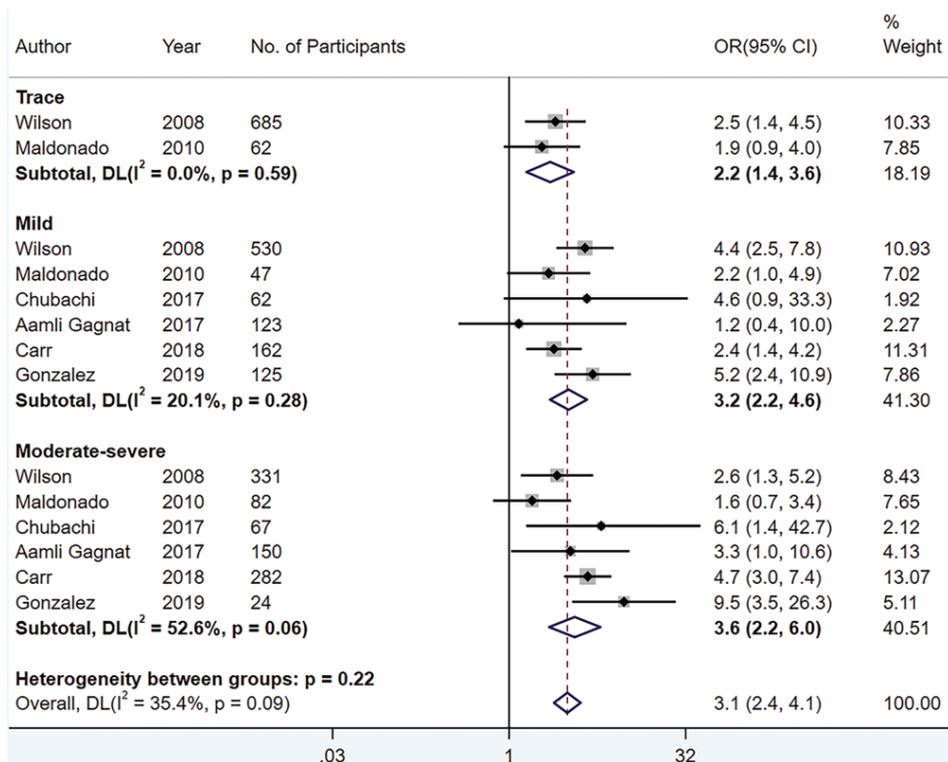


Figure 4: Forest plot of the random-effects meta-analysis for the association between emphysema severity (assessed visually and/or quantitatively) and lung cancer. The overall pooled odds ratios (ORs) of trace, mild, and moderate to severe emphysema for lung cancer were 2.2 (95% CI 1.4 – 3.6 [$p = 0.001$]), 3.2 (95% CI 2.2 – 4.6 [$p < 0.001$]) and 3.6 (95% CI 2.2 – 6.0 [$p < 0.001$]), respectively. Adjusted factors in these mixed-effects models varied, as shown in Table S3.2 [p. 192]. Squares and horizontal lines represent estimates and 95% CIs, respectively, for each study part. Diamonds indicate pooled effect sizes with 95% CIs. DL: DerSimonian and Laird.

Sources of heterogeneity

In the additional stratified analyses, the potential reasons for heterogeneity were explored (Table S3.6 [p. 213]), but we could not find any explanation. The pooled ORs were comparable between case-control (2.2; 95% CI 1.8 – 2.8; $I^2 = 55.0\%$) and cohort (2.3; 95% CI 2.0 – 2.7; $I^2 = 0\%$) studies ($p = 0.46$). Population-based studies, which had moderate heterogeneity ($I^2 = 27.0\%$), had a comparable pooled OR (2.2; 95% CI 1.9 – 2.5) to that of hospital-based studies (2.6; 95% CI 1.9 – 3.6; $I^2 = 32.7\%$ [$p = 0.06$]). The variation in study characteristics and study quality did not affect our results (Table S3.6 [p. 213]). The pooled effect sizes were comparable between studies that reported hazard ratios (2.3; 95% CI 1.9 – 2.9; $I^2 = 19.3\%$) and those that reported ORs (2.3; 95% CI 1.9 – 2.8;

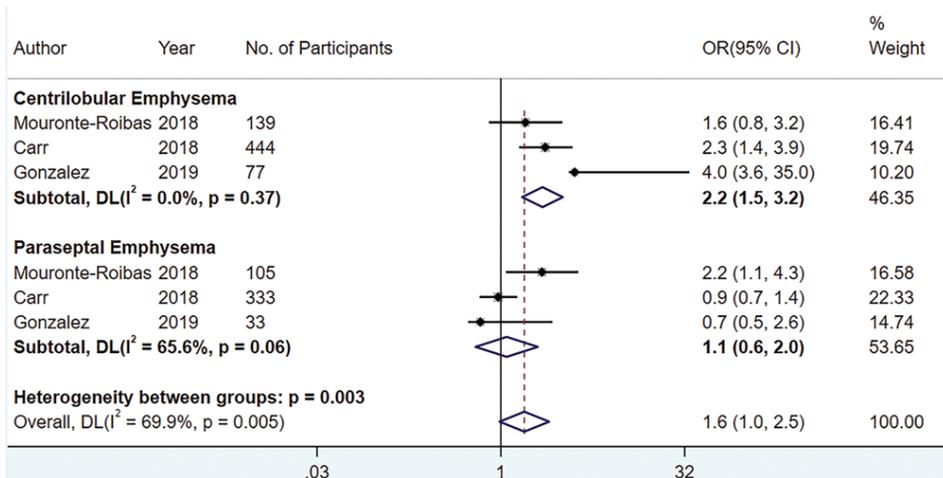


Figure 5: Forest plot of the random-effects meta-analysis for the association between emphysema subtype (assessed visually only) and lung cancer. The pooled odds ratios (ORs) for lung cancer odds in the presence of centrilobular and paraseptal emphysema were 2.2 (95 % CI 1.5 – 3.2 [$p < 0.001$]) and 1.1 (95 % CI 0.6 – 2.0 [$p = 0.71$]). Adjusted factors in these mixed-effects models varied, as shown in Table S3.2 [p. 192]. Squares and horizontal lines represent estimates and 95% CIs, respectively, for each study part. Diamonds indicate effect sizes with 95% CIs.

DL: DerSimonian and Laird.

$I^2 = 47.6\%$ [$p = 0.64$]). Emphysema assessed quantitatively based on thin CT sections was associated with lung cancer (pooled OR, 2.2; 95 % CI 1.3 – 3.7; $p = 0.002$), while this was not the case for the assessment based on normal section thickness. Similarly for LAA HU thresholds, an association with lung cancer was found based on a cutoff of -950 HU (pooled OR, 2.6; 95 % CI 2.0 – 3.4; $p < 0.001$), but not for -900 HU.

Secondary caption Table 1:

See Table S3.4 [p. 203] for full details.

HB: hospital-based; HR: hazard ratio; NS: not specified; OR: odds ratio; PB: population-based; RR: risk ratio. For full details, see Table S3.4 [p. 203].

*: Data are numbers of patients; †: Unless otherwise specified, data are means±SDs;

‡: Data in parentheses are 95% CIs. All effect sizes are adjusted for smoking status, except for the study by Henschke et al. For specific adjusted factors, see Table S3.2 [p. 192]; §: Data are medians, with ranges in parentheses.

Secondary caption Table 2:

See Table S3.5 [p. 208] for full details.

Lung Ca: lung cancer; HB: hospital-based; HR: hazard ratio; NS: not specified; OR: odds ratio; PB: population-based; RR: risk ratio. For full details, see Table S3.5 [p. 208].

*: Data are numbers of patients; †: Data are means±SDs; ‡: Data in parentheses are 95% CIs; §: Effect size when emphysema was assessed as a continuous variable. All effect sizes are adjusted for smoking status. For specific adjusted factors, see Table S3.2 [p. 192].

Table 3: Association between emphysema and lung cancer stratified by emphysema assessment method

Assessment method	Visual	Quantitative
No. of studies	12	8
No. of participants	95 561	5531
No. of lung cancers	2330	1616
Pooled odds ratio	2.3	2.2
95% CI	1.9 – 2.6	1.8 – 2.8
I ² (%)	48.4	3.7
p-value for heterogeneity	0.03	0.40
p-value for method	0.61	

Unless otherwise specified, analysis was based on emphysema when measured as a dichotomous variable.

Table 4: Association between emphysema severity and lung cancer.

Emphysema severity	No. of studies	No. of participants	No. of lung cancers	Pooled odds ratio	95% CI	I ² (%)	p-value heterogeneity	p-value severity
Overall								0.22
Trace	3	747	34	2.2	1.4 – 3.6	0	0.59	
Mild	6	1049	140	3.2	2.2 – 4.6	20.1	0.28	
Moderate to severe	6	936	168	3.6	2.2 – 6.0	52.6	0.06	
Visual								0.27
Trace	1	685	22	2.5	1.4 – 4.5	
Mild	3	817	118	3.7	2.3 – 5.8	42.9	0.17	
Moderate to severe	3	637	124	4.5	2.5 – 8.3	55.9	0.10	
Quantitative*								0.94
Trace	1	62	12	1.9	0.9 – 4.0	
Mild	3	232	22	2.2	1.1 – 4.3	0	0.56	
Moderate to severe	3	299	44	2.5	1.2 – 5.1	23.0	0.27	

*: Cutoff value for emphysema severity varied among six studies.

Table 5: Association between emphysema subtype (visual assessment) and lung cancer

Emphysema subtype	Centrilobular emphysema	Paraseptal emphysema
No. of studies	3	3
No. of Participants	660	471
No. of Lung Cancers	258	153
Pooled Odds Ratio	2.2	1.1
95 % CI	1.5 – 3.2	0.6 – 2.0
I ² (%)	0	65.6
p-value for heterogeneity	0.37	0.06
p-value for subtype	0.003	

3.4 Discussion

In this systematic review and meta-analysis comparing the association of emphysema at chest CT with the presence of lung cancer, we found that both the visual and quantitative CT assessments of emphysema were associated with a higher risk of lung cancer (pooled odds ratio [OR], 2.3; 95 % CI 1.9 – 2.6; $p < 0.001$), and the odds increased with emphysema severity. Regarding subtype, only centrilobular emphysema was associated with lung cancer (pooled OR, 2.2; 95 % CI 1.5 – 3.2; $p < 0.001$).

Our study showed that emphysema at CT was associated with a 2.3-fold increased odds of lung cancer, comparable to that reported by Brenner et al. [120] and Zhang et al. [121]. However, Smith et al. [93] only found this association for visually diagnosed emphysema, whereas our study demonstrated it for both visual and quantitative methods, irrespective of whether emphysema was analysed as a dichotomous or continuous variable. An explanation for this difference may be that Smith et al. only included two quantitative CT studies in 2012 (1549 patients), while in our analysis, 10 studies were included (12 841 patients).

There was no evidence showing that source of population or study design influenced the overall association between emphysema and lung cancer. Besides, in our study, we found comparable pooled ORs for visual and quantitative assessment, implying no difference between them. Nonetheless, each method of emphysema assessment has its own limitations. Visual assessment is time-consuming, subjective, and experience-dependent and has high inter- and intraobserver variability despite well-established and standardised criteria [40, 103]. In contrast, although quantitative assessment is objective, quick, and highly reproducible when similar devices and protocols are used, it is hampered by inconsistencies in factors like the section thickness, HU threshold (-900 HU or -950 HU), and LAA% cutoffs ($1 - 25$ %). To illustrate this, we found no evidence of an association ($p = 0.09$) between emphysema and lung cancer when emphysema was quantitatively assessed at thick-section chest CT with a cutoff value of -900 HU. Therefore, it is recommended that a thin section thickness (≤ 1.5 mm) and -950 HU cutoff value are used for quantitative emphysema assessment. Given that each of these factors may affect emphysema detection with the quantitative method [92], standardisation is needed to ensure the precision, reliability, and robustness required for widespread use [122–124].

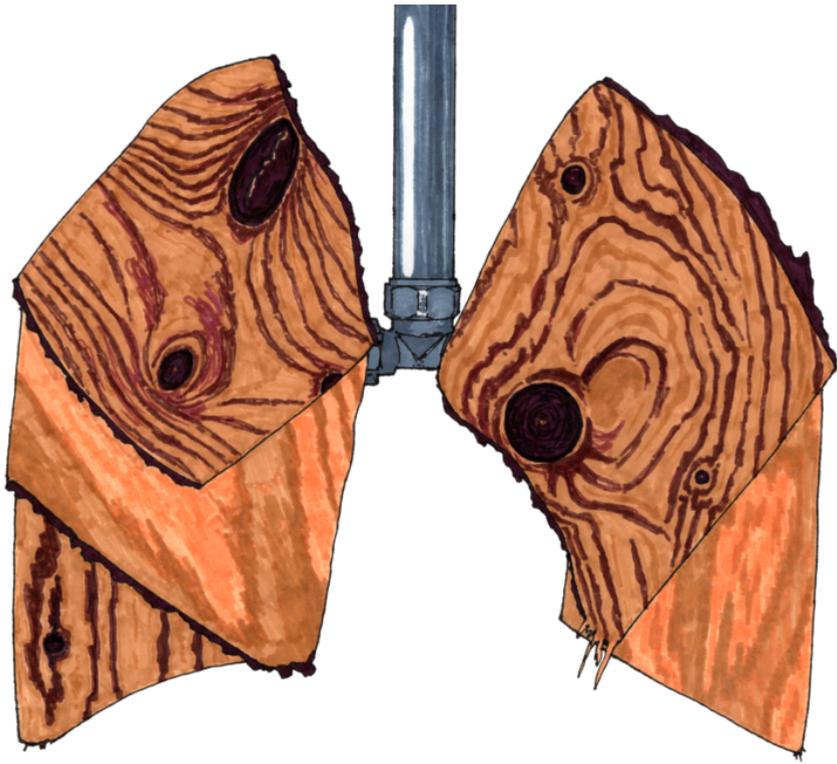
The presence of emphysema, irrespective of its severity, was related to the presence of lung cancer. The odds of lung cancer increased with increasing levels of emphysema severity. We identified several studies that reported inconsistent results regarding the association between increasing emphysema severity and increasing lung cancer odds, with some suggesting that this trend existed [95, 98] and others suggesting the opposite [87, 107]. It may be that the limited sample

sizes for severe emphysema in the studies (82 and 135 patients) resulted in no trend being visible. The analysis stratified by assessment method showed that ORs for lung cancer increased with increasing emphysema severity and that this association was higher for visual assessment. This is not surprising, given that visual assessment relies on subjective estimation of emphysema severity and not a pre-specified HU threshold. Validated or cross-calibrated quantitative and visual assessments of severity have not previously been well established in the literature. Our cutoff values for categorising emphysema severity were generally higher for the visual (mild, $\leq 25\%$; moderate, $>25\%$) than for the quantitative (mild, $\leq 10\%$; moderate, $>10\%$) assessments [87, 90].

Centrilobular emphysema, but not paraseptal emphysema, was independently associated with an increased odds of lung cancer. Although these results should be interpreted cautiously due to their reliance on only three studies, the large sample of 1370 participants should increase the reliability (48% centrilobular, 34% paraseptal, 15% controls) [97, 98, 107]. If paraseptal emphysema truly has no association with lung cancer, its presence may also explain existing discrepancies.

Our study has limitations. First, airflow obstruction is an independent risk factor for lung cancer [125], yet some included studies did not adjust for its presence (62% [13 of 21 studies]). This confounder could have affected the pooled OR for lung cancer. Second, only six studies reported the effect of emphysema severity on lung cancer, and only two reported the association for trace emphysema. Third, based on the included data in this meta-analysis, it was not possible to determine whether the presence of CT-defined emphysema leads to incremental and independent prognostic value over that of already known (shared) risk factors of emphysema and lung cancer. Finally, the cutoff value for the presence of emphysema and its severity varied among the studies, and this may likely have affected the pooled ORs.

In conclusion, emphysema diagnosed at chest CT was independently associated with a higher odds of developing lung cancer, regardless of whether it was assessed visually or quantitatively. Moreover, this risk increased with emphysema severity. Concerning visual assessment by subtype, only centrilobular emphysema was significantly associated with lung cancer. To benefit from the potential value of visual and quantitative CT assessments in early emphysema detection and lung cancer screening, research must now establish guidelines for scanning protocols, evaluation, and nodule risk stratification.



Chapter 4

Predicted versus CT-derived total lung volume in a general population: the ImaLife study

Published in PLoS ONE.

DOI: 10.1371/journal.pone.0287383

*Hendrik Joost Wisselink, Danielle J.D. Steerenberg,
Mieneke Rook, Gert Jan Pelgrim,*

*Marjolein A. Heuvelmans, Maarten van den Berge,
Geertruida H. de Bock, Rozemarijn Vliegthart*

Abstract

Predicted lung volumes based on the Global Lung Function Initiative (GLI) model are used in pulmonary disease detection and monitoring. It is unknown how well the predicted lung volume corresponds with computer tomography (CT) derived total lung volume (TLV). The aim of this study was to compare the GLI-2021 model predictions of total lung capacity (TLC) with CT-derived TLV. 142 female and 131 male healthy participants (age 45 – 65 years) were consecutively selected from a Dutch general population cohort, the Imaging in Lifelines (ImaLife) cohort. In ImaLife, all participants underwent low-dose, inspiratory chest CT. TLV was measured by an automated analysis, and compared to predicted TLC based on the GLI-2021 model. Bland-Altman analysis was performed for analysis of systematic bias and range between limits of agreement. To further mimic the GLI-cohort all analyses were repeated in a subset of never-smokers (44 % of the cohort). Mean \pm SD of TLV was 4.7 ± 0.9 L in women and 6.1 ± 1.2 L in men. TLC overestimated TLV, with systematic bias of 1.0 L in women and 1.7 L in men. Range between limits of agreement was 3.2 L for women and 4.2 L for men, indicating high variability. Performing the analysis with never-smokers yielded similar results. In conclusion, in a healthy cohort, predicted TLC substantially overestimates CT-derived TLV, with low precision and accuracy. In a clinical context where an accurate or precise lung volume is required, measurement of lung volume should be considered.

4.1 Introduction

Pulmonary conditions are common, with two major diseases - asthma and chronic obstructive pulmonary disease (COPD) - adding up to a global prevalence of 13.1 % [126]. For diagnosis and disease monitoring of COPD, several lung volumetric parameters are determined, including the total lung capacity (TLC) [48]. While the diagnosis of COPD is still based on the results of spirometry, the (separately measured) TLC is often of great importance as an additional measure.

There are three methods to measure the TLC. If performed at end-tidal volume, the gas dilution method (often performed with helium) and body plethysmography (often called body box), provide the functional residual capacity [48, 127] that can be added to the inspiratory capacity to obtain the TLC [128, 129]. The third method is the use of an inspiratory computer tomography (CT) scan, on which the lungs can be segmented, generally without the conducting airways [130–132]. This method relies on the assumption that the CT scan is acquired at full inspiration. Gas dilution and body box will mostly have matched results for subjects without air trapping [48]. While a CT scan allows diagnostic evaluation of both airways and parenchyma, the CT-derived total lung volume (TLV) tends to differ slightly from the first two methods, although there is a strong correlation between TLV and gas dilution or body box (r 0.87 – 0.90) [133–137]. Which of these three should be considered the reference standard depends on the specific clinical question or research goal [128].

To give a correct interpretation of lung volume measurements with regard to potential disease presence, severity and progress in time, expected values are required for reference [36]. Accurate prediction of TLC is of importance in some clinical applications, such as in lung transplantation where a potential lung donor is matched to a recipient [138, 139]. Recently, the Global Lung Function Initiative (GLI) published a guideline with updated models to predict the median values for several static lung volumes for healthy individuals, among which the TLC [48]. This model was endorsed by the European respiratory society (ERS) [48]. The 2021 TLC model is a generalised additive model of location, shape, and scale (GAMLSS), which is mathematically similar to a logistic model with age and height as parameters. It also includes a spline term that depends on sex and age. To the best of our knowledge, the GLI-2021 model has not been directly compared to CT-derived lung volume. It is unknown how well the new GLI model corresponds with CT-derived lung volumes.

The GLI models are often applied to clinical non-healthy populations, for instance to provide a baseline estimation at time of diagnosis and for follow-up purposes, expressing measurements as percentage of expected or predicted [41]. This may lead to a mismatch in clinical practice if the goal is to estimate the expected lung volume in a normal healthy person instead of the idealised reference population used by the GLI [140]. The aim of this study was to compare the outcomes of

the GLI-2021 model with CT-derived lung volumes in a healthy consecutively selected sample from a Western European general population-based study cohort.

4.2 Materials and methods

Participant selection

CT scans in this study were acquired as part of the ongoing ImaLife study. ImaLife is embedded in Lifelines, a population-based cohort study in the northern part of the Netherlands [10, 59, 141]. Lifelines is a multi-disciplinary prospective population-based cohort study examining - in a unique three-generation design - the health and health-related behaviours of 167 729 individuals living in the North of the Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics [59]. In ImaLife, participants from Lifelines aged 45 or older are invited to undergo a low-dose chest CT scan. Ethical approval for the ImaLife study was given by the institutional ethical review board and all participants provided written informed consent. For our present study, the aim was to select a sample of 400 participants from this cohort, by consecutively including 50 women and 50 men per 5-year age group, with an age range of 45 – 65 years. This was done to achieve an even distribution across age. Participants with incomplete imaging data ($n=3$) or missing weight information ($n=5$) were replaced by continuing the sampling. To reach the goal of 400 included participants, a selection was performed from the 1421 CT scans that were acquired between June and December 2018. Other than data availability, age, and sex, no special inclusion or exclusion criteria were applied for this initial selection. Prior to the main analyses of this study, we excluded participants with COPD or self-reported lung disease, as well as participants who received a follow-up CT scan or were referred to primary care ($n=127$). The main analysis was performed on the 273 healthy participants (cohort H). Additional analyses were performed for the full general population sample (cohort GP, $n=400$) and including only healthy never-smokers (cohort HNS, $n=119$). A flowchart detailing the selection steps is shown in Figure 1.

Lifelines parameters

This study uses data from the second assessment round of Lifelines (2014-2018), which includes questionnaire answers, as well as results from a pulmonary function test [10, 59]. The questionnaire data included smoking status, pack-years, and self-reported lung disease. The spirometric data included the Forced Expiratory Volume in 1 second (FEV_1) and Forced Vital Capacity (FVC), which allows determination of the GOLD stage, but this does not allow derivation of the TLC [41]. Participant height and weight were self-reported during the assessment and shortly

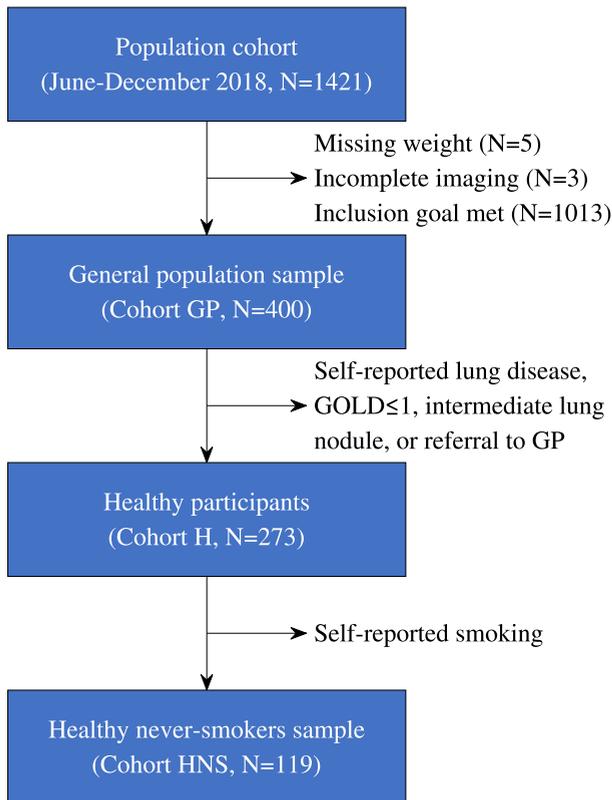


Figure 1: Flow chart describing the selection of the GP (general population), H (healthy participants), and HNS (healthy never smokers) cohorts from the ImaLife study.

before the CT scan, respectively. The body mass index (BMI) was computed from body weight and height. For the purposes of the analyses in this study, a participant was considered to be healthy if the spirometry did not indicate COPD and if she/he reported no COPD, emphysema, chronic bronchitis, or asthma. If a participant was invited for a follow-up scan for an intermediate lung nodule or was referred to primary care for an incidental finding, this participant was considered non-healthy. The exact criteria for a follow-up or referral can be found in the ImaLife design paper [10]. In case of missing data, participants were considered ever-smokers or non-healthy, respectively.

Data availability

The data used for this study can be requested through the Lifelines biobank catalogue (<https://data-catalogue.lifelines.nl>), except for the imaging data. All identifying participant information is stored with Lifelines, in their role as Trusted Third Party [10]. This includes the written informed consent specific to the ImaLife study pursuant to the ethical approval by the institutional review board of the University Medical Center Groningen. The ImaLife study was registered with the Dutch Central Committee on Research Involving Human Subjects (<https://www.toetsingonline.nl>, Identifier: NL58592.042.16).

Given the larger file sizes and specialised analysis tools required, there are no automatic systems to request the imaging data. Despite this practical limitation, the imaging data are available for research. Lifelines or the corresponding author can be contacted for a tailored data sharing solution.

TLC prediction model

For this study, the ERS-endorsed GLI guideline model was used [48]. The GLI model and its predecessors were developed with the use of participants without a history of smoking or lung disease only [48, 142, 143]. The 2021 model equations look like a stratified logistic regression, although the method used to derive these equations is a generalised additive models of location, shape and scale (GAMLSS) [48]. The model is based on age and height, see Equation 4.1 [48]. Because one of the parameters in this model (M_{spline}) is a variable based on sex and age, a lookup table is required to use this model, which is provided as a supplementary material to the original publication (permanently archived at <http://web.archive.org/web/20210629151841/https://erj.ersjournals.com/content/erj/57/3/2000289/DC1/embed/inline-supplementary-material-2.xlsx?download=true>).

$$\begin{aligned} V_{women} &= e^{-10.1128+0.1062*\ln(age)+2.2259*\ln(height)+M_{spline}} \\ V_{men} &= e^{-10.5861+0.1433*\ln(age)+2.3155*\ln(height)+M_{spline}} \end{aligned} \quad (4.1)$$

CT scan data collection

Low-dose CT scans were acquired on a third-generation dual source CT system (Somatom Force, Siemens Healthineers) with a tube potential of 120 kV and a reference current-time product of 20 mAs (median dose-length product for cohort H 58 mGy, range 29 – 113 mGy) [10]. The field of view was 350 mm (with a pitch of 3.0), or, in case of a large body habitus, 400 mm (pitch 2.5). Scans were reconstructed with a slice thickness/increment of 1.0/0.7 mm, yielding approximately isotropic voxels. For this study, the reconstruction with a medium-smooth (Br40) kernel was used. The scans were acquired at inspiration according to clinical standard breath coaching.

Image analysis

Image analysis consisted of a fully automatic extraction of the lung volume from the CT scan. This was performed with the Syngo.Via Pulmo3D package (version VB40A-HF02, Siemens Healthineers), which did not require manual interaction. A trained researcher (DS) checked the segmentation quality. This quality check consisted of confirming all lung parenchyma was included. Lobar segmentation failures were accepted as long as the overall lung volume was correct. An example of the segmentation result is included in the supplemental materials. CT scans of cases with a large difference between the GLI model prediction and

the CT-derived lung volume, i.e. a difference in the upper and lower 5 % extremes, were visually inspected. This visual inspection was performed to ensure acquisition problems (e.g. substantial omission of an apical or caudal section of the lungs) or major pathology (e.g. severe emphysema/fibrosis and marked pleural disease) were not present and could therefore not bias the lung volume. Technical physicians (HJW and GJP, 4 years of experience in chest CT research/scan evaluation) performed visual review of these 28 cases.

Statistical analyses

The two-sample Kolmogorov–Smirnov test was used to determine whether TLV, weight, and height for women and men are from the same distribution. Differences in age, height, weight, and TLV between women and men were tested with t-tests. Then, linear regression was used to predict TLV stratified by sex, where age, height, and weight were included as parameters. Model performance of the two models was quantified by correlating the predicted model values with the observed values of TLV by using Pearson's ρ to estimate correlation and R^2 to estimate model fit. Then Bland-Altman analyses were performed to evaluate the systematic differences between the estimated values and the observed TLV values. The mean difference was considered as the estimated bias, and the variability is indicated by the difference between the 95 % limits of agreement (ΔLoA). Levene's test was used to test whether the ΔLoA was the same between models and the Wilcoxon rank-sum test was used to test difference between systematic biases. All analyses were stratified by sex. The results of the Bland-Altman analysis were shown in a residual plot, showing the measured volume on one axis and the difference between TLV and TLC on the other axis.

The Bland-Altman analyses were repeated with the original consecutively selected general population sample (cohort GP, $n=400$) and with only the healthy never-smokers (cohort HNS), see Figure 1. The cohort HNS was used to further mimic the cohort used for the GLI model [48].

For a sensitivity analysis, the volumes reported by Yamada et al. were used to correct for the positional difference between the CT (supine) and PFT (sitting). Yamada et al. performed standing and supine CT scans on 32 healthy volunteers [133]. Since they found the standing CT volume to be 10.9 % higher than the supine CT volume, this percentage was added to the CT measurements in this study. This corrected TLV was then compared to the GLI-predicted TLC in a Bland-Altman analysis.

Statistical analysis of derived data was performed with SPSS 26 (IBM). Data visualisation and simple computations were done with MATLAB R2022b (MathWorks).

Table 1: Population characteristics stratified by sex.

Variable	Women (N=142)	Men (N=131)	p-value
Age (years)	54±5.4	53±5.5	0.331
Weight (kg)	74±12	87±11	<0.001
Height (m)	1.70±0.07	1.84±0.07	<0.001
BMI (kg/m ²)	25.6±3.8	25.7±2.9	0.790
Smoking status	Never: 61 (43%) Past: 57 (40%) Current: 20 (14%) Missing: 4 (3%)	Never: 58 (44%) Past: 51 (39%) Current: 17 (13%) Missing: 5 (4%)	N.A.
Pack-years	7.7±7.4	8.9±7.7	0.350
Emphysema score	4.0±3.3	6.2±4.0	<0.001
CT-diagnosed emphysema	None: 98 (69%) Trace: 43 (30%) Mild: 1 (1%)	None: 57 (44%) Trace: 70 (53%) Mild: 4 (3%)	N.A.
FEV ₁ (L)	2.9±0.5	4.2±0.6	<0.001
FVC (L)	3.8±0.6	5.3±0.8	<0.001

Values are mean±SD or N (percentage).

Pack-years were calculated for all ever-smokers. Emphysema score was quantified as LAV%_(-950 HU). CT-diagnosed emphysema categories were based on LAV%: none (<5%), trace (5 – 15%), and mild (>15%).

BMI: Body-mass index; FEV₁: forced expiratory volume in 1 second;
FVC: forced vital capacity; HU: Hounsfield unit.

4.3 Results

Visual review of cases with a large difference between predicted TLC and measured TLV did not reveal any anomalies substantial enough to warrant exclusion of any participant. Mean participant age was 54 and 53 years for women and men, respectively (Table 1). Mean weight was 74 kg for women and 87 kg for men, and mean height was 1.70 m for women and 1.84 m for men (mean BMI 25.6 kg/m² for women and 25.7 kg/m² for men). The prevalence of smoking or ever-smoking was 57% for women and 56% for men (including missing data in 4 and 5 cases, respectively).

The plots in Figure 2 show TLV, height, and weight versus age. None of the scatter plots suggest a strong correlation with age.

Observed mean TLV was lower for women than for men: 4.7 L (SD 0.9 L) versus 6.1 L (SD 1.2 L), respectively ($p < 0.0001$). Mean TLC according to the GLI-2021 was 5.7 L for women (SD 0.5 L) and 7.8 L for men (SD 0.7 L). Compared to TLV, the systematic bias of the TLC was 1.0 L for women and 1.7 L for men, indicating on average overestimation of lung volume based on the GLI model. The difference ranged from 0.9 L underestimation to 4.0 L overestimation.

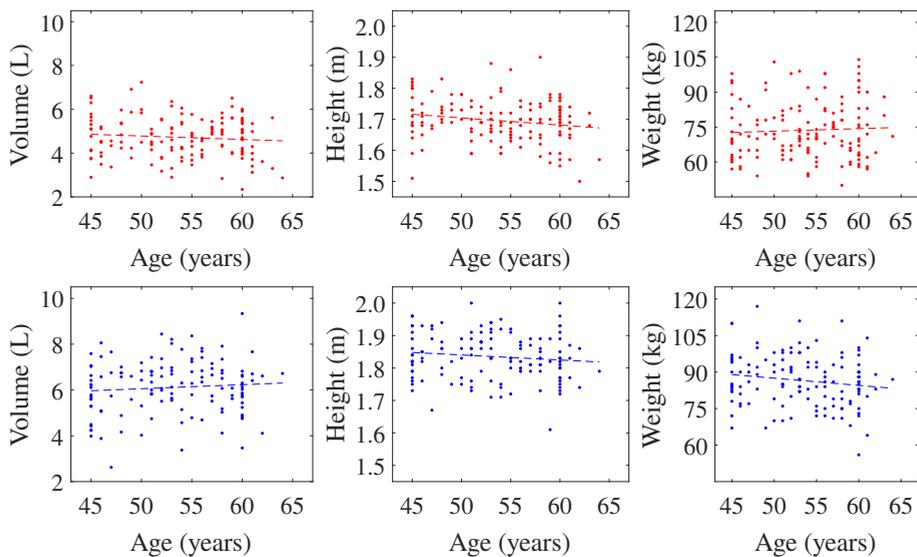


Figure 2: Explorative scatter plots showing age plotted against **A**) lung volume (measured on CT), **B**) height, and **C**) weight. The dotted lines are linear trend lines, determined separately for women (top row, red markers/line) and men (bottom row, blue markers/line).

The residual plots in Figure 3 show the results of the Bland-Altman analysis. The difference between the limits of agreement (ΔLoA) was 3.2 L for women and 4.2 L for men, indicating large variability of GLI-model results to TLV.

The correlation plots for the TLC and the TLV are shown in Figure 4. For larger lung volumes, the TLC and TLV were mostly the same, but for smaller lungs there was a progressive difference, with the predicted TLC increasingly overestimating the actual measured lung volume.

Re-including participants with lung disease or performing a sub-analysis on the healthy never-smokers did not result in significantly different systematic bias or ΔLoA ($p = 0.094 - 0.784$). A full population description of the three cohort subgroups (general population sample, healthy participants, and healthy never-smokers) is available in Table S4.1 [p. 217]. Analysis outcomes for the three cohort subgroups including p-values are shown in Table S4.2 [p. 218].

An optimised linear regression model based on the study population (i.e. the healthy participants) resulted in prediction formulae of lung volume for women and men (Equation 4.2). The mean difference between the predictions and the TLV was -0.0155 L for women and 0.0003 L for men, indicating that the rounded parameters fit the data. The linear regression resulted in ΔLoA values of 3.1 L and 3.9 L, compared to the GLI-model a reduction of 1.2 % (women), and 6.8 % (men).

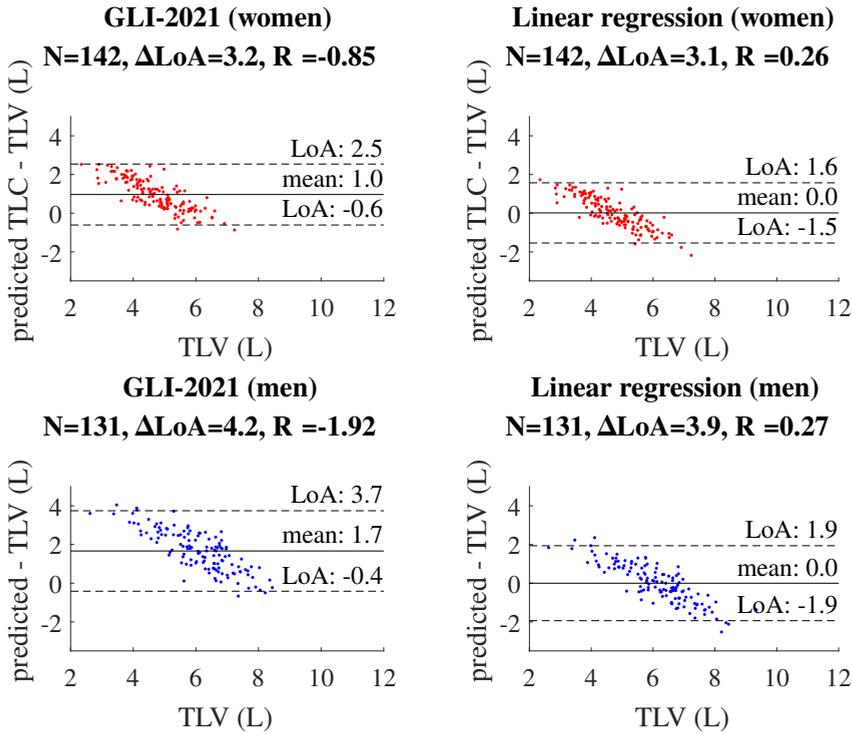


Figure 3: Dashed lines show limits of agreement, solid lines show the mean. GLI-2021: Global Lung Function Initiative prediction model; TLC: total lung capacity; TLV: total lung volume.

ΔLoA values were not equal between the linear regression and GLI model ($p > 0.259$).

$$\begin{aligned} V_{\text{women}} &= -7.296 + 7.554 * H - 0.010 * W - 0.000 * A \\ V_{\text{men}} &= -9.641 + 9.535 * H - 0.033 * W + 0.022 * A \end{aligned} \quad (4.2)$$

When applying the position-correction to the CT-measured TLV, the results did not meaningfully change. The ΔLoA increased slightly to 3.5/4.6 L compared to the original values of 3.2/4.2 L. The mean difference was reduced slightly from 1.0/1.6 L to 0.4/0.9 L.

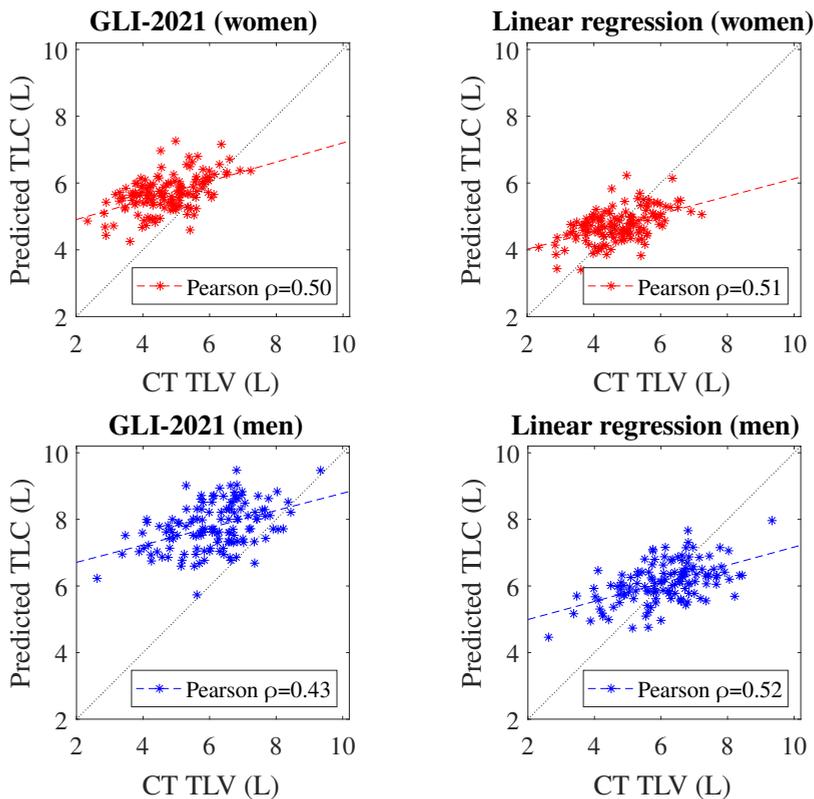


Figure 4: Dotted lines are the $y=x$ lines, dashed lines are linear trend lines. GLI-2021: Global Lung Function Initiative prediction model; TLC: total lung capacity; TLV: total lung volume.

4.4 Discussion

This study found a substantial mismatch between the predicted total lung capacity based on the recent GLI model and CT-measured lung volume. The GLI model tended to overestimate the lung volume compared to the actual, measured TLV, by 1.0 – 1.7 L (24 – 32%), with larger overestimation in individuals with a lower TLV. The Δ LoA was high (3.2 – 4.2 L), indicating low precision of the GLI model compared to TLV. When restricting the analyses to healthy never-smokers or expanding the analyses to include non-healthy participants, the precision and accuracy did not meaningfully change. This implies a prediction may not be sufficiently accurate or precise in clinical situations where true lung volume matters.

CT is an increasingly important modality in the evaluation of quantitative lung parameters [128]. There are suggestions that CT-derived parameters might be more sensitive than PFT measurements [144, 145]. Others have suggested that CT measurements are more reproducible than a body box [146]. This has led to the argument in a recent review by Bakker et al. that CT-derived parameters can,

now or in the future, replace some or all of the spirometry-based parameters [128]. The current study adds further evidence for this argument. When considering the difference between the GLI model predictions and actual CT-derived lung volume found in this study, there are several possible explanations.

Firstly, CT is normally acquired in supine position, while spirometry is performed in a sitting position; this in itself leads to a positional difference in lung volume. This is supported by the finding by Yamada et al., who compared supine and standing CT in healthy volunteers, and also reported sitting pulmonary function test measurements [133]. They reported that the mean lung volume measured in a supine position was 9.9% smaller than the mean lung volume measured in a standing position. The (unexplained) difference between standing TLV and sitting TLC was 7.5%. Since the difference between (supine) TLV and (sitting) TLC in the present study was 24 – 31%, this suggests only a proportion of the systematic bias may be due to the difference in position, but a third to half of the difference is likely due to an overestimation by the GLI model. Furthermore, the high variability cannot be explained by the positional difference. A sensitivity analysis based on the findings by Yamada et al. provided concordant results.

Secondly, there are technical differences between CT and spirometry. To compare CT-derived volumes with other types of measurements, it is important to be aware of the intrinsic differences between the body box measurement (or gas dilution) and the measurement on a CT scan. Normally, the CT volume measurement will exclude conducting airways, while the volume of these airways is included for body plethysmography [132]. However, since this difference would be approximately 20 mL (trachea only) up to 60 mL (full bronchial tree), it is not clinically relevant [147, 148]. It should furthermore be emphasised that this would only affect the systematic difference between CT and spirometry, and not the variability. The lung segmentation might include air pockets that are not actually ventilated (or exclude air pockets that are) due to imaging artifacts. This kind of segmentation issues should be rare in the absence of pathology and was not observed in this study.

Thirdly, pathology may influence the measurements. On the one hand, it may be difficult to reach maximal inspiration for patients with restrictive lung disease; on the other hand, there may be hyperinflation in patients with COPD. Garfield et al. compared body plethysmography to CT for a cohort of COPD GOLD 3 and 4 patients [134]. They found the TLV to be 17.3% lower than the measured TLC. As we excluded patients with COPD (based on spirometry) in our study, this did not play a role in the current results.

Despite the differences outlined above, the correlation between measured TLC and TLV is high (r 0.87 – 0.90), regardless of the TLC measurement method (body box or gas diffusion) [133–137].

As outlined by Hall et al., the differences in predicted TLC between different models are minor in the age range 45 – 65. Of the six TLC prediction models spanning this age, four are within 250 mL of each other [48]. Most prediction

models from the past decade (including the GLI-2012 and GLI-2021) make use of complex formulae, e.g. using logistic regressions with model parameters derived from splines in a separate lookup table [48, 143]. Despite this more mathematical approach, the GLI model did not result in a better fit for our study population than our linear regression. The reason is either the difference in population, or a difference in parameter choice. The cohort in this study includes participants with a smoking history and pulmonary pathology, as it is a population cohort. The linear regression includes weight as a parameter, while most other prediction models do not. Both of these differences could lead to a difference in model performance. However, we specifically performed analyses in healthy (i.e. no positive GOLD stages, no reported lung disease) and never-smoking subcohorts, to eliminate possible effects by pulmonary pathology and smoking history.

The main strength of this study is the cohort. As a sample from a population-based cohort, it matches the characteristics of the general population more closely than a hospital sample would. This is for instance important in early disease detection and monitoring and also particularly valuable in the context of lung transplantation donors where size does matter. The difference in age range between the current study population (45 – 65 years) and a hospital sample can be reasonably expected to be of lesser importance. This is because participants in early disease detection programs and candidates for lung donation tend to be younger than a typical hospital patient. For this study three different population types were used: a general population (cohort GP), healthy subjects (cohort H), and healthy never-smokers (cohort HNS). In general only never-smokers without pulmonary conditions (cohort HNS) have been used to develop prediction formulae [48, 142, 143]. This ignores the reality that a substantial proportion of the population are ever-smokers. Even among never-smokers there may be undiagnosed emphysema as found on CT [53]. In addition to this, there are non-pulmonary conditions that may affect the lung volume, like obesity or neuromuscular disease [149, 150]. The prediction formulae should be regarded as providing expected values for healthy never-smokers and may consequently not be accurately predicting normal values for more general populations.

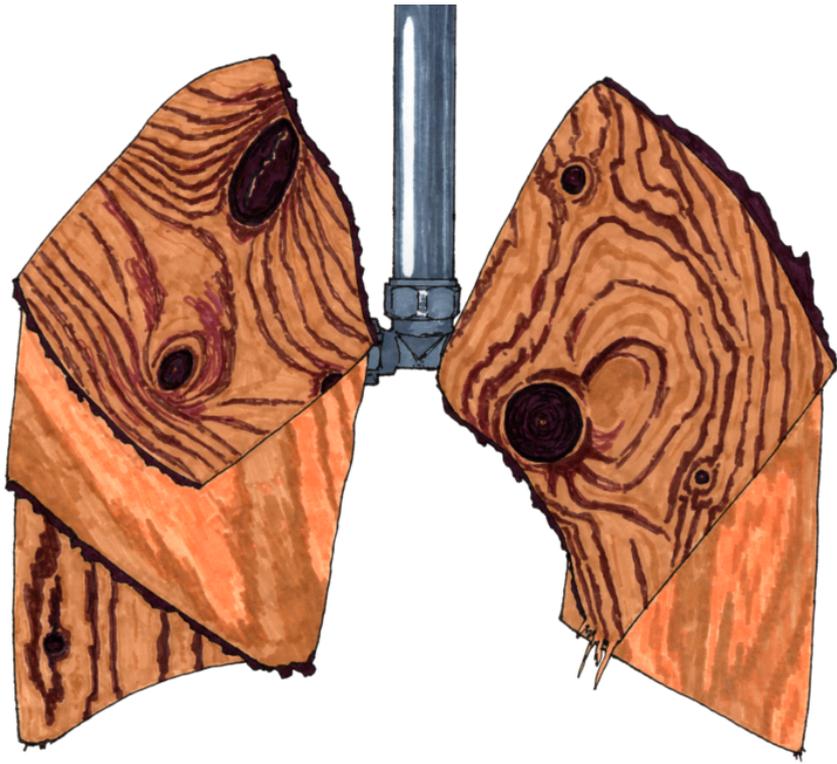
The assumption that our population matches the GLI-2021 population is both a strength and a limitation of this study. The difference in age range between this study (45 – 65) and the GLI-2021 population (5 – 80) is not expected to have a large impact, as the GLI model is reasonably linear in the age range 45 – 65 years. To mitigate this, a visual review (including fibrosis, emphysema, and incomplete inclusion of lungs on CT) was performed on subjects with a large difference between predicted and measured lung volume. In this review no obvious disease was found. Furthermore, sub-analyses were performed with only never-smokers without pulmonary disease, as well as with the general population sample. These sub-analyses did not yield a meaningfully different variability or systematic bias.

One limitation of this study is the use of the clinical standard breath coaching, which does not completely ensure full inspiration. The breath coaching during spirometry tends to more effectively ensure maximal effort. A further limitation is the lack of external validation for our linear regression model. The same cohort was used for the creation of the model and to test the performance of the model. This limits the generalisability. Combining this limitation with the particular height and weight of the study population, it would be interesting to repeat this study in a country with a different distribution of height and weight. Since the spirometry did not allow derivation of the TLC, it was not possible to compare a measured TLC with a measured TLV. A final limitation concerns the cohort size. While it is unlikely the results would substantially improve with a large cohort size, a large number of cases would increase confidence in the conclusions, especially in the case of the HNS cohort (never-smokers without pulmonary disease).

The current prediction models have a poor performance for lung volume as compared to actual measurements on CT in a general population cohort. Even without external validation (allowing for over-fitting of parameters), our linear regression only yielded a marginal reduction in variability of 1 – 7 %. Combining this with the inherent population spread, as evident from the data reported by Hall et al. [48], it does not seem likely that a model with easily obtainable parameters will be able to predict lung volume with reasonable precision [140]. Future research should evaluate the possibility of machine learning to assist in accurate and precise predictions, which should be tested in populations of different ethnicity. Moreover, future research should be aimed at exploring the potential value of CT derived lung volume and other parameters for lung disease detection and monitoring.

Conclusions and implications

This study found that there is a substantial mismatch between the GLI-predicted TLC and CT-derived TLV. The predicted TLC generally overestimates actual, measured lung volume, and has a high variability compared to TLV. A measurement (CT or otherwise) rather than a prediction should be performed in situations where size matters.



Chapter 5

Potential for dose reduction in CT emphysema densitometry with post-scan noise reduction: a phantom study

Published in the British Journal of Radiology.

DOI: 10.1259/bjr.20181019

Hendrik Joost Wisselink, Gert Jan Pelgrim,

Mieneke Rook, Maarten van den Berge, Kees Slump,

Yeshu Nagaraj, Peter M.A. van Ooijen,

Matthijs Oudkerk, Rozemarijn Vliegenthart

Abstract

Objective The aim of this phantom study was to investigate the effect of scan parameters and noise suppression techniques on the minimum radiation dose which still results in acceptable image quality for CT emphysema densitometry.

Methods The COPDGene phantom was scanned on a third generation dual-source CT system with 16 scan setups ($CTDI_{vol}$ 0.035 – 10.680 mGy). Images were reconstructed at 1.0/0.7 mm slice thickness/increment, with three kernels (one soft, two hard), filtered backprojection and three grades of third-generation iterative reconstruction (IR). Additionally, deep learning-based noise suppression software was applied. Main outcomes: overlap in area of the normalised histograms of CT density for the emphysema insert and lung material, and the radiation dose required for a maximum of 4.3 % overlap (defined as acceptable image quality).

Results In total, 384 scan reconstructions were analysed. Decreasing radiation dose resulted in an exponential increase of the overlap in normalised histograms of CT density. The overlap was 11-91 % for the lowest dose setting ($CTDI_{vol}$ 0.035 mGy). The soft kernel reconstruction showed less histogram overlap than hard filter kernels. IR and noise suppression also reduced overlap. Using intermediate grade IR plus noise suppression software allowed for 85 % radiation dose reduction while maintaining acceptable image quality.

Conclusion CT density histogram overlap can quantify the degree of discernibility of emphysema and healthy lung tissue. Noise suppression software, IR, and soft reconstruction kernels substantially decrease the dose required for acceptable image quality.

Advances in knowledge Noise suppression software, IR, and soft reconstruction kernels allow radiation dose reduction by 85 % while still allowing differentiation between emphysema and normal lung tissue.

5.1 Introduction

In industrialised countries, many people suffer from cardiovascular disease (CVD), lung cancer and chronic obstructive pulmonary disease (COPD). These three diseases are collectively referred to as the “big three”. CVD, lung cancer and COPD have high rates of mortality and morbidity [151–154]. Early detection combined with early treatment may reduce the disease burden, which has been shown for lung cancer detected by screening with low-dose computer tomography (CT) [155]. Screening for lung cancer has been introduced in the USA, and is under consideration in Europe [156]. A low-dose CT scan made for lung cancer screening may also allow evaluation of imaging biomarkers of emphysema [157]. CT-assessed emphysema has been linked to increased mortality in the MESA and COPDGene Study [43, 158]. Emphysema can be quantified by analysing the density of the lung parenchyma by measuring the Hounsfield units (HU) of lung voxels [157]. The underlying rationale is that destruction of alveolar walls and air trapping will result in an increased air content of lung tissue, lowering its density. HU density values correlate with pulmonary function test (PFT) and pathology results, the gold-standard for diagnosing COPD and quantifying emphysema [159, 160].

One of the main challenges in low-dose CT screening is to achieve adequate image quality, while limiting radiation exposure. Standardised phantoms may help to reliably compare imaging biomarker results that were obtained with different CT scanning and reconstruction methods. A phantom simulating COPD has previously been developed (the COPDGene phantom, see Figure 1) [161].

The quantitative imaging biomarkers alliance (QIBA) is developing a profile for quantifying lung density on CT. They aim to define what is sufficient “image quality”, meaning suitable for quantitative densitometry analysis. The most recent proposed maximum standard deviation (SD) for the CT density of the water insert and air insert is 20 HU [162]. They further propose that the deviation of the mean from the true value should be at the most 6 HU for water and air inserts. Therefore, the mean value for water and air should be between -6 and 6 HU, and between

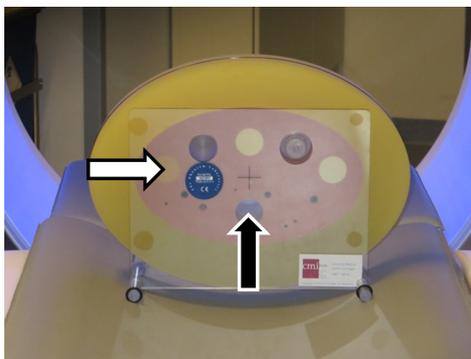


Figure 1: COPDGene phantom (CTP698). Materials used in this study: lung-like material (pink material surrounding the inserts), emphysema-like insert (left-most larger insert, white arrow), and air hole (hole in the lower centre, black arrow)

–1006 and –994 HU, respectively.

Noise reduction methods can be employed to allow dose reduction while still preserving adequate quality for visual reading and quantitative analysis, as decreasing radiation exposure will increase noise. One well-known method of reducing image noise is the use of iterative reconstruction. Another method of noise reduction is using a non-iterative technique artificial neural network (NiTANN) deep learning algorithm, trained with pairs of normal and low-dose CT scans. A NiTANN uses a complex arrangement of simple computational steps to achieve a mathematically defined goal, which in this case is to train the software to “reconstruct” a normal dose image from the low dose acquisition.

As stated in its FDA-clearance, the NiTANN used in this study can be used for the “processing and enhancement of CT images”. “It is specifically indicated for assisting professional Radiologists and Specialists in arriving at their own diagnosis.” [163] This product can be integrated in the normal workflow by adding a separate DICOM network node running the NiTANN software.

The aim of this phantom study was to study the effect of scan parameters and noise suppression techniques on the minimum radiation dose resulting in images that are suitable for CT emphysema densitometry.

5.2 Methods

Phantom and CT acquisition protocol

The COPDGene phantom was used [161]. It is approximately 35 cm wide, 25 cm high and 6 cm deep and contains inserts of different densities, one of which has a HU value low enough to simulate emphysema and has a reported density of –937 HU. The phantom also has an empty hole, simulating air trapping or bullae (Figure 1).

Scans were acquired using a third-generation dual-source CT system (SOMATOM Force, Siemens Healthineers, Forchheim, Germany) with 96×0.6 mm collimation, a pitch of 2.5, and a field of view diameter of 400 mm.

The selected kV values were 70, 100 (with and without Sn filter), and 120 kV, to cover the range of tube voltages in thoracic imaging. The effective tube current time products used were 10, 20, 30 mAs, as well as the maximum tube current setting that the system allowed for the selected kV, namely 260 mAs for 70 kV, 240 mAs for 100 kV with and without Sn filter, and 200 mAs for 120 kV. The maximum mAs scan was not made as a normal dose reference, but to determine the maximum quality for a given kV setting. The associated volumetric CT dose

Table 1: CTDI_{vol} in mGy for each combination of kVp and mAs.

	70 kVp	100 kVp (with Sn filter)	100 kVp (without Sn filter)	120 kVp
10 mAs	0.09	0.03	0.32	0.53
20 mAs	0.18	0.07	0.63	1.07
30 mAs	0.27	0.10	0.95	1.60
max mAs	2.34	0.83	7.60	10.68

CTDI_{vol}: computer tomography dose index (volumetric).

Max mAs is 260 mAs for 70 kVp, 240 mAs for 100 kVp (independent of tin filter), and 200 mAs for 120 kVp.

index (CTDI_{vol}) for each kV-mAs-combination is shown in Table 1.

Acquisitions were reconstructed with a slice thickness of 1 mm and a slice increment of 0.7 mm, with standard filtered backprojection (FBP), and with third-generation iterative reconstruction (IR) settings, advanced modelled iterative reconstruction (ADMIRE), grades 1, 3 and 5. Scans were reconstructed with a soft tissue kernel (Br40), a hard quantification kernel (Qr59), and a very hard kernel used for lung imaging (BI57). Further processing was done with NiTANN. In this study, the first FDA-cleared market version of PixelShine (AlgoMedica, Palo Alto, CA) was used (version 1.2.18). PixelShine is a noise-reduction algorithm that is based on deep learning. The training of NiTANN was performed with both phantom scans and human scans, and tested with human scan images (Algomedica technical staff, oral communication, June 2019). During normal use there will be no training, so the same input will always result in the same output [164]. This lack of training during use also means that the hardware requirements are much lower, resulting in a processing speed of several slices per second on a consumer-grade system [164].

Although integration of this software in the clinical workflow as a DICOM network node is possible, the processing for this study was performed on a separate laptop provided by the vendor.

In total, we acquired 384 reconstructions (4 kVp settings, 4 mAs settings, 3 kernels, 4 reconstruction options, and 2 options for PixelShine). The phantom was not moved between the different scans.

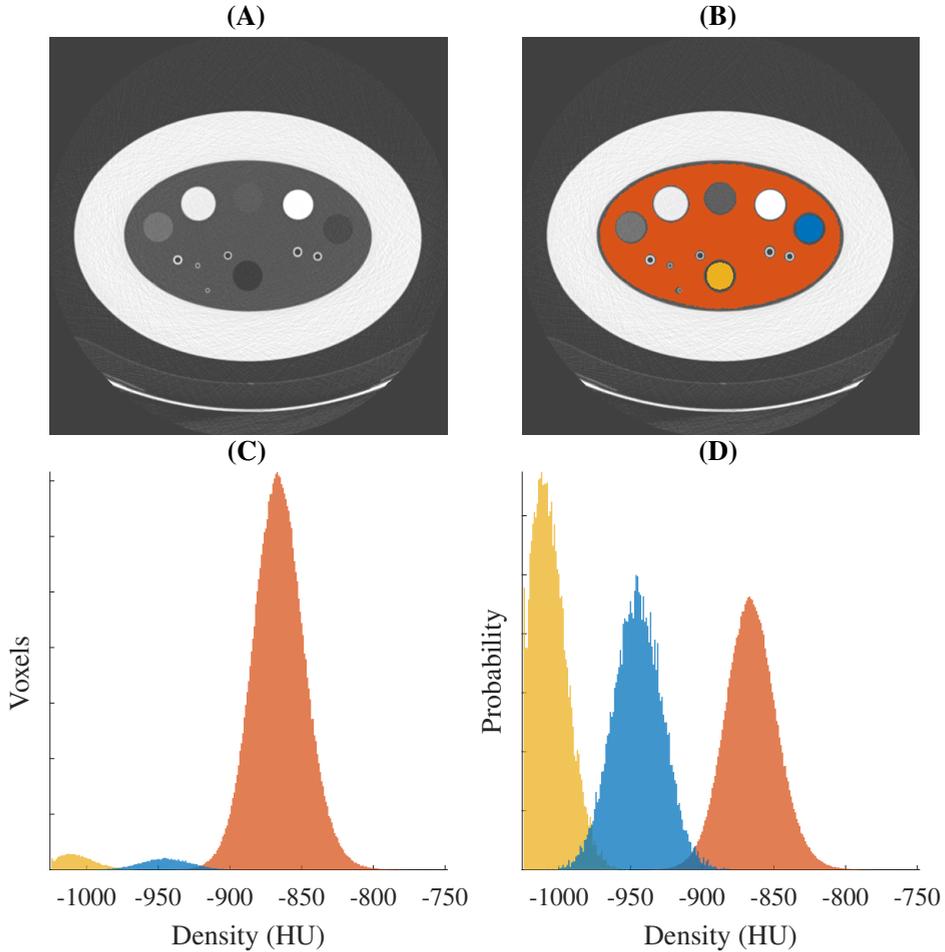


Figure 2: Steps of calculating overlap in density distributions. **(A)** Transverse CT image (120 kVp, 20 mAs, Br40, FBP, no NiTANN applied). **(B)** shows the same CT image with the LabelMap overlay (lung-like material in orange, air insert in yellow, emphysema-like insert in blue). **(C)** CT density histogram for voxels with lung, air and emphysema, same colour scheme as in (B). **(D)** Normalised histograms (i.e. the total area of each histogram was made the same). The overlap between emphysema and lung (blue and orange) was 2.5 % in this case.

Outcome metric development

Quantification of emphysema depends on distinguishing the voxels with emphysema from those with healthy parenchyma. The present study focuses on density analysis, by which emphysema can be differentiated from “normal” lung-like material if the CT density distributions do not overlap. Histograms of CT density were made of all voxels labelled as emphysema and as lung material. These two histograms were then normalised (i.e. divided by the total number of voxels). This enabled comparison of lung material and emphysema material, even though they had a different number of voxels. Next, the overlap in histogram distributions was calculated. The workflow of calculating the overlap in the phantom CT density histograms is described in Figure 2.

A simulation was performed to determine a threshold for acceptable overlap. The overlap was calculated for varying values of SD and differences of mean value between the simulated density histograms. Both distributions were assumed to be normal distributions for this calculation. The result of this simulation is shown in Figure 3. At 81 HU separation (the separation between the values for the lung and emphysema inserts reported by Newell et al.: -937 HU and -856 HU) and a SD at the limit proposed by QIBA (20 HU), the overlap was calculated at 4.3 % [161, 162]. This percentage was then considered the upper cut-off value for acceptable overlap in the remainder of this study.

Graphical user interfaces

Two graphical user interfaces (GUIs) were developed to provide more intuitive insight into the effects of the chosen parameters. These tools were used to visually compare scans, and to assess the effect of each parameter on the density distributions. Screenshots of these tools are shown in Figure 4 and Figure 5. Interactive versions are available as supplementary material.

Data analysis

Data processing and characterisation were performed with MATLAB R2018b [165]. A labelled mask was generated to enable consistent CT density analysis. To create the mask, the water in the bottle, the emphysema insert and the inside air from the phantom were segmented. The mask was based on the physical dimensions reported in the manual, and was created from the scan with the highest dose in combination with the highest IR setting [166]. To avoid partial volume effects, the edges of each volume of interest (VOI) were discarded. This was done using a morphological erosion with a spherical structuring element with a radius of two voxels. The eroded cylindrical VOIs had a diameter of approximately 27 mm.

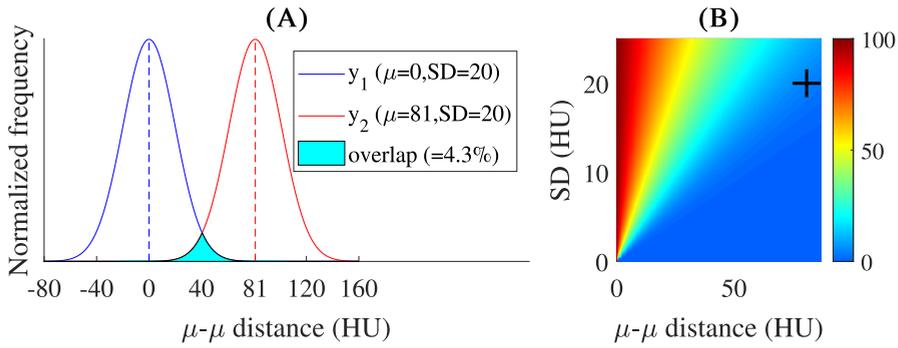


Figure 3: Results of overlap simulation. (A) Simulated overlap percentage calculation was based on two normal distributions with equal SD and a specific distance between their means. The image shows the overlap as the filled cyan area. In this example the separation is 81 HU (mean difference between emphysema insert and lung material [161]) and SD is 20 HU (upper limit suggested by the QIBA [162]). (B) Three-dimensional plot that shows the histogram overlap for each combination of SD and $\mu-\mu$ distance. The crosshair marks the case of the A part.

SD: standard deviation; HU: Hounsfield unit; QIBA: quantitative imaging biomarker alliance.

The radiation dose was correlated to the overlap percentage using an exponential function. The exponential function was fit to the dose-overlap data by fitting a linear function to the dose against the logarithm of overlap. This function was then intersected with the 4.3% threshold (based on the previously mentioned simulation) to determine the minimally required dose to reach the QIBA recommendations. For an example, see Figure 6.

The image noise was defined as the SD of the HU values within each material.

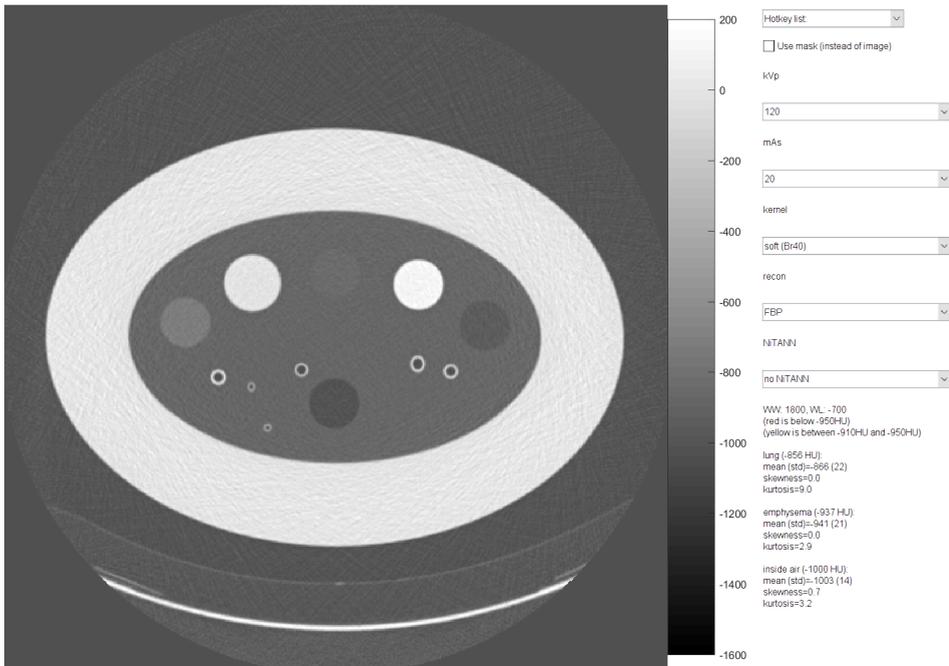


Figure 4: Layout of the user interface used to assess the visual differences caused by changing acquisition parameters and post-filtering parameters. The top drop-down menu can be used to change several parameters at once. The check box can be used to switch between normal view and mask view. In the mask view, voxels with a density below -950 HU are marked red, and all voxels with a density between -910 HU and -950 HU are marked yellow. The window level setting is adjusted by dragging, and the setting is shown in the text area.

HU: Hounsfield unit.

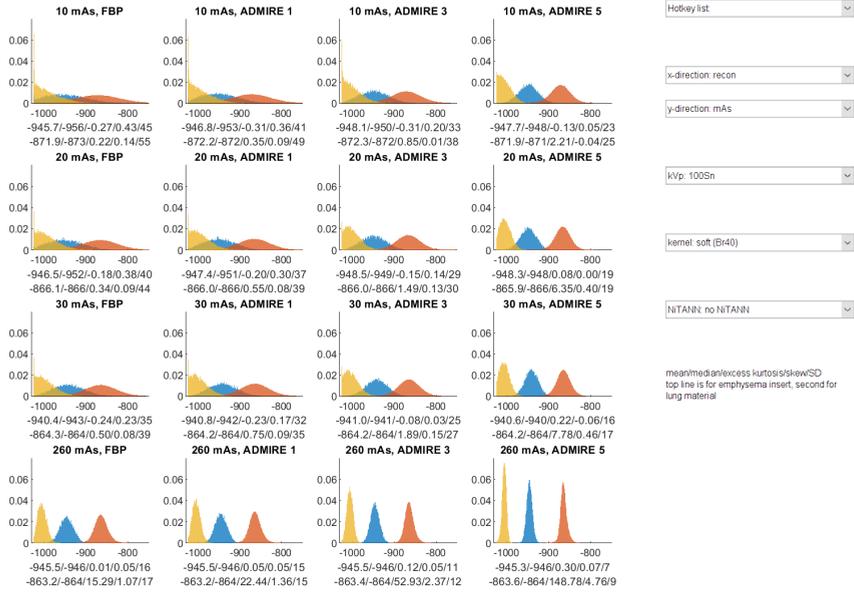
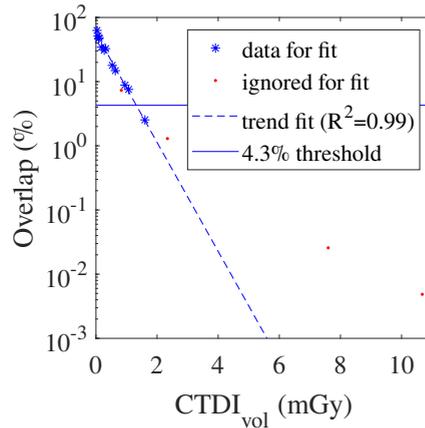


Figure 5: Layout of the user interface used to view the histogram characteristics. The “x-direction” drop-down menu controls which parameter is varied between columns, the “y-direction” drop-down menu controls which parameter is varied between rows. The colours of each histogram correspond to the material: yellow is for the air inside the phantom, blue is for the emphysema insert, orange is for the lung material. The shown histograms are normalised, meaning that their total area is 1.

Figure 6: Percentage of overlap between the CT density histograms of lung material and emphysema insert plotted against the CTDI_{vol} (for this example, data from the Br40, FBP, no NiTANN scan was used). Maximum mAs setting for each kV was ignored for the fit. Fit parameters and R^2 were calculated with the log of the overlap. CTDI_{vol} : volumetric CT dose index; FBP: filtered backprojection; NiTANN: non-iterative technique artificial neural network.



5.3 Results

An overview of the normalised histogram data of CT density is available in the GUIs, which are available on <http://tiny.cc/QUykRbrc> (usage instructions included). From the histogram GUI it becomes clear that a softer kernel resulted in much less overlap for the same dose, while visual inspection with the other GUI showed an image with more noise for hard kernels. Four example slices can be found in Figure 7.

The kVp had a small effect on the HU values: approximately 0.5 % for the described levels (Figure 8). This means that mAs and kV can be considered together to determine the dose effects. Decreasing the dose exponentially increased the measured SD (Figure 9). A decrease in dose from 10 to 8 mGy resulted in only a minor change in SD, while a decrease from 2.5 to 0.5 mGy resulted in a tripling of the SD. All combinations of kernel, IR and NiTANN settings showed a similar trend as shown in Figure 9B.

ADMIRE and NiTANN decreased the image noise (Figure 10). The relative decrease for ADMIRE was approximately the same for air, emphysema and water inserts, while the reduction in SD by NiTANN was more profound for air and emphysema, i.e. for the very low-density inserts. For the air insert, NiTANN had approximately the same effect on image noise as ADMIRE 5, while for the lung and water inserts the effect was in-between ADMIRE 3 and 5. Neither ADMIRE or NiTANN caused a substantial shift in median HU.

The correlation between dose and overlap percentage is shown in Figure 6. An exponential function was fitted to the data and then intersected with the horizontal line. This horizontal line denotes an overlap of 4.3 %, which is the maximum overlap allowed when conforming to the QIBA profile. The dose that is required for acceptable imaging to be able to have no more than 4.3 % overlap in CT density histograms between emphysema material and lung-like material, is shown in Table 2 for all combinations of ADMIRE and NiTANN. The tabulated values are the values found for the intersection of the trend line and the threshold in Figure 6. Each value is based on 12 scans (4 kV levels and 3 mAs levels, as the maximum mAs was ignored for the trend line fit). This table shows that ADMIRE and NiTANN both allow a substantial reduction in the minimal dose required to conform to the quality standard suggested by the QIBA profile.

Table 2: $CTDI_{vol}$ in mGy required to comply with the QIBA profile, meaning that the expected SD is at most 20 HU [162]. This should allow differentiating healthy from emphysematous lung tissue.

	No NiTANN	NiTANN
FBP	1.32	0.48
ADMIRE 1	1.07	0.39
ADMIRE 3	0.66	0.19
ADMIRE 5	0.25	≤ 0.07

ADMIRE: advanced model iterative reconstruction; $CTDI_{vol}$: computer tomography dose index (volumetric); FBP: filtered backprojection; HU: Hounsfield unit; NiTANN: non-iterative technique artificial neural network; QIBA: Quantitative Imaging Biomarkers Alliance.

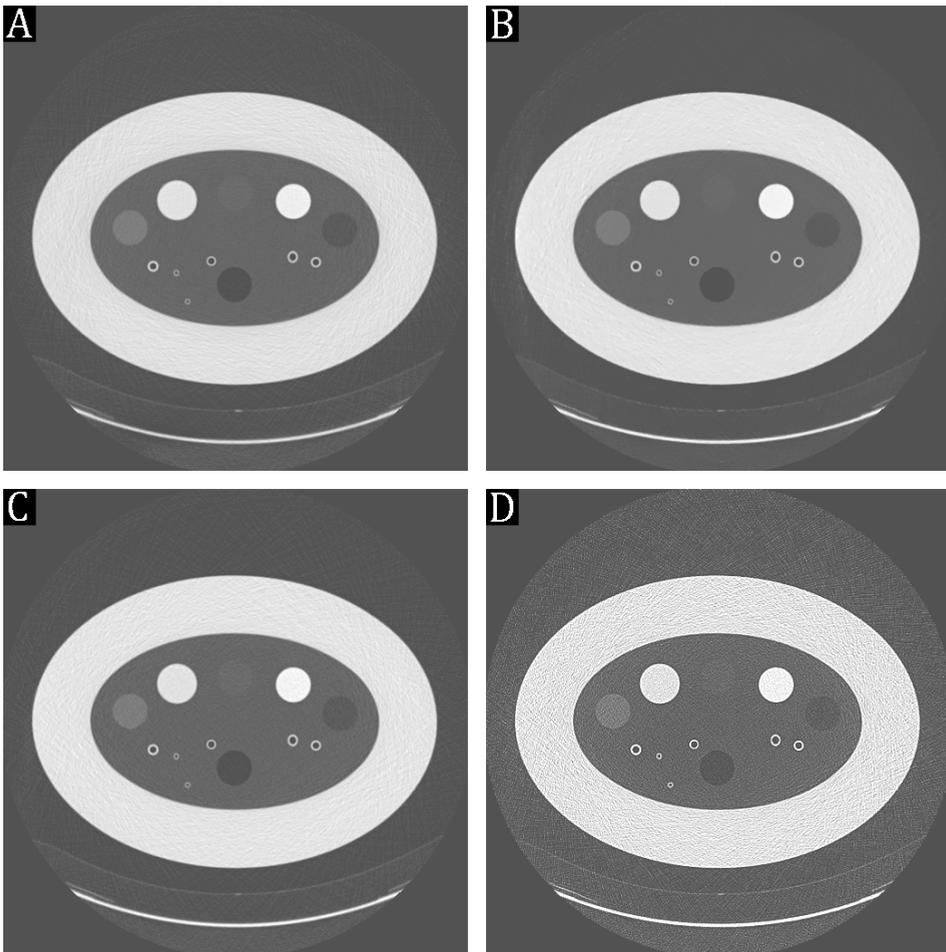


Figure 7: Example slices for different settings (shown at WW1800WL-700). (a) is an 86% lower dose than (b), with ADMIRE 3 and NiTANN to reduce noise. (c) is 120 kVp, 20 mAs with a soft kernel, (d) is with a hard kernel. ADMIRE: advanced model iterative reconstruction; NiTANN: non-iterative technique artificial neural network; WL: window level; WW: window width.

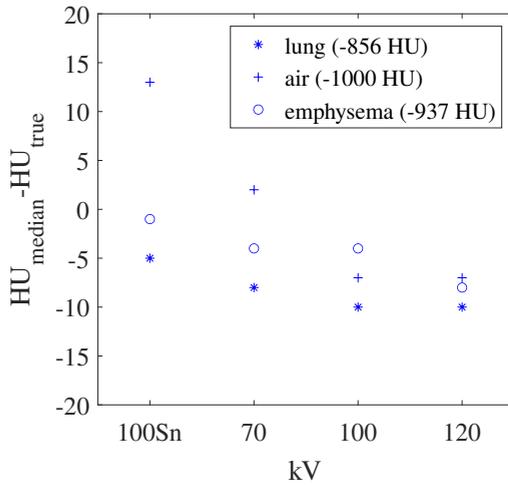


Figure 8: Difference between median HU of each insert and true value (see Newell et al. [161]), points shown here are from the 30 mAs scans. To prevent mixing of effects in this example, no ADMIRE or NiTANN data was used.

HU: Hounsfield unit; FBP: filtered backprojection; ADMIRE: advanced model iterative reconstruction; NiTANN: non-iterative technique artificial neural network.

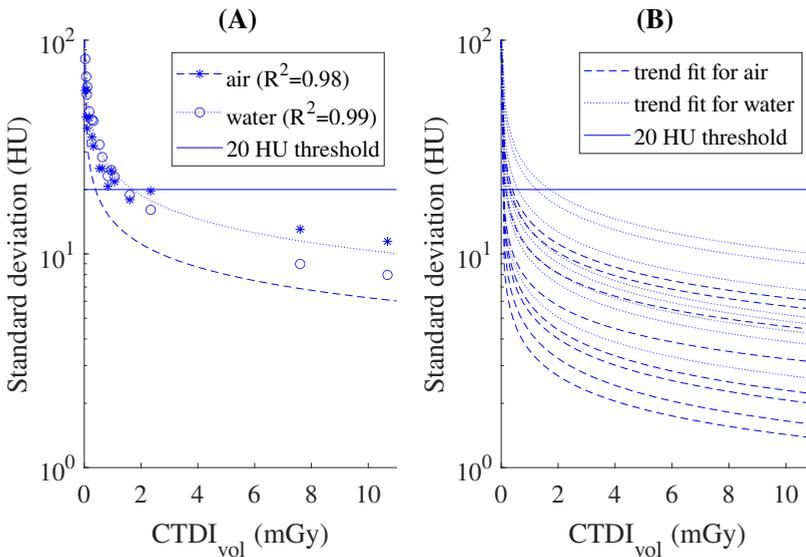


Figure 9: (A) Measured standard deviation of air and water plotted against CTDI_{vol} (data used as example: Br40, FBP, no NiTANN). The threshold is 20 HU (QIBA threshold for air and water inserts [162]). (B) All trend lines for soft kernel (different ADMIRE levels and with/without NiTANN). R^2 values of the fits range: 0.96 – 0.99 (median 0.98).

ADMIRE: advanced model iterative reconstruction; CTDI_{vol} : volumetric CT dose index; FBP: filtered backprojection; HU: Hounsfield unit; NiTANN: non-iterative technique artificial neural network; QIBA: quantitative imaging biomarkers alliance.

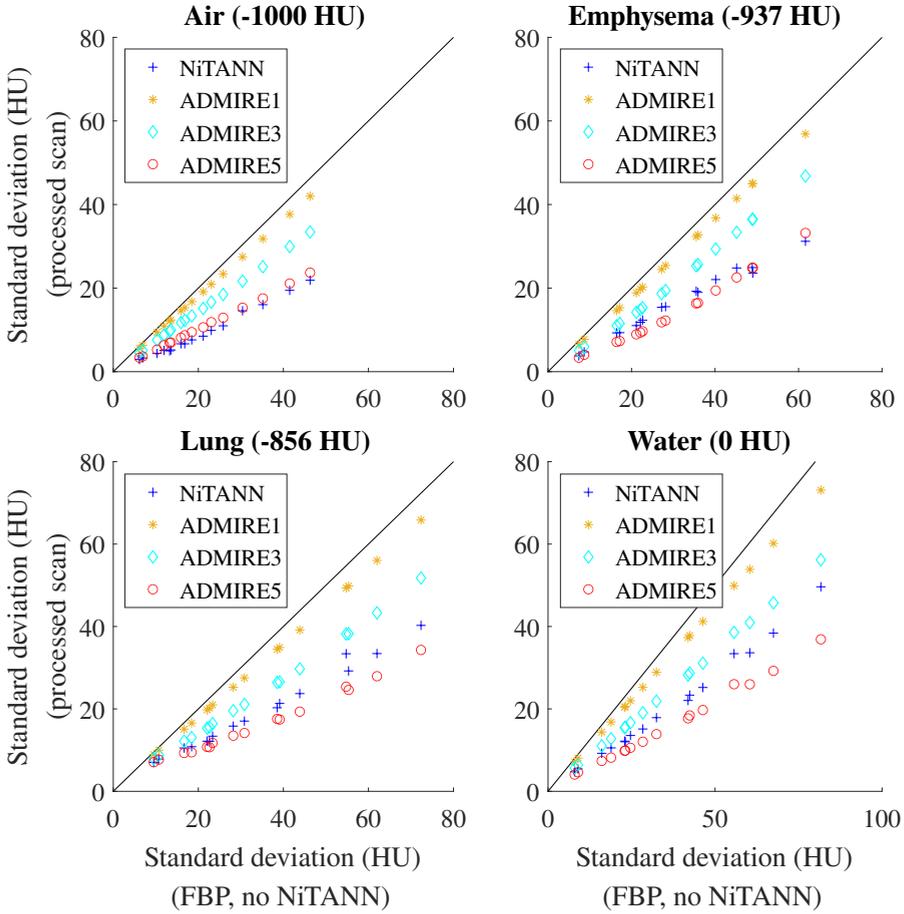


Figure 10: Effect of ADMIRE and NiTANN on SD of the density distribution of air insert (upper left), emphysema material (upper right), lung material (lower left) and water (lower right). Only soft kernel scans were used for these plots; hard kernels showed similar results, but with a wider range of SD values. The black line shows the equality line, so SD values to the right of the line are lower in the processed scan than in the unprocessed scan.

ADMIRE: advanced model iterative reconstruction; NiTANN: non-iterative technique artificial neural network; SD: standard deviation.

5.4 Discussion

The aim of this phantom study was to study the effect of scan parameters and noise suppression techniques on the minimum radiation dose that results in reconstructed images that are suitable for CT emphysema densitometry.

This study showed that decreasing CT dose increases histogram overlap, which could be mitigated by using (higher levels of) IR and/or NiTANN. The use of moderate level IR (e.g. ADMIRE 3) allowed 50% reduction in dose without loss of power to differentiate between emphysematous and lung-like material in this phantom study. The use of NiTANN allowed at least 64% dose reduction compared to the standard FBP reconstruction. It is important to note that these results are only applicable to emphysema densitometry. For other evaluations and applications such as measurement of bronchopathy, lung nodules, or coronary calcium, this low dose might result in inadequate images.

A commonly used method of quantifying pulmonary emphysema on CT relies on the assumption that emphysematous tissue has a CT density below -950 HU. So, to quantify for emphysema, the amount of tissue with CT density below this threshold is calculated. However, the emphysema insert in the COPDGene phantom has a homogeneous density above -950 HU, so this thresholding method cannot be used to calculate sensitivity and specificity. Based on the requirements suggested in the QIBA profile draft, the overlap percentage was used as a proxy for emphysema densitometry performance, as described in the methods section [162].

In this study, different scan parameters and reconstruction possibilities which potentially allow reducing CT dose while maintaining discernibility between normal and emphysematous lung-like materials were studied.

The small shift caused by changing kVp will generally have little effect, but it is nonetheless important to keep in mind when comparing quantifications based on scans with slightly different scan parameters. Adding the Sn filter for 100 kV approximately lowers the dose by a factor of 10, while lowering the kV to 70 results in an approximate dose reduction by a factor of 3. The mAs has a linear effect on the dose, but the kV does not.

Hard kernels are generally not recommended for emphysema quantification, due to a flat density distribution [167]. This means that a single material density will result in a wide spread of HU values. With high emphysema thresholds ($-900/-890$ HU), hard kernels yield the same results as soft kernels, but this is not the case for more usual thresholds such as -950 HU [168]. Using a high threshold would then seem a good solution for making the quantification method more robust. However, choosing a high threshold will result in more healthy tissue being marked as emphysema, reducing the specificity. Therefore, we propose a low

threshold (like -950 HU) with a soft kernel (like Br40) for emphysema densitometry.

The effect of iterative reconstruction on emphysema quantification has been extensively studied. A frequent conclusion is that IR reduces the measured emphysema index compared to FBP. This is thought to be especially true for low-dose CT, although this difference was not always statistically significant [157, 169–173]. Lower radiation dose leads to higher emphysema index values and higher SD, which can be mitigated with iterative reconstruction [167, 170, 173].

Our results suggest that applying NiTANN to a CT scan may increase its quality substantially. Very little is known about the effects of PixelShine on CT. Cross et al. performed a study in which 10 CT images were sent to radiologists in a survey (five low-dose, five low-dose with PixelShine) [174]. 75 % of the respondents classified the processed images as being acquired with a standard dose protocol [174]. This suggests that NiTANN has a large potential for allowing dose reduction without adversely affecting the visual quality. Although it was partially trained with phantom data, PixelShine is intended for the processing of human CT data, and it is therefore unknown whether the effects in our phantom study are comparable to the effects in human data. When indeed any difference is found, it is to be expected that PixelShine will perform better on human data than on phantom data, meaning that an even larger dose reduction might be possible. Recently, the use of PixelShine was shown in pelvic CT in 33 patients [164]. In that study, the use of NiTANN lowered the image noise by 30 % and increased the signal to noise ratio by 58 %.

ADMIRE and the NiTANN provide the potential for substantial dose reduction, down to the dose level of a two-direction chest radiograph. This finding agrees with De Margerie-Mellon et al., who compared a standard-of-care CT to a reduced dose protocol with different types of IR [175]. It should be noted their study shares the potential dose underestimation because of low body weight study participants.

One of the strengths of this study is the standardised analysis of several CT parameters influencing the quantification of emphysema. This makes it easier to compare the newly tested NiTANN with more common methods of influencing image noise. Furthermore, the use of the graphical user interface did not result in numerical outcomes, but did substantially contribute to the understanding of the effect each parameter has on visual image quality and on the CT density distributions of each insert. A weakness of this study is that the phantom does not mimic a typical western body habitus. This likely results in an underestimation of the minimum required dose. It is worth noting that QIBA proposes the use of a slightly different phantom to the one used in this paper, although this is not expected to have a substantial effect, as the materials and shape are very similar.

This study describes an objective method of determining the possible dose reduc-

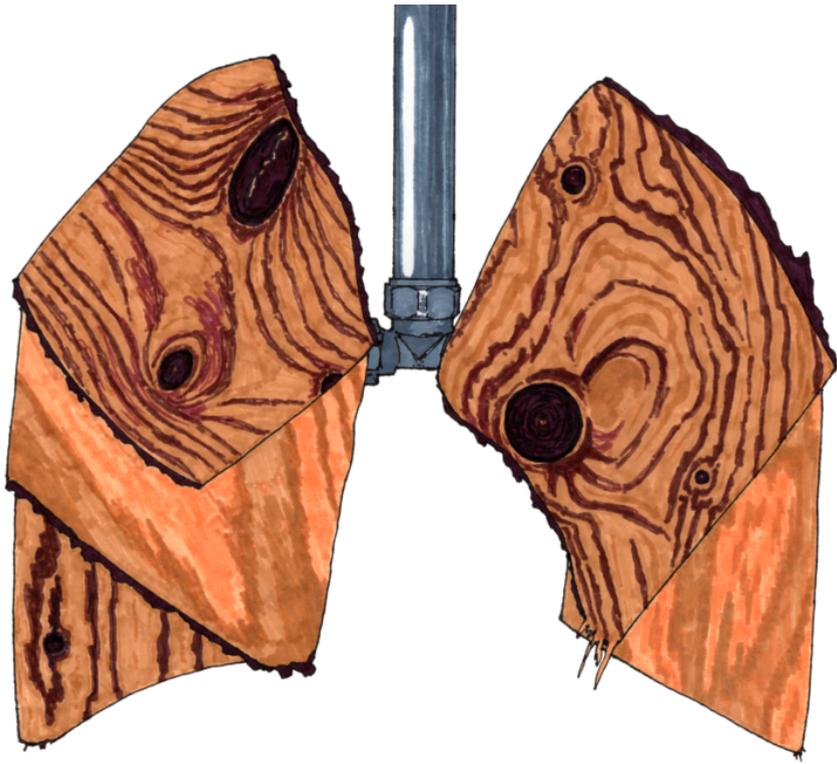
tion and the minimum dose required for CT densitometry in emphysema estimation. This is especially important for the clinical implementation of a low-dose chest CT screening programme, which has been introduced for lung cancer in the USA, and is under consideration for lung cancer in Europe [156]. Lung cancer screening provides the opportunity to simultaneously evaluate the presence and extent of emphysema, which gives additional information about its prognosis. Based on the requirements described in the QIBA profile draft, our study suggests that modern CT systems with new iterative reconstruction techniques can yield images that are acceptable for quantitative emphysema evaluation with substantially lower dose than normal dose levels [162]. The presented results are very promising for densitometry-based automated analysis of lung parenchyma on low-dose CT, but the assessment of emphysema and other thoracic diseases do not solely depend on densitometry. High levels of denoising may potentially remove structural information, which could make the scan quality insufficient for reading by a radiologist. Structural information is also very important for correct segmentation of the blood vessels and airways, which may also be quantified. Future research should focus on assessing what level of denoising yields acceptable images in human CT scanning.

5.5 Conclusion

The aim of this phantom study was to investigate the effect of scan parameters and noise suppression techniques on the minimum radiation dose that results in reconstructed images that are suitable for CT emphysema densitometry. Reducing the dose reduced discernibility of emphysema and healthy lung tissue.

A soft reconstruction kernel yielded markedly better results than harder kernels.

ADMIRE reduced image noise. Using NiTANN and/or ADMIRE substantially decreased the dose required to obtain low-dose CT that can differentiate between emphysematous and normal lung tissue.



Chapter 6

Ultra-low-dose CT combined with noise reduction techniques for quantification of emphysema in COPD patients: An intra-individual comparison study with standard-dose CT

Published in the European Journal of Radiology.

DOI: 10.1016/j.ejrad.2021.109646

*Hendrik Joost Wisselink, Gert Jan Pelgrim,
Mieneke Rook, Kai Imkamp, Peter M.A. van Ooijen,
Maarten van den Berge, Geertruida H. de Bock,
Rozemarijn Vliegenthart*

Abstract

Purpose Phantom studies in CT emphysema quantification show that iterative reconstruction and deep learning-based noise reduction (DLNR) allow lower radiation dose. We compared emphysema quantification on ultra-low-dose CT (ULDCT) with and without noise reduction, to standard-dose CT (SDCT) in chronic obstructive pulmonary disease (COPD).

Method Forty-nine COPD patients underwent ULDCT (third generation dual-source CT; 70 ref-mAs, Sn-filter 100 kVp; median $CTDI_{vol}$ 0.38 mGy) and SDCT (64-multidetector CT; 40 mAs, 120 kVp; $CTDI_{vol}$ 3.04 mGy). Scans were reconstructed with filtered backprojection (FBP) and soft kernel. For ULDCT, we also applied advanced modelled iterative reconstruction (ADMIRE), levels 1/3/5, and DLNR, levels 1/3/5/9. Emphysema was quantified as Low Attenuation Value percentage (LAV%, ≤ -950 HU). ULDCT measures were compared to SDCT as reference standard.

Results ULDCT, the median radiation dose was 84% lower than for SDCT. Median extent of emphysema was 18.6% for ULD-FBP and 15.4% for SDCT (inter-quartile range: 11.8 – 28.4% and 9.2 – 28.7%, $p = 0.002$). Compared to SDCT, the range in limits of agreement of emphysema quantification as measure of variability was 14.4 for ULD-FBP, 11.0 – 13.1 for ULD-ADMIRE levels and 10.1 – 13.9 for ULD-DLNR levels. Optimal settings were ADMIRE 3 and DLNR 3, reducing variability of emphysema quantification by 24% and 27%, at slight underestimation of emphysema extent (–1.5% and –2.9%, respectively).

Conclusions Ultra-low-dose CT in COPD patients allows dose reduction by 84%. State-of-the-art noise reduction methods in ULDCT resulted in slight underestimation of emphysema compared to SDCT. Noise reduction methods (especially ADMIRE 3 and DLNR 3) reduced variability of emphysema quantification in ULDCT by up to 27% compared to FBP.

6.1 Introduction

Lung tissue densitometry is a common method for quantifying emphysema on CT in patients with chronic obstructive pulmonary disease (COPD) [40, 42, 157, 170]. These CT-derived estimates of emphysema severity correlate well with pulmonary function test results, pathology results, and mortality rates [43, 158, 159, 176, 177]. CT is often used to monitor COPD progression, assess causes of COPD exacerbations, and to assess bronchiectasis [41]. Thus, cumulative radiation exposure resulting from standard-dose CT (SDCT) scans in COPD patients throughout their lifetime can be considerable.

Of late, several studies have focused on the feasibility of quantifying emphysema based on ultra-low-dose CT (ULDCT, < 1 mSv). Dose reduction increases image noise and can thus negatively affect image quality. Specifically for quantitative emphysema analysis high levels of noise lead to misclassification of voxels as either emphysematous or healthy [9]. Methods are being investigated to reduce CT image noise to levels similar to SDCT [157, 170, 172].

Iterative reconstruction (IR) is a method often used to reduce noise to an acceptable level for clinical decision making, at the cost of affecting the noise texture and spatial resolution [171, 178, 179].

Recently, a fundamentally different method of noise reduction has become available: deep learning-based noise reduction (DLNR) [180, 181]. Deep learning can either be employed to reconstruct the image from the raw data, or to reduce noise on an already reconstructed DICOM image [180–183]. A recent phantom study suggests that both IR and DLNR allow for substantial dose reduction in CT for emphysema quantification [9]. IR and DLNR generally remove high spatial frequencies, which reduces both image noise and structure detail [180]. The decrease in detail may reduce the differentiation between emphysema and healthy lung tissue by blurring the image. It is likely that there is an optimal setting that removes a substantial part of the noise but allows structural details to remain mostly visible in the image, allowing accurate quantification of emphysema. The aim of this study was to compare emphysema quantification on ULDCT with and without state-of-the-art noise reduction techniques to SDCT in COPD patients.

6.2 Materials and methods

Patient cohort

In an on-going treatment study in COPD patients, patients underwent a non-contrast high-resolution chest CT scan (SDCT). Inclusion criteria for this study were age 40 – 80 years, smoking history >10 pack-years, and spirometry-confirmed COPD. Patients with asthma were excluded. For the current sub-study,

Table 1: CT scan parameters.

Parameter \ Protocol	Standard dose	Ultra-low-dose
Scanner model	SOMATOM Definition AS, Siemens Healthineers	SOMATOM Force, Siemens Healthineers
Tube current-time product	40 mAs (fixed)	70 mAs (ref)
Tube potential	120 kVp	100 kVp
Spectral shaping	None	Tin filter
Scanner pitch	1.5	1.6
Slice thickness	1.0 mm	1.0 mm
Slice increment	0.7 mm	0.7 mm
Kernel	Smooth (B30f)	Smooth (Br40)
Field of view	317 – 450 mm	346 – 500 mm

50 consecutive participants who were scanned from February 2018 to June 2018 additionally underwent ULDC. The order (i.e. whether the ULDC was acquired first or the SDCT first) was randomised between participants, and the two scans were made within 30 minutes of each other. Prior to the first scan, 100 μg Sabutamol was administered via inhalation as part of the COPD treatment study protocol. The institutional ethical board gave approval for this study, and participants provided written informed consent (METC 2015/335, clinicaltrials.gov NCT02477397). The sample size for this study was based on the cohort size of prior studies comparing emphysema in ULDC to chest CT [157, 173, 178, 184, 185]. One participant was excluded due to an inspiration issue during acquisition leading to a 3 L difference in lung volume between the two acquisitions. The severity of COPD was graded according to the GOLD 2017 guidelines [186].

CT scans

The high-resolution SDCT scans (CTDI_{vol} 3.04 mGy) were acquired, according to standard clinical protocol. SDCT involved fixed mAs, conform the protocol used in the COPDGene and SPIROMICS studies [187, 188]. The ULDC was acquired with automatic exposure control enabled to ensure sufficient and uniform image quality despite the very low radiation dose (median CTDI_{vol} 0.39 mGy, range 0.19 – 1.34 mGy). The field of view was adapted for each individual participant. A more detailed list of acquisition and reconstruction parameters can be found in Table 1.

All scans were reconstructed with filtered backprojection (FBP). For the ULDC scans, additional reconstructions were performed at advanced modelled iterative reconstruction (ADMIRE) levels 1, 3, and 5 (Siemens Healthineers). Deep

learning-based noise reduction (DLNR) processing was based on the ULD-FBP reconstructions. DLNR (PixelShine v1.2.102.07, Algomedica) processing was performed with levels 1, 3, 5, and 9. The ADMIRE and DLNR levels were chosen to analyse the full spectrum of settings while limiting the number of scans to be analysed.

Analysis

To measure noise, the standard deviation of Hounsfield units (HU) of air in the trachea was measured in a circular region of interest about 1 cm above the carina with an area of 1 cm². The same voxels were measured for the different reconstructions, and a visual check was performed to confirm that the tracheal wall was not included in the measurement.

A trained technical physician (HJW), supervised by a radiologist (MR, 3 years of post-residency experience in chest radiology), performed visual emphysema assessment. Technical physicians are well-trained dedicated technicians with a medical background and the supervision during this study consisted of review and consensus read on request. The visual assessment was used to describe the population, and was therefore only recorded for the reference images (SDCT). The scoring was performed according to the Fleischner criteria [40].

Low attenuation value percentages (LAV%) and lung volumes were measured using fully automated analysis software (Syngo.Via Pulmo3D, Siemens Healthineers) with the default threshold set at -950 HU [157, 189]. A screenshot of this software is available in the supplement as Figure S6.1 [p. 221]. As differences in segmented lung volume can alter the total measured lung volume and possibly influence the LAV, an incorrect segmentation could alter the emphysema extent. The automated segmentation was visually checked for errors by a technical physician (HJW) to prevent this potential bias. All further analyses and data processing were performed with MATLAB R2018a (MathWorks, Natick, Massachusetts, USA). Sub-analyses were performed for participants with a high (≥ 30), medium (25-30), or normal/low (≤ 25) body mass index (BMI).

To determine the effect of CT dose setting on LAV%, two potential sources of bias were analysed. First, as described, an incorrect segmentation could lead to an incorrect value of LAV%. This was ruled out by visual inspection of the segmentations. Second, as each patient was scanned two times the scans may have been performed at slightly different inspiration levels. A difference in inspiration levels could introduce a difference in measured emphysema extent, as deeper inspiration lowers the density of lung tissue. Comparing the ratio of lung volume (of the ULDC and SDCT scan) and the difference in LAV% (of the same pairs) provides an indication of the occurrence of this phenomenon. Thus, we analysed the lung volume of both CT scans, based on FBP reconstruction, and

then compared the difference in LAV% to lung volume ratio (i.e. by comparing $LAV\%_{ULDCT} - LAV\%_{SDCT}$ and V_{ULDCT}/V_{SDCT}).

Statistical analysis

For this study we defined systematic bias as the structural difference in measurements between two measurement methods. Variability describes how far measured values tend to deviate from the true value, based on SDCT as reference method.

Bland-Altman analysis was performed to compare the ULDCT lung volume and LAV% measurements to the SDCT measurements and derive systematic bias and limits of agreement. Since SDCT was reference standard for LAV% measurements, the results are shown in a residual plot with the SDCT measurement plotted against the difference between LAV% for the ULDCT and SDCT value. The distance between the limits of agreement (ΔLoA) for emphysema quantification was taken as an indicator of variability between ULDCT and SDCT. Because the ΔLoA is related to the variance, Levene's test was used to test whether the ΔLoA for the ADMIRE/DLNR reconstructions was significantly different from the ULD-FBP ΔLoA . The Wilcoxon signed rank sum test was used to test absolute differences. Normality of continuous variables was tested with the Shapiro-Wilk test. MATLAB R2018a (MathWorks, Natick, Massachusetts, USA) was used for the statistical analysis.

Table 2: Patient characteristics (N=49). Values are given as number (%), mean (SD), or median (25th–75th percentile)

Characteristic	Value
Age (years)	65.3 (7.4)
Male sex	33 (67 %)
Body mass index (kg/m ²) [†]	27.3 (25.3–31.4)
Smoking history (pack-years) [†]	36.0 (27.5–59.5)
FEV ₁ (L)	1.6 (0.5)
FEV ₁ %predicted	53 (16)
FVC (L)	3.8 (1.0)
FVC %predicted	95 (19)
Tiffeneau index (FEV ₁ /FVC)	0.43 (0.11)
TLC (L)	7.5 (1.6)
GOLD stage (number of patients)	I: 0, II: 30 (61 %), III: 14 (29 %), IV: 5 (10 %)
CTDI _{vol} (mGy), standard dose CT	3.04
DLP (mGy·cm), standard dose CT	Overall: 104.3 (9.5) BMI ≤25: 110 (11.4) BMI 25–30: 103.8 (97.8 – 107.9) BMI ≥30: 101.2 (9.1)
CTDI _{vol} (mGy), ultra-low dose CT	0.39 (0.35–0.53)
DLP (mGy·cm), ultra-low dose CT	Overall: 16.6 (12.2 – 20.7) BMI ≤25: 11.4 (3.0) BMI 25–30: 15.9 (4.6) BMI ≥30: 20.3 (18.3 – 27.8)
Emphysema severity score	Trace: 9 (18 %), Mild: 15 (31 %), Moderate: 9 (18 %), Severe: 16 (33 %)
CT-based lung volume (L)	6.8 (1.6)
LAV% (on standard dose CT)	Overall: 15.4 (9.2 – 28.7) BMI ≤25: 31.5 (9.8) BMI 25–30: 18.0 (11.3) BMI ≥30: 11.6 (6.6)

FEV₁ forced expiratory volume in 1 second; FVC forced vital capacity;
CTDI_{vol} volumetric CT dose index; DLP dose length product.

[†] $p < 0.05$ (Shapiro-Wilk).

6.3 Results

The characteristics of the study cohort are shown in Table 2. The median DLP of the ULDCT was 16.6 (range 7.3 – 47.6 mGy·cm), on average 84 % lower than the radiation dose of the SDCT (range, 53 – 93 %). Visual evaluation showed at least moderate severity of emphysema in 51 % of patients. No errors in segmentation of lung volume were visually apparent. For SDCT, the mean total lung volume was 6.8 L (standard deviation 1.6 L, range 3.9 – 10.9). The mean absolute difference in volume between SDCT and ULD-FBP was 270 mL (standard deviation 300 mL, range 2 – 1285 mL), without systematic bias (Figure 1). The relative difference in volume between the scans was 0.9 ± 6.4 %. The mean \pm SD LAV% per BMI group was 31.5 ± 9.8 (BMI ≤ 25), 18.0 ± 11.6 (BMI 25-30), and 12.0 ± 6.6 (BMI ≥ 30).

The noise level was 20.9 HU for SDCT, and 33.9 HU for ULD-FBP (Table 3). For ULDCT, at increasing ADMIRE and DLNR level, noise decreased. The noise level for ADMIRE 3 and DLNR 3 was closest to SDCT reconstructed with FBP (24.6 HU and 22.5 HU, respectively). Figure 2 illustrates the effect of increasing levels of ADMIRE and DLNR on visual appearance in a typical emphysema case. A chart in which the image noise is plotted against the BMI is included in the supplement as Figure S6.2 [p. 221]. Differences in measured lung volume for the denoised ULD reconstructions versus ULD-FBP were minimal, with a maximal difference of 68 mL (0.66 % of the lung volume) for one outlier (Figure 3).

Median extent of emphysema was 18.6 % for ULD-FBP and 15.4 % for SDCT (inter-quartile range: 11.8 – 28.4 % and 9.2 – 28.7 %, $p = 0.0026$).

Table S6.1 [p. 220] contains the full description for each separate reconstruction. In Figure 4, the difference in LAV% (Δ LAV%, ULD-FBP–SDCT) is plotted against the lung volume ratio (the lung volume on ULDCT as a percentage of SDCT). The difference in LAV% between SDCT and ULDCT ranged from -9.6 % to 10.7 %. Only a weak trend is visible (R^2 for linear trend line 0.36), which disappears when scan pairs with a larger difference in lung volume are omitted from the analysis. Other ULDCT reconstructions showed a similar weak trend.

Compared to SDCT, the systematic bias in emphysema extent based on ULDCT was minimal for ADMIRE 1 and DLNR 1 (0.7 LAV%-point and 0.1 LAV%-point, respectively), and increased for higher levels of noise reduction (up to -4.8 LAV%-point for ADMIRE 5 and -8.8 LAV%-point for DLNR 9), with more underestimation of LAV% (Figure 5). Low levels of denoising had high variability as assessed by distance between limits of agreement (Δ LoA 13.1 %-point for ADMIRE 1, 12.9 %-point for DLNR 1); this decreased at intermediate ADMIRE and DLNR settings (11.0 %-point for ADMIRE 3, 10.1 %-point for DLNR 5). In contrast, for the highest levels of denoising, the variability increased (11.2 %-point for ADMIRE 5, 13.9 %-point for DLNR 9). The optimal settings in terms of Δ LoA were ADMIRE 3 and DLNR 3, representing a reduction in variability of

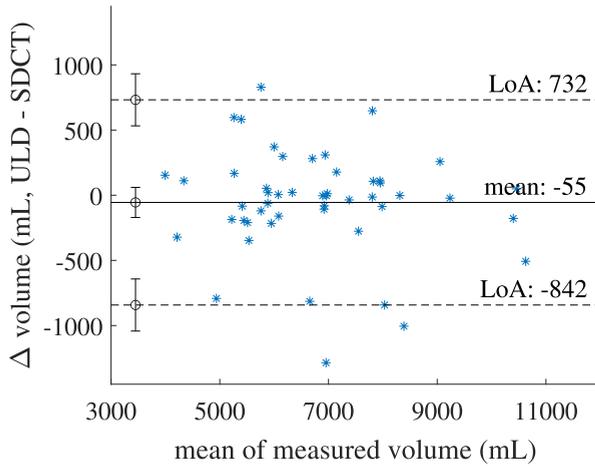


Figure 1: Bland-Altman plot of the lung volume for the ULDC and SDCT scans.

The continuous line denotes the mean value and the dotted lines mark the upper and lower limits of agreement.

ULDCT: ultra-low dose CT; SDCT: standard dose CT; LoA: limit of agreement.

Table 3: Image noise level by reconstruction.

Scan and reconstruction type	Strength/level of denoising	Noise (HU)
SDCT	N.A.	20.9 (17.7 – 24.1)
ULDCT FBP	N.A.	33.9 (30.2 – 36.6)
ULDCT ADMIRE	1	30.9 (27.6 – 33.3)
	3	24.6 (22.0 – 26.9)
	5	17.4 (15.9 – 18.9)
ULDCT DLNR (PixelShine)	1	29.9 (26.2 – 32.5)
	3	22.5 (19.2 – 24.6)
	5	16.8 (14.2 – 18.9)
	9	7.0 (5.8 – 8.1)

SDCT: standard dose CT; ULDCT: ultra-low-dose CT; FBP: filtered backprojection; ADMIRE: advanced modelled iterative reconstruction; DLNR: deep learning-based noise reduction.

24 % and 27 %, respectively, at a systematic bias of -1.5 and -2.9 . Stratifying by normal/low (≤ 25), medium (25–30) and high (≥ 30) BMI did not reveal an additional trend (Figure S6.3 [p. 222]).

In Figure 6, box plots show the difference between LAV% derived from ULDCT reconstructions and SDCT as reference standard, as well as the range in LAV%. All median differences for denoised reconstructions were significantly different from ULD-FBP ($p < 0.0001$). The Δ LoA value was significantly different for DLNR 3 ($p = 0.0498$).

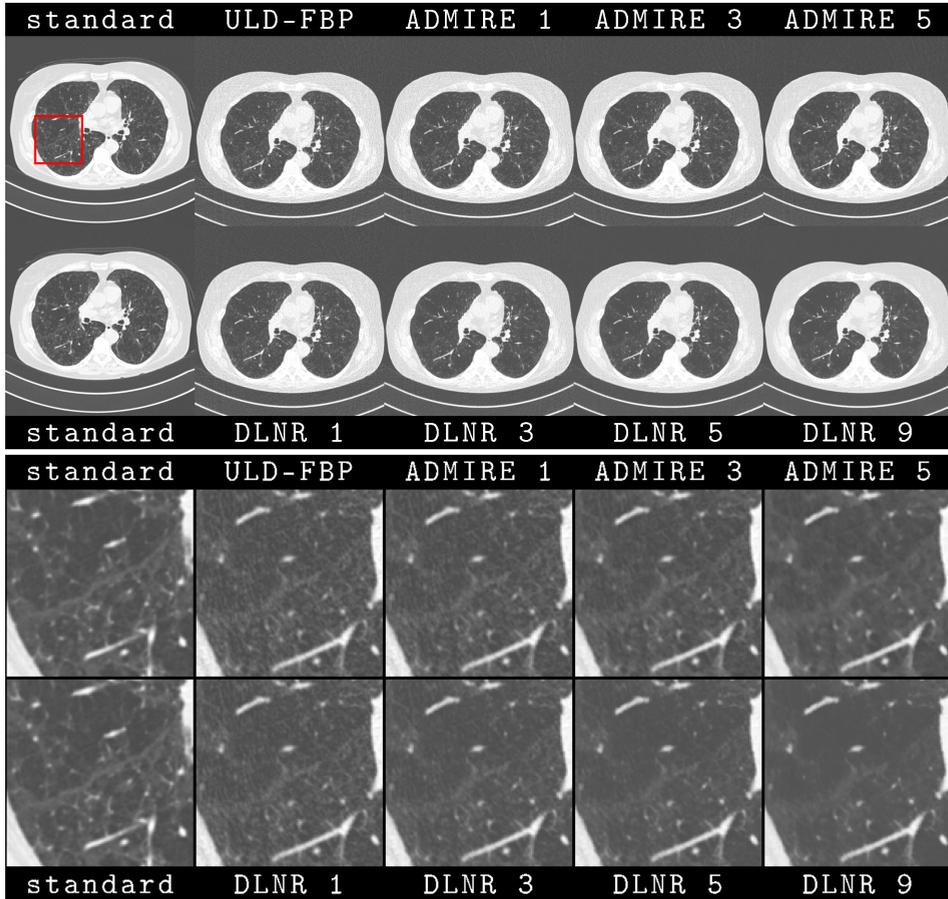


Figure 2: Axial CT slices from a typical case with visible emphysema.

Top row (from left to right): clinical baseline CT (SD-FBP), ULD-FBP and ADMIRE level 1, 3, and 5.

Bottom row: SD-FBP, and DLNR 1, 3, 5, and 9 for ULDCT. The window level is WW1600/WL-700.

Part A contains the full-size slices, part B contains a cropped area indicated by the red box.

SD-FBP: standard dose CT filtered backprojection; ULD-FBP: ultra-low-dose CT filtered backprojection; ADMIRE: advanced modelled iterative reconstruction; DLNR: deep learning-based noise reduction; ULDCT: ultra-low-dose CT.

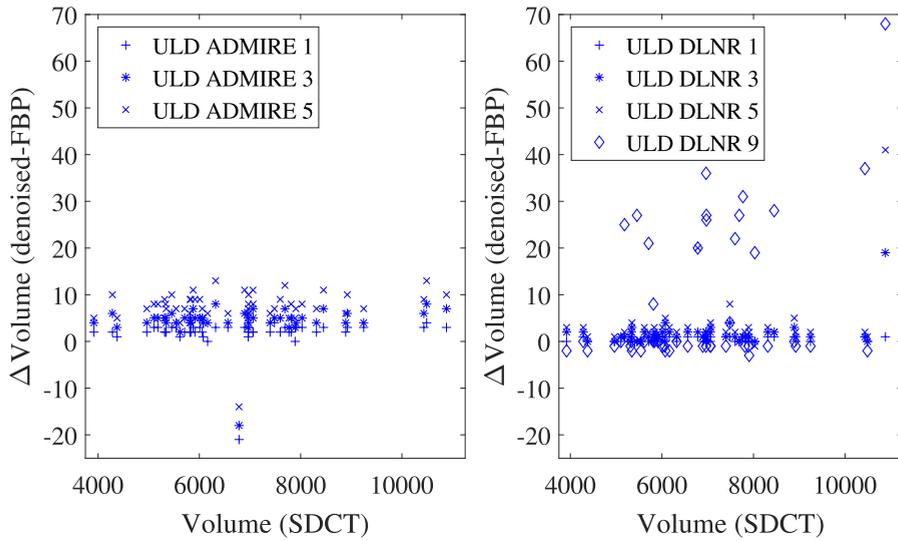


Figure 3: Lung volume measurement difference for ULDCT reconstructions (ADMIRE and DLNR volume minus FBP volume) compared to the lung volume based on SDCT.

Part A contains the ADMIRE results, part B contains the data for the DLNR. ULDCT: ultra-low dose CT; FBP: filtered backprojection; ADMIRE: advanced modelled iterative reconstruction; DLNR: deep learning-based noise reduction; SDCT: standard dose CT.

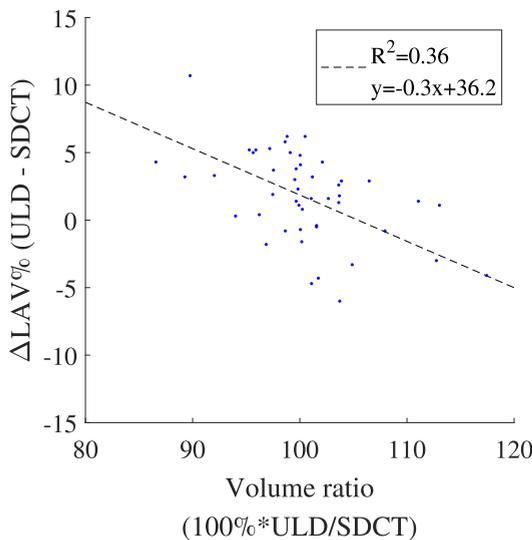


Figure 4: Ratio of the measured volume plotted against the $\Delta\text{LAV}\%$ ($\text{LAV}\%$ ULD-FBP minus $\text{LAV}\%$ SDCT). The dotted line is a linear trendline.

SDCT: standard dose CT; ULD-FBP: ultra-low dose CT filtered back-projection; LAV: low attenuation value.

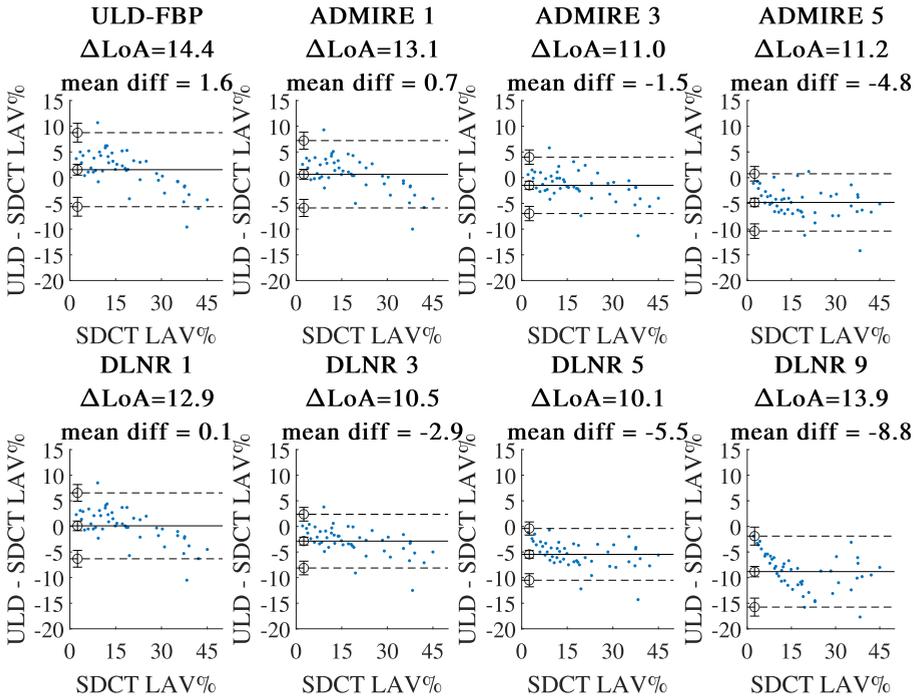
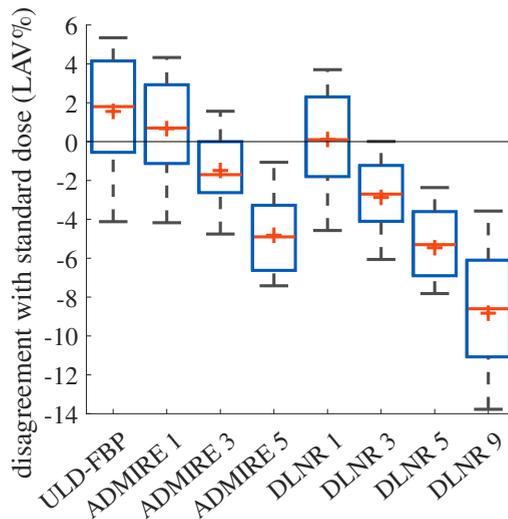


Figure 5: Residual plots showing the results of the Bland-Altman analysis for LAV%, including the confidence intervals for the mean and limits of agreement. Each subplot compares a different ULDCCT reconstruction to SDCT.

SDCT: standard dose CT; ULDCCT: ultra-low dose CT; FBP: filtered backprojection; ADMIRE: advanced modelled iterative reconstruction; DLNR: deep learning-based noise reduction; LAV: low attenuation value; ΔLoA : distance between limits of agreement.

Figure 6: Boxplots showing the difference between the LAV% derived from ULDCCT (FBP, ADMIRE and DLNR) and LAV% based on SDCT as reference standard.

SDCT: standard dose CT; ULDCCT: ultra-low dose CT; FBP: filtered backprojection; ADMIRE: advanced modelled iterative reconstruction; DLNR: deep learning-based noise reduction; LAV: low attenuation value.



6.4 Discussion

Our study shows that ULDCT reduced radiation dose by 84% compared to standard-dose CT in COPD patients. State-of-the-art noise reduction techniques significantly reduced variability in emphysema quantification compared to FBP. While low levels of ADMIRE/DLNR in ULDCT had low systematic bias, they had relatively wide limits of agreement. Higher levels of noise reduction techniques reduced variability, at the cost of underestimation of emphysema. ADMIRE 3 and DLNR 3 provided an optimal balance for emphysema quantification in ULDCT, with a decrease in variability by up to 27% compared to FBP, at a slight underestimation of the extent of emphysema.

Both the segmentation of the lung and the inspiration level during the scan can affect the emphysema extent. Because the LAV% is the percentage of low attenuation voxels, measuring different voxels can lead to different outcomes. As lung segmentation is computationally easy, the segmentation itself is not expected to differ much between different reconstructions of the same scan. A difference in inspiration level for the SDCT and ULDCT scans could introduce a difference in LAV by affecting the density of the tissue itself, but our results show that in nearly all patients the inspiration level for both scans was similar (mean difference 270 mL). The volume differences between denoised reconstructions for ULDCT, and ULD-FBP were not dependent on the actual lung volume, and were minimal (−21 mL to +68 mL). This suggests that the segmentations are sufficiently similar to not expect any LAV difference caused by the segmentation alone.

Prior literature

In a study by Iyer et al., participants were coached during spirometry-guided CT scans in one scan session with two standard dose acquisitions on the same CT system [190]. They found a ΔLoA of 1.77 %-point, which represents the inherent variability in a best-case scenario. Under more usual clinical circumstances, the variation in LAV has been studied in lung cancer screening trials [191–194]. Compared to our study cohort, there tended to be only a limited amount of emphysema in these studies. By design, the paired scans in prior studies were made with the exact same CT protocol, dose level, and CT system. The present study did not use the same protocol twice. Thus, the results from prior studies are not fully applicable.

There are two factors in this study that may have a major influence on the emphysema quantification. The first aspect is the effect of denoising on emphysema quantification. Two studies with a study design close to ours are by Messerli et al. and Den Harder et al. [157, 195]. Both looked at the effect of iterative reconstruction on emphysema quantification in ULDCT with a clinical protocol as reference standard. The general trend in their results, and that of extensive prior research on iter-

ative reconstruction, was similar to ours: higher levels of iterative reconstruction reduce image noise, and lower the measured LAV [157, 178, 184, 185, 195, 196]. The change in LAV may be related to the effect of IR on the HU value of tissues with a density close to air [9]. This same trend is visible in the results with DLNR, although the research on this topic has hitherto been limited [164, 174]. It should be noted that it is not a given that all DLNR systems will have the same effect on emphysema quantification. If suppression of high spatial frequencies is the principal consequence of both, that would explain the similar effect on LAV underestimation. The second aspect is the effect of dose reduction itself. Dose reduction seems to have the opposite effect of IR on LAV quantification, resulting in LAV overestimation [173, 178, 195]. This indicates that the right combination of scan and reconstruction parameters is required to minimise the differences in emphysema quantification.

Variability in LAV% measurement is also affected by these two factors. Messerli et al. did not report the Δ LoA, but reported a confidence interval for the LAV% difference instead. The narrowest confidence interval was 9.7 %-point for ADMIRE 4 [157]. A more recent study analysed the interscan variability in LAV% for different radiation dose levels and reconstruction kernels of ULDCT scans in 49 patients without confirmed COPD, using 120 kV at low mA setting [178]. The Δ LoA in that study was 14.7 %-point, compared to 14.4 %-point for the equivalent comparison in our study (i.e. ‘uFBP-Stnd vs sFBP-Stnd’). Thus, variability was comparable between our studies. This suggests that the differences between ULDCT and SDCT are not caused by the differences in CT system (e.g. tin filtration) but are primarily related to the radiation dose itself.

Strengths

A strength of this study is the well-described cohort of COPD patients with a distribution of emphysema, unlike other studies [178, 191, 194]. Another strength is that the reference scan was specifically acquired for parenchymal analysis, and therefore did not involve intravenous contrast. Furthermore, SDCT and ULDCT scans were performed on the same day with a standardised protocol, ruling out disease progression. This also mimics the clinical situation where frequent disease monitoring or screening would be performed with ULDCT, while diagnostic scans would often be made with a standard-dose protocol on a routine CT system.

One particular advantage of the use of DLNR is that this software can be applied to CT scans from any CT system or vendor, even long after scan acquisition. This adds to the generalisability of the results of this study, although the specific ULDCT protocol in our study is so far only available from one CT vendor.

Limitations

The reference standard to determine the severity of emphysema is pathology. Many studies use LAV as a proxy measure, since LAV correlates well with pulmonary function and pathology results [43, 158, 159, 176, 177, 197]. This correlation is not perfect [39, 43, 176], but is at this moment the best non-invasive measure available.

Another possible limitation of this study was the presence of potential differences due to a second acquisition, i.e. different inspiration levels between SDCT and ULDCCT and the inherent differences of the CT systems (e.g. different reconstruction kernels).

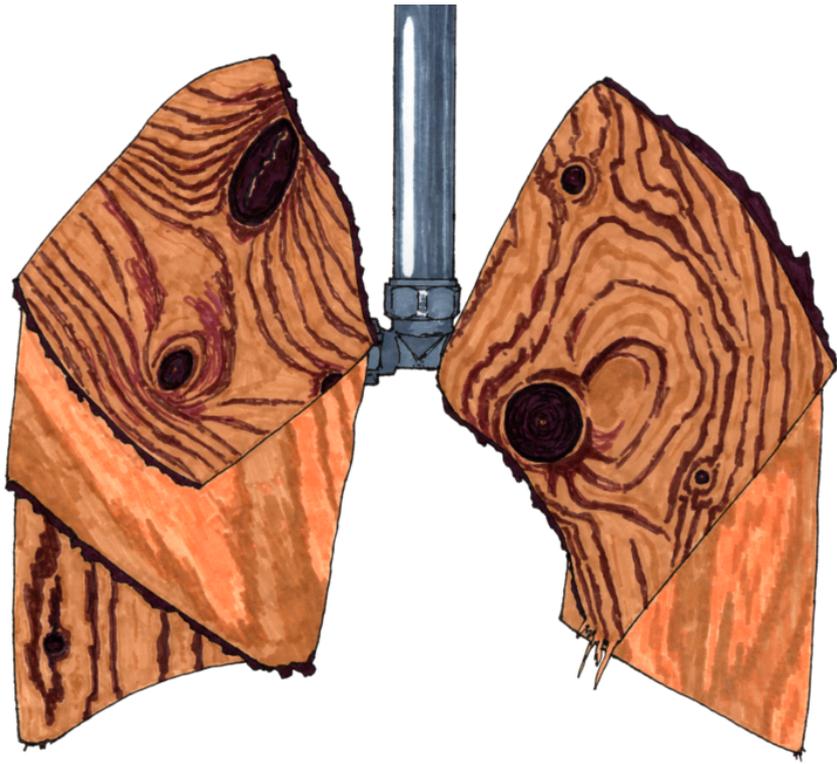
Future research

The results of this study suggest that ULDCCT at 84 % reduced radiation dose is able to yield emphysema measurements close to SDCT, although agreement was not perfect. Subsequent investigations should determine if a simple baseline correction is sufficient to correct for the systematic bias and reliably determine the level of parenchymal destruction. Alternatively, as suggested by Den Harder et al., the threshold value could be changed depending on the reconstruction parameters [195].

Future research is needed to assess if the scans are sufficiently accurate and detailed, so that no relevant structural information required for visual assessment is lost. This is of additional relevance when studying bronchial wall thickness, which is an important parameter in the bronchopathy phenotype of COPD. Whether the results of this study are generalisable to CT systems from different vendors remains to be seen, especially in the case of DLNR. Future research is additionally required to confirm these results in a larger cohort, before clinical implementation can be proposed. In the context of such a study, it would also be interesting to see whether the correlation between LAV% and pulmonary function test parameters is similar for ULDCCT.

6.5 Conclusions

Ultra-low-dose CT in COPD patients allows dose reduction by 84 %. State-of-the-art noise reduction methods in ULDCCT resulted in a slight underestimation of emphysema compared to SDCT. Noise reduction methods (especially ADMIRE 3 and DLNR 3) reduced variability of emphysema quantification in ULDCCT by up to 27 % compared to FBP.



Chapter 7

Improved precision of noise estimation in CT with a volume-based approach

Published in European Radiology Experimental.

DOI: 10.1186/s41747-021-00237-x

Hendrik Joost Wisselink, Gert Jan Pelgrim,

Mieneke Rook, Ivan Dudurych,

Maarten van den Berge, Geertruida H. de Bock,

Rozemarijn Vliegenthart

Abstract

Assessment of image noise is a relevant issue in computed tomography (CT). Noise is routinely measured by the standard deviation of density values (Hounsfield units, HU) within a circular region of interest (ROI). We explored the effect of a spherical volume of interest (VOI) on noise measurements. Forty-nine chronic obstructive pulmonary disease patients underwent CT with clinical protocol (regular dose [RD], volumetric CT dose index [CTDI_{vol}] 3.04 mGy, 64-slice unit), and ultra-low dose (ULD) protocol (median CTDI_{vol} 0.38 mGy, dual-source unit). Noise was measured in 27 1 cm² ROIs and 27 0.75 cm³ VOIs inside the trachea. Median true noise was 21 HU (range 17-29) for RD-CT and 33 HU (26-39) for ULD-CT. The VOI approach resulted in a lower mean distance between limits of agreement compared to ROI: 5.9 versus 10.0 HU for RD-CT (-40 %); 4.7 versus 9.9 HU for ULD-CT (-53 %). Mean systematic bias barely changed: -1.6 versus -0.9 HU for RD-CT; 0.0 to 0.4 HU for ULD-CT. The average measurement time was 6.8 s (ROI) versus 9.7 (VOI), independent of dose level. For chest CT, measuring noise with a VOI-based instead of a ROI-based approach reduces variability by 40 – 53 %, without a relevant effect on systematic bias and measurement time.

7.1 Background

In computed tomography (CT) imaging, the call for dose reduction has led to ongoing efforts to mitigate the effects of increased noise. Current strategies include iterative reconstruction methods and artificial intelligence-based techniques. Less attention is given to the optimisation of noise measurement. The common definition of image noise is the standard deviation (SD) of the measured Hounsfield units (HU) in a physically homogeneous volume [49]. The noise level depends on the specific acquisition and reconstruction parameters, total attenuation of the scan subject, absolute density of the tissue of interest, and on the location in the scanner bore (i.e., the distance of a given voxel to the centre of the field of view). For that reason, it is important to measure a calibration structure with a density and location similar to the tissue of interest. By using a standardised location, the noise measurement provides a good indication for inherent image noise, except in cases of local image artifacts like beam hardening [157, 195].

In chest CT, optimal representation of image noise may be obtained by segmenting the entire tracheobronchial tree lumen, and measuring the SD of this air. However, this is not feasible in most clinical software programmes, due to software limitations and/or time constraints. Because of this, the current clinical practice is to measure the SD in a 1 cm² circular region of interest (ROI) inside the trachea [157, 185]. Accurate noise measurements are important for protocol optimisation and quantification processes [9, 198, 199]. For instance, in emphysema quantification by CT lung densitometry, image noise may affect the threshold needed for reliable distinction between emphysema and normal lung tissue [9, 157].

Moreover, reducing variability of HU measurements may have other clinical implications. The ROI-based technique is commonly used for the assessment of liver parenchyma density and for kidney stone density. These measurements, too, are prone to variation, partly inherent to the ROI-based approach and exacerbated by the sensitivity of mean to outliers [198, 199]. This suggests that the results of this study are applicable to more CT scan indications than just lung CT imaging and assessment of noise. Since reproducibility largely depends on the number of voxels included in the calculation, using a volume-based approach with a volume of interest (VOI) may result in greater precision, without requiring more complicated processing (e.g., by measuring multiple ROIs). Despite this, many studies over the years, including recent studies, have used an ROI-based approach [200–204].

The aim of this study was to determine the systematic bias and variability of ROI-based and VOI-based noise measurements in CT scans obtained at two radiation doses, regular dose (RD) and ultra-low dose (ULD), resulting in low and high noise levels, respectively. These two study arms were independently analysed.

Table 1: Patient cohort characteristics (N=49).

Values are given as mean±standard deviation or median (range), unless stated otherwise. To facilitate comparison, the DLP for the regular-dose CT is expressed as median (range), despite a normal distribution ($p = 0.103$).

Characteristic	Value
Age (years)	66±7
Sex	34 males (69 %), 15 females (31 %)
Body mass index (kg/m ²)	28.0±5.3
FEV ₁ (% of predicted)	53±16
FEV ₁ /FVC (%)	42.4±11.2
CTDI _{vol} (mGy), regular CT protocol	3.04
DLP (mGy·cm), regular CT protocol	105.1 (86.3 – 134.1)
CTDI _{vol} (mGy), ultra-low CT protocol	0.38 (0.19 – 1.06)
DLP (mGy·cm), ultra-low CT protocol	16.6 (7.3 – 29.8)

CT: computed tomography; FEV₁: forced expiratory volume in 1 second; FVC: forced vital capacity; CTDI_{vol}: volumetric CT dose index; DLP: dose length product.

7.2 Methods

Patient cohort

In an on-going chronic obstructive pulmonary disease (COPD) patient study, 50 patients underwent non-contrast clinical chest CT at RD as well as ULD CT between February 2018 and June 2018. The two scans were made on the same day and the order was randomised between participants. The institutional ethical board gave approval for this study and participants provided written informed consent (METC 2015/335, clinicaltrials.gov NCT02477397). Table 1 shows a summary of the patient characteristics. One patient was excluded due to a body habitus far outside the normal range for COPD patients: a body mass index of 56, over 5 standard deviations (SDs) above the mean of the remainder of the cohort.

CT scans

The RD-CT scans were acquired on a routine 64-slice CT system (Somatom Definition AS, Siemens Healthineers, Forchheim, Germany) with routine high-resolution CT protocol of 40 mAs (fixed tube current) and 120 kVp (volumetric CT dose index [CTDI_{vol}] 3.04 mGy). The ULD-CT scans were acquired on a third generation dual-source CT system (Somatom Force, Siemens Healthineers, Forchheim, Germany) with 70 mAs (reference tube current), at 100 kVp with Sn filter (median CTDI_{vol} 0.38 mGy, range 0.19 – 1.34 mGy). The pitch was 1.5 for RD-CT and 1.6 for ULD-CT. The field of view was adjusted to the individual patient size for each scan (range 317 – 500 mm). Scans were reconstructed with

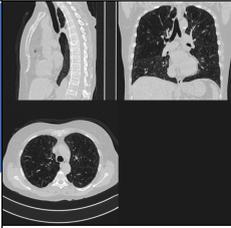
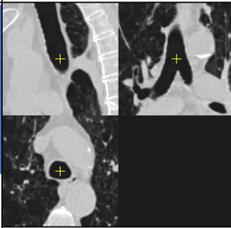
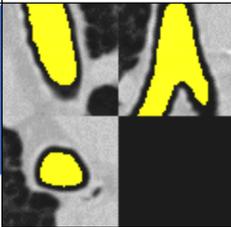
Stage	Region	Task
Manual selection	 Full scan	The approximate location of the trachea is selected.
Isocentre refinement	 151*151*151 voxels	The location of the carina is determined. Then the centre of the trachea 1.0/1.5/2.0 cm above the carina ridge is determined.
Measurement	 61*61*61 voxels	The trachea and main bronchi are segmented and eroded to obtain the ground truth. ROI and VOI measurements use the same isocentre.

Figure 1: Flow chart of the steps to determine the ground truth noise and the isocentre for the measurements. ROI: region of interest; VOI: volume of interest.

slice thickness/increment of 1.0/0.7 mm, filtered back projection and a soft kernel. The two kernels used B30f and Br40, respectively, are suggested by the vendor as similar and are generally treated in literature as comparable [203].

Image analysis

Analysis was performed with an in-house developed MATLAB script (MATLAB R2020b, MathWorks, Natick, Massachusetts, USA). The complete function is available online via <http://tiny.cc/YL3BNUQ4>. The choice for a stand-alone analysis script was made to avoid time-consuming efforts to determine the variability of manual measurements. The simulation method is a best-case scenario for what a human reader would achieve. The noise level was defined as the SD of the selected voxels. To obtain the ground truth for the noise level for intra-thoracic air, a section of the tracheobronchial tree (caudal trachea and proximal bronchi) was segmented in a $61 \times 61 \times 61$ voxel region (referred to as ‘trachea segmentation’ or ‘segmentation’ in the remainder of this paper). Due to the patient-specific field of view, the size in millimetres of this cubic region differed case by case. See the flow chart in Figure 1 for a description of each step in this process. For the ROI and VOI, a standardised measurement location was used (a fixed distance above the carina ridge). The edge of the segmentation was removed with a morpholog-

ical erosion (a mathematical operation removing boundary pixels) to avoid edge artifacts.

To simulate repeated manual measurements, a jitter was applied, meaning the centroid was moved one voxel in x , y , and z -direction, resulting in 27 possible locations. For all 27 centroids, the noise was measured with both a circular ROI and spherical VOI. The radius was based on an area of 1.0 cm^2 , resulting in a VOI of approximately 0.75 cm^3 . Due to these definitions, the number of voxels used for these analyses depended on the FOV and the slice thickness. For the ROI, between 101 and 261 voxels were included (median 177 voxels), for the VOI between 1117 and 2789 voxels (median 1849 voxels). If either the ROI or VOI contained voxels outside the segmentation (prior to the previously mentioned morphological erosion), both ROI and VOI were excluded from further analysis for that measurement position, mimicking manual measurements. The values obtained at the level above the carina ridge that resulted in the fewest rejections were used for the remainder of the analysis (at either 1.0, 1.5, or 2.0 cm), to further mimic a manual measurement accounting for anatomical variation. This height selection was done separately for each scan.

To estimate the extra time required for a VOI-based measurement, a trained researcher (HJW) measured the noise ten times manually with each strategy. The Syngo.Via software (version VB40A, Siemens Healthineers, Forchheim, Germany) was used to perform the measurements. To account for the imprecision of a manual measurement and considering that a precise area or volume may not be possible given the voxel size of a specific scan, a radius difference of up to 5% with the area or volume described below was considered acceptable when measuring the noise. The order of the measurements was randomised.

Statistical analysis

Statistical analysis was performed with MATLAB R2020b (MathWorks, Natick, Massachusetts, USA). Bland-Altman analysis was used to determine the systematic bias between the true noise level and measured noise [205]. The difference between the systematic biases of the two measurement strategies was tested with the Wilcoxon signed-rank sum test. Variability was defined as the distance between the limits of agreement. Because this is directly related to the variance, Levene's test was used. Each characteristic in Table 1 (except sex) was tested separately for normality with the Shapiro-Wilk test.

7.3 Results

The seed point location and the segmentation of the air in the trachea was visually confirmed for each reconstruction. One representative case is depicted in Figure 2, showing successful segmentation without excluding large parts of the trachea or

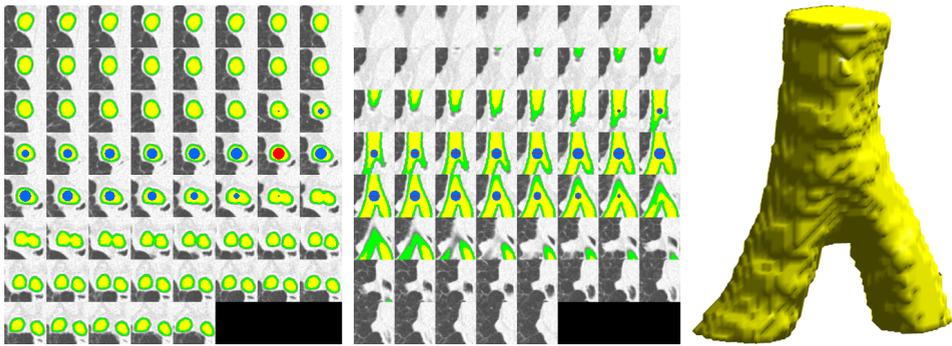


Figure 2: Subsection of the CT images around the carina (window width 1600 HU, window level -700 HU). The red part is the position of the region of interest, the blue is the volume of interest, the yellow is used to measure the ground truth, and the green area was removed from the segmentation to prevent edge artifacts like the partial volume effect.

This image shows the measurement with the isocentre 1.0 cm above the carina ridge.

Left: Axial images. Middle: Coronal images, interpolated to account for the anisotropic dimensions of the voxels. Right: Volume render of the yellow segmentation.

main bronchi, or including parenchyma or bronchial wall. For RD-CT, 66 of 1323 jitter-scan combinations (5.0 %) were discarded because the ROI or VOI contained voxels outside the trachea. For ULD-CT, 84 of 1323 combinations (6.3 %) were discarded. This led to a total exclusion rate of 150 of 2646 values (5.7 %). The range of true noise based on the trachea segmentation was 17 – 29 HU for RD-CT and 26 – 39 HU for ULD-CT. As these ranges are based on the true noise, only a single value per patient was obtained. For the ranges of the noise measured with a ROI or a VOI, all valid measurements were considered. The range of noise measured with a ROI was 11 – 32 HU for RD-CT (based on 1257 measurements) and 23 – 44 HU for ULD-CT (based on 1239 measurements). The respective ranges for the VOI-based measurement were 13 – 30 HU for RD-CT and 25 – 43 HU for ULD-CT.

The results of the Bland-Altman analysis in residual plots are shown in Figure 3. As the noise was measured in 27 different locations, there are multiple dots for each scan. Because every scan has only one ground truth noise value, this results in vertical patterns. For the VOI-based approach, the distance between limits of agreement, compared to the ROI-based approach, decreased from 10.0 to 5.9 for RD-CT (40 % reduction, $p < 0.001$) and from 9.9 to 4.7 for ULD-CT (53 % reduction, $p < 0.001$), indicating a lower inter-measurement variation when using the VOI-based method. There was a minimal effect on the systematic bias for both the RD-CT (-1.6 to -0.9 HU, $p < 0.001$) and ULD-CT (0.0 to 0.4 HU, $p < 0.001$).

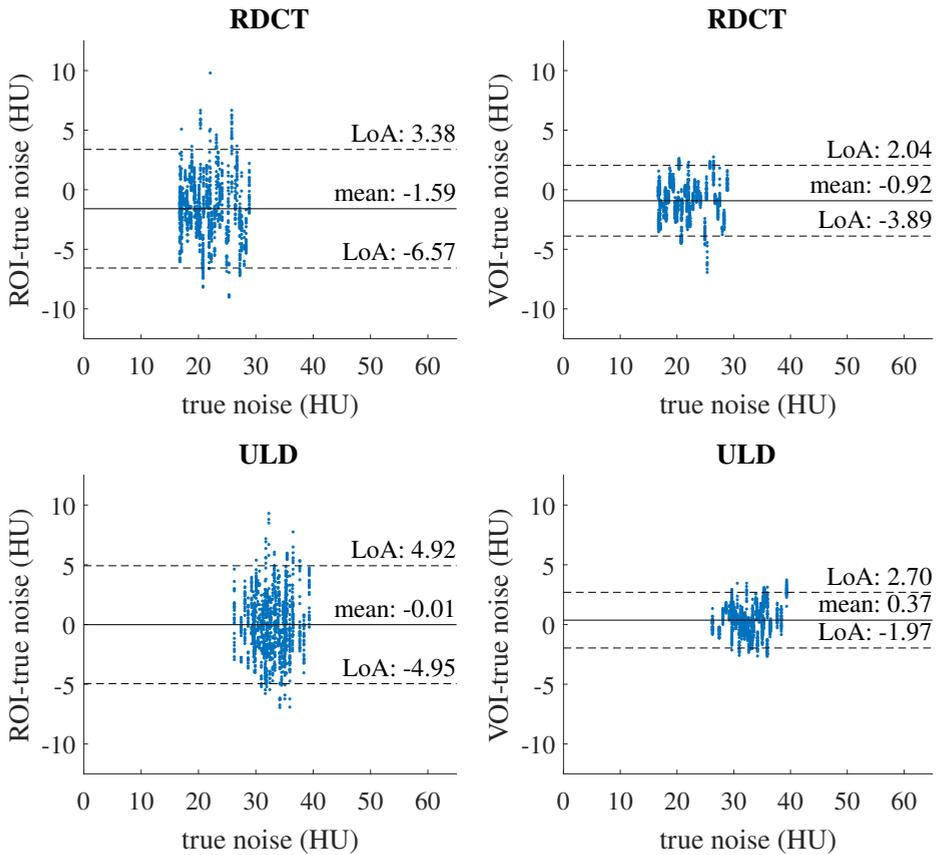


Figure 3: Results of the Bland-Altman analyses. Each plot shows the difference between the noise measured with either ROI or VOI and ground truth noise on the y-axis, versus ground truth on the x-axis. Regular radiation dose computed tomography protocol measured with an ROI (a) or a VOI (b), same data for ultra-low dose protocol (c and d, respectively).

ROI: region of interest; VOI: volume of interest; LoA: limits of agreement; HU: Hounsfield units.

The manual ROI measurement by the trained researcher took 6.8 s on average; for the VOI measurement, this increased by 2.9 s to 9.7 s (+43 %) and would therefore not meaningfully increase the time required to read a CT scan.

7.4 Discussion

In this study, we showed that a VOI-based noise measurement approach significantly improves precision compared to a ROI-based approach, especially in CT scans with a higher intrinsic noise level, without a relevant trade-off in terms of measurement time.

As early as 1978, an alternative method for objective measurement of image noise was published: a noise power spectrum (NPS) [206]. This has the benefit of not relying on the measured region being homogeneous and of providing a more detailed description of noise, instead of relying on a single descriptive value. Despite the NPS method having been available for decades, clinical studies have continued to use the ROI method [200–204] while the NPS method is only used in highly technical applications [180]. To our knowledge, no clinical system provides the option to compute the NPS. Thus, the calculation of the NPS will most likely require exporting the scans for external processing, making it less desirable for either research or clinical use. This same limitation applies to using the segmented trachea to measure the noise.

Other studies proposed other methods to improve on the ROI-based method, e.g., by subtracting two adjacent slices (similar to how digital subtraction angiography works) before calculating either a local (pixel-by-pixel) SD, a regional SD, or multiple regional SDs [207–210]. Such methods are particularly useful in situations where noise does not have a Gaussian distribution, or where pixel value differences exist due to anatomical structures [207, 209]. Another commonly proposed method is to average multiple regions [208, 209]. This is mostly used for liver parenchyma, where multiple smaller ROIs are sometimes used to ensure a measurement area that better reflects the organ as a whole [200]. To our knowledge, none of the previously mentioned alternatives to the ROI-based method are available for routine clinical use.

Given the increased use of artificial intelligence (AI), any specific application of a ROI-based measurement may eventually be replaced by an AI tool. Such tools may forgo measuring a specific density or noise level in favour of directly assessing the intended biomarker [211, 212]. Up to the moment that an AI tool (for this specific application) does become available, the VOI-based method proposed in this work is a simple and quick option, to be preferred over ROI evaluation.

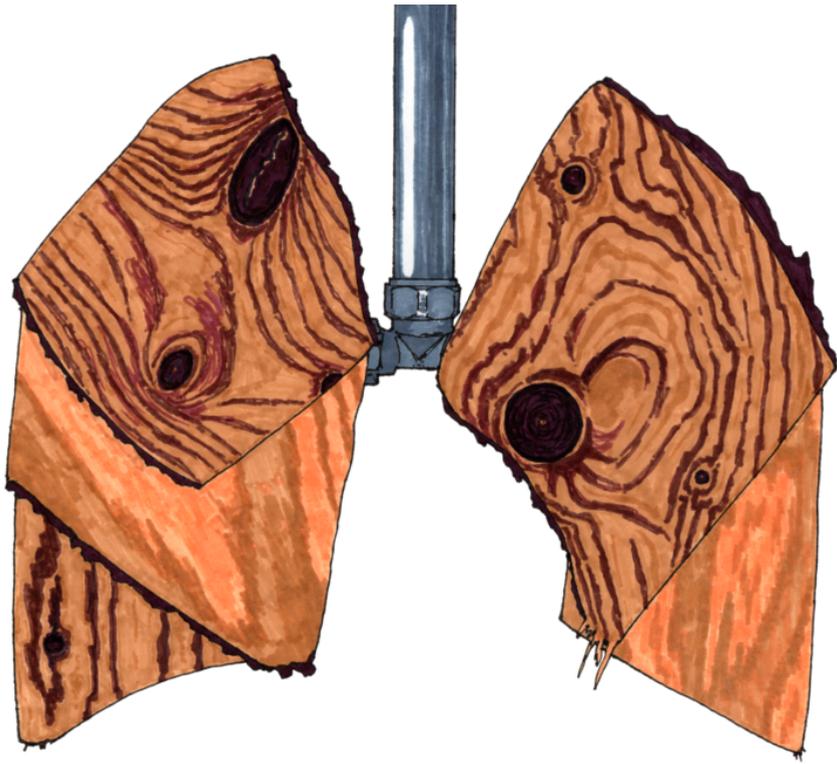
The potentially quick and easy applicability is one of the main advantages of using a volume-based approach, which may help implementation in both research and clinical practice. A VOI-based measurement should be widely available in PACS reading systems, often in the same drop-down menu as the ROI-based measurement option. The extra time required is limited.

More generally, volumetric analyses on CT scans are increasingly common. An example of this is the volumetric assessment of lung nodules, which increasingly replaces the diameter-based approach [213]. Additionally, some nuclear medicine guidelines also require the use of volumetric measurements [214]. To our knowledge, only one previous study has focused on the use of volume-based noise measurements in radiology [215], outside of recent technical quality standards like the QIBA lung density profile [79]. This is unfortunate, as the applicability

is likely not limited to measuring noise, but may also extend to other situations in which a density measurement is performed, e.g., when measuring liver density or muscle density [200, 203]. Future research should be conducted to confirm this expectation.

Some aspects of this study may potentially limit the generalisability of these results. The scans were made on CT systems from one vendor only in a relatively small COPD patient cohort, without including healthy controls. However, only testing scans from a single vendor is not expected to influence the conclusion. To improve generalisability of the results, scans were acquired with many differences in the scan protocol, like radiation spectrum, mAs, and reconstruction kernel. Importantly, the aim of our study was not to compare noise between an RD and an ULD CT scan protocol, but to investigate the method to quantify the noise. This means the scans should not be analysed as pairs, but should be treated as two study arms that are independently analysed. The results from both scanners support the same conclusion, even with the different scan protocols. The small size of the cohort is unlikely to affect the conclusion, even if a larger cohort size would further increase confidence in quantifying the difference between the two methods. Similarly, there is no technical reason why the presence or absence of COPD would influence the noise characteristics in the trachea of an ROI compared to a VOI. Lastly, switching from an automated script to a human reader is unlikely to substantially change the results.

In conclusion, in chest CT protocols, measuring image noise with a VOI-based approach instead of a ROI-based approach reduces variability by 40 – 53 %, without a relevant effect on systematic bias and measurement time.



Chapter 8

CT-based emphysema characterisation per lobe: a proof of concept

Published in the European Journal of Radiology.

DOI: 10.1016/j.ejrad.2023.110709

Hendrik Joost Wisselink, Xiaofei Yang,

Mieneke Rook, Marjolein A. Heuvelmans,

Wenzhen Jiang, Jianing Zhang, Yihui Du,

Marleen Vonder, Monique D. Dorrius, Zhaoxiang Ye,

Geertruida H. de Bock, Rozemarijn Vliegenthart

Abstract

Purpose The Fleischner society criteria are global criteria to visually evaluate and classify pulmonary emphysema on CT. It may group heterogeneous disease severity within the same category, potentially obscuring clinically relevant differences in emphysema severity. This proof-of-concept study proposes to split emphysema into more categories and to assess each lobe separately, and applies this to two general population-based cohort samples to assess what information such an extension adds.

Method From a consecutive sample in two general population-based cohorts with low-dose chest CT, 117 participants with more than a trace of emphysema were included. Two independent readers performed an extended per-lobe classification and assessed overall severity semi-quantitatively. An emphysema sum score was determined by adding the severity score of all lobes. Inter-reader agreement was quantified with Krippendorff's alpha.

Results Based on Fleischner society criteria, 69 cases had mild to severe centrilobular emphysema, and 90 cases had mild or moderate paraseptal emphysema (42 had both types of emphysema). The emphysema sum score was significantly different between mild (10.7 ± 4.3 , range 2–22), moderate (20.1 ± 3.1 , range: 15–24), and severe emphysema (23.6 ± 3.4 , range: 17–28, $p < 0.001$), but ranges showed significant overlap. Inter-reader agreement for the extended classification and sum score was substantial (alpha 0.79 and 0.85, respectively). Distribution was homogeneous across lobes in never-smokers, yet heterogeneous in current smokers, with upper-lobe predominance.

Conclusions The proposed emphysema evaluation method adds information to the original Fleischner society classification. Individuals in the same Fleischner category have diverse emphysema sum scores, and lobar emphysema distribution differs between smoking groups.

8.1 Introduction

Chronic obstructive pulmonary disease (COPD) is a disease with high global prevalence and substantial disease burden, consisting of chronic bronchitis and emphysema [13, 41]. While the gold standard to diagnose COPD is a pulmonary function test, CT scans are increasingly used for both visual and quantitative analysis of the airways and lung parenchyma [39, 41, 42]. Like COPD in general, emphysema as assessed on CT-scans is heterogeneous. Emphysema is subdivided into three subtypes: centrilobular emphysema (CLE), paraseptal emphysema (PSE), and panlobular emphysema (PLE), with CLE and PSE being the most prevalent subtypes, often co-existing [40]. The three subtypes may have a different aetiology and may have distinct symptoms and outcome [38, 46, 216, 217]. The three subtypes can be diagnosed and assessed by CT. Both visually and quantitatively assessed emphysema are associated with an increased risk of mortality and lung cancer [4, 43, 158]. The increased mortality associated with visually assessed emphysema persists even after adjusting for the results of quantitative CT-based analyses [43].

The current standard scoring method to assess emphysema on CT is described in a statement from the Fleischner society [40]. This allows capturing expert opinion semi-quantitatively with five severity levels of CLE, two severity levels of PSE, and a dichotomous classification for PLE (present/absent). Previous emphysema scoring methods generally used a five-point scale that ignored subtype [218–221]. The current criteria yield a general overview of the emphysema presence and severity of lungs as a whole, which can be linked to clinically relevant outcomes [4]. Nevertheless, the presence and extent of emphysema can vary within and between lung lobes. Because the current criteria do not allow consideration of inter-lobe differences, they may obscure differences in overall severity. Extending the criteria allows a more detailed capture of the expert opinion, which may allow more detailed analysis of the development of emphysema and effect of risk factors, as well as the relation with outcomes. This has the potential to provide more information and reduce the impact of a disagreement in classification. The clinical impact of a difference between two scores will potentially be reduced by increasing the number of possible scores and score range, and therefore the level of detail of a classification. This is beneficial, since there is an inherent inter-reader variability in subjective scores.

The aim of this study was to study our proposed extended criteria for a per-lobe characterisation of emphysema with regard to its reproducibility, and to provide a proof of concept by applying it to two population-based samples. The proposed method extends the Fleischner society criteria in two regards: it splits PSE and PLE into more severity levels, and the criteria are applied to each lobe separately.

8.2 Materials and methods

Population

For this study, cases were retrieved from the Dutch ImaLife and Chinese NELCIN B3 general population-based cohorts [10, 47]. The ImaLife cohort comprises participants from the Lifelines population study aged 45 years and older who completed a pulmonary function test (PFT) as part of the Lifelines second round visit and did not receive a chest CT scan in the past year [10]. Lifelines is a study with 167 729 participants recruited from the northern part of the Netherlands. It is a prospective multi-disciplinary three-generation cohort study examining health and health-related behaviours. In this biobank, biomedical, socio-demographic, behavioural, physical and psychological information is gathered, allowing analysis of health and disease, with an additional focus on complex genetics and multi-morbidity [59, 141]. The NELCIN B3 cohort comprises participants aged 40-74 years from the Hexi district of Tianjin who did not report a history of malignancy [47]. All participants in these two cohorts underwent a low-dose chest CT scan. Both cohorts completed a questionnaire, from which the population characteristics (sex, age at time of scan, smoking status, and pack-years) were extracted. In the ImaLife cohort, former smokers were defined as participants having smoked for at least a year and quit smoking at least a month before the questionnaire [222]. In the NELCIN B3 cohort, former smokers were defined as participants who reported having smoked at least 1 cigarette a day for 6 months, and who reported having quit at the time of the interview [64]. Pack-years were based on self-reported daily tobacco use and duration of smoking. The PFTs performed for the ImaLife study allowed determination of the Global initiative for chronic Obstructive Lung Disease (GOLD) stage [41].

Acquisition

CT scans of the ImaLife cohort were acquired using a third-generation dual-source CT system (SOMATOM Force, Siemens Healthineers). Scans were acquired with a reference tube current time product of 20 mAs at 120 kVp, and were reconstructed with filtered backprojection with slice thickness/increment of 1.0/0.7 mm, and a soft tissue kernel (Br40). The pitch was either 2.5 or 3, depending on the required field of view of 400 mm or 350 mm. Standard breath coaching was used to ensure acquisition at full inspiration.

CT scans of the NELCIN B3 cohort were acquired using a 64-multidetector CT system (SOMATOM Definition AS, Siemens Healthineers). The scans were acquired with a reference tube current time product of 35 mAs at 120 kVp, and were reconstructed with filtered backprojection with a soft tissue kernel (B30f) with slice thickness/increment of 2.0/1.0 mm. The pitch was set to 1, with a reconstruction field of view of 400 mm. Standard breath coaching was used to ensure acquisition at full inspiration.

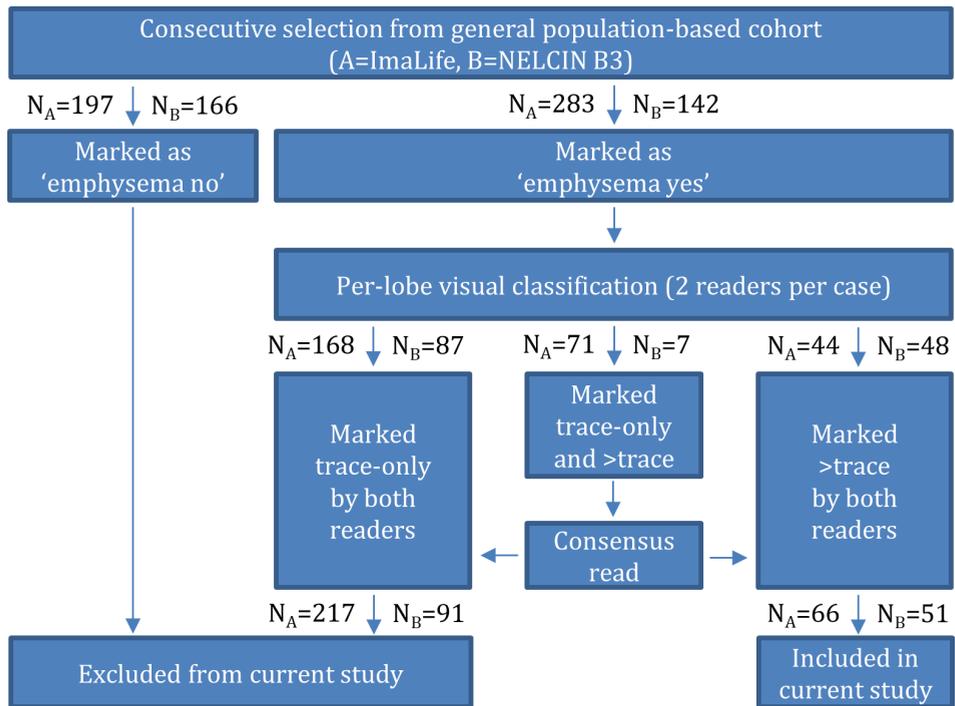


Figure 1: Inclusion flow chart.

Sample selection and scan reading

Sample selection for the current analyses was performed stepwise and separately for the two cohorts (Figure 1). During the initial evaluation of scans, 480+308 (ImaLife + NELCIN B3, respectively) consecutive chest CT scans were marked by a board-certified radiologist (MR for ImaLife; MDD for NELCIN B3) as '(any) emphysema yes or no'. For the current study, the 283+142 cases with any emphysema were further analysed. In this second step, two trained readers independently performed a per-lobe classification for each case with emphysema. Cases with a final conclusion of 'more than trace' (at least in one lobe) were included in this study. Cases with disagreement in the final conclusion were reviewed in a joint session consensus read. In the consensus read, the per-lobe classification was determined for any lobes with disagreement by the same readers who performed the independent reading. The exclusion of trace-only cases was intended to leave only cases with non-trivial emphysema, as trace emphysema of one subtype is unlikely to have a clinical impact in the absence of a more substantial degree of disease of another subtype. This stepwise selection resulted in the inclusion of 66 ImaLife cases and 51 NELCIN B3 cases. For the per-lobe classification the readers were the same in both cohorts (HJW, a trained technical physician, 3 years of experience; XY, a radiologist, 5 years of experience). The readers were blinded to smoking status and pack-years. A random sample of 40 cases was selected from the 117 included cases. These scans were re-assessed

after a gap of several months by the first reader (HJW) to allow determination of the intra-observer variability.

Visual assessment

A detailed visual assessment of CT-based emphysema was performed with our proposed extended Fleischner society criteria as described in Table 1. These criteria were adapted from the original Fleischner society criteria by extending the PSE and PLE levels to homogenise the scale of score levels per subtype, and evaluating emphysema per lobe [40]. The CLE level definitions were not changed. Because it is difficult to visually estimate the cut-off percentages used for CLE classification, computer-generated example distributions were used as a visual reference by the readers. These images are included in the supplemental materials as Figure S8.1 [p. 227]. The sum of the grades of all subtypes is a semi-quantitative indicator of the overall severity of emphysema. To compute an emphysema sum score, trace was scored as 1, mild as 2, etc. See Table 1 for the full conversion from category to score. Such a sum score facilitates interpretation of the overall severity, avoiding the need to consider 15 scores simultaneously. The theoretical range for this sum score is 0 (no emphysema in any of the lobes) to 65 (a hypothetical case with a combination of advanced destructive CLE (5), substantial PSE (4), as well as severe PLE (4) in all five lobes). In practice the maximum score will be around 30. The reason for this is that the subtypes are competing pathologies, and at the more severe end of the spectrum differentiation is no longer possible [38]. Two example images are shown in Figure 2. Additional example images are available in the supplementary materials, see Figure S8.2 [p. 228]. Visual analysis was performed using visualisation software (Syngo.via, version VB40A-HF02, Siemens Healthineers). The initial settings were W1600L–700 for multi-planar reconstructions and W800L–900 for 10 mm slab minimal intensity projection. Readers were given broad discretion to change window level, and to use axial, coronal and sagittal planes.

The original Fleischner society criteria score was inferred from the recorded extended score as follows: to determine a single severity for a participant, the most severe score of any lobe and subtype was taken. Trace PSE was mapped to no PSE, and moderate PSE was mapped to mild PSE. There were no participants with PLE. Calculation examples of this mapping are included in the supplementary materials [p. 224].

Except when stated otherwise, the results from the first reader were used, unless a consensus read was available from the inclusion stage. A lobe was considered affected by emphysema if there was more than trace emphysema of any subtype.

Table 1: Extended Fleischner criteria. The definitions in the original Fleischner criteria presented by Lynch et al. [40] were not changed in the case of centrilobular emphysema, nor was the definition for substantial paraseptal emphysema. These were marked in the table with an asterisk.

Grade Subtype	1	2	3	4	5
Centrilobular emphysema	Trace*: minimal centrilobular lucencies, occupying <0.5 % of a lung zone.	Mild*: scattered centrilobular lucencies, usually separated by large regions of normal lung, involving an estimated <0.5 % of a lung zone.	Moderate*: many well-defined centrilobular lucencies, occupying more than 5 % of any lung zone.	Confluent*: coalescent centrilobular or lobular lucencies, including multiple regions of lucencies that span several secondary pulmonary lobules (no hyperexpansion or distortion of pulmonary architecture).	ADE*: advanced destructive emphysema consists of panlobular lucencies with hyperexpansion and distortion of pulmonary architecture.
Panlobular emphysema	Trace: N.A.	Mild: generalised destruction with hyperexpansion involving a segment.	Moderate: generalised destruction with hyperexpansion involving more than a segment.	Severe: generalised destruction of an entire lobe.	
Paraseptal emphysema (juxtaleural includes adjacent to an interlobar fissure)	Trace: <5 small (< 1 cm) lucencies.	Mild: >5 small (< 1 cm), well-demarcated rounded juxtaleural lucencies, usually aligned in a row, or 1 large (> 1 cm) juxtaleural lucency.	Moderate: multiple groups of juxtaleural lucencies, any lucencies centrally attached to juxtaleural lucencies, or more than 1 large (> 1 cm) non-apical juxtaleural lucency.	Substantial*: mainly large (> 1 cm) juxtaleural cyst-like lucencies or bullae, involving more than the lung apices, aligned in a row along a pleural margin, and sometimes including adjacent to an interlobar fissure.	

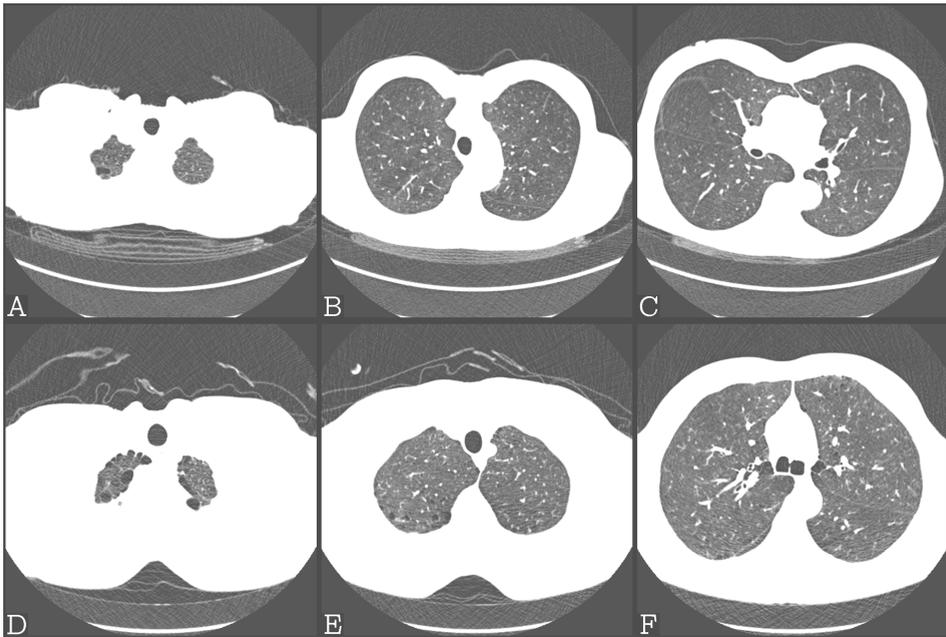


Figure 2: Examples of cases with the same nominal (Fleischner) emphysema severity. The upper section (A-C) shows a case with mild paraseptal emphysema. Since no other emphysema was found in this case, the sum score was 2. The lower section (D-F) shows a case with mild centrilobular and moderate paraseptal emphysema. The sum score for this case was 19. According to the original Fleischner classification both would be marked as mild emphysema. The images shown here are slightly cropped, evenly spaced 1 mm transversal slices with W800L-900.

Statistical analysis

Differences between the two cohorts regarding population characteristics (age, sex, smoking status, and pack-years) were tested with a χ^2 -test for categorical characteristics. For continuous variables, a t-test or a Wilcoxon rank-sum test was used, depending on the normality of the distribution. Normality was tested with the Kolmogorov-Smirnov test.

Krippendorff's alpha was used to estimate the inter-observer agreement and intra-observer agreement. Similar to Cohen's Kappa, an alpha of -1 indicates inverse agreement, 0 indicates an absence of apparent agreement, and a value of 1 indicates perfect agreement [223]. Alpha > 0.66 is commonly cited as acceptable agreement, although strict cut-offs are generally discouraged [223, 224]. Despite the recommendation against strict cut-offs, the same ranges are generally used to convert a numerical result to interpretative agreement labels: 0.0–0.2 none to slight, 0.2–0.4 fair, 0.4–0.6 moderate, 0.6–0.8 substantial, 0.8–1.0 excellent. To account for the 15 scores per case for the ordinal scores (five lobes and three

subtypes), a bootstrapping procedure was used to determine the confidence interval (CI), as well as the value of alpha itself. In addition to Krippendorff's alpha, a Bland-Altman analysis was used to explore the inter-observer agreement between the readers and to assess the overall agreement.

To explore the lobar distribution of emphysema, the percentage of participants with more than trace emphysema (of any subtype) per lobe was calculated and displayed in a diagram. The standard deviation (SD) of this percentage was used to express the heterogeneity of this distribution numerically.

Comparisons of emphysema sum scores between categories were visualised with violin plots, as they provide the summarisation of a boxplot, without hiding the actual values. For violin plots of the emphysema sum score, differences between pairs of violins (i.e. Fleischner categories) were tested with t-tests or Wilcoxon rank-sum tests, and the overall significance was tested with an ANOVA or with a Kruskal-Wallis test, depending on normality. Non-parametric tests were used if the sum scores for any of the categories were not normally distributed. The difference in SD between groups was tested with Levene's test.

To assess the direct clinical impact of the proposed classification system, the pulmonary function test (PFT) results (FEV₁ and GOLD stage) were compared to the Fleischner society criteria and the emphysema sum score. This comparison, reported with two violin plots, a heatmap plot, and a correlation plot, is included in the supplement. The PFT results were only available for ImaLife, so the 66 participants from that cohort were included in the analysis.

Statistical analysis was performed with R 4.1.2. Data visualisation was performed with MATLAB R2022b.

8.3 Results

Population characteristics

In total, 117 participants were included, with median age 65 years old (range 45 – 83 years old) and 72 % male sex. The selection process explained in Figure 1 resulted in the selection of 42 men (64 %) and 24 women from the ImaLife cohort (starting from a base selection of 204 men and 276 women), and 42 men (82 %) and 9 women from the NELCIN B3 cohort (starting from a base selection of 142 men and 166 women). All excluded individuals had only trace emphysema or had no emphysema. The median age of the two cohorts was similar (65 years). For more details, see Table 2.

Table 2: Cohort characteristics. Values are given as number (percentage), or median (25th–75th percentile). For the pack-years only ever-smokers were considered.

	ImaLife (N=66)	NELCIN B3 (N=51)	p-value
Male sex	42 (64 %)	42 (82 %)	0.043
Age at time of scan	65.0 (60.0 – 72.0)	65.0 (62.0 – 68.0)	0.694
Smoking status			<0.001
Never smokers	8 (12 %)	16 (31 %)	
Former smokers	34 (52 %)	10 (20 %)	
Current smokers	24 (36 %)	25 (49 %)	
Pack-years	17.9 (9.5 – 24.3)	22.5 (10.0 – 40.0)	0.694
Fleischner society criteria conclusion	No/trace: 0 (0 %) Mild: 58 (88 %) Moderate: 6 (9 %) Severe: 2 (3 %) ADE: 0 (0 %)	No/trace: 0 (0 %) Mild: 43 (84 %) Moderate: 2 (4 %) Severe: 6 (12 %) ADE: 0 (0 %)	0.112

Observer agreement

As part of the case selection, participants were classified as either trace or >trace emphysema by two readers. Krippendorff's alpha for this classification was 0.57 (95 % CI 0.39 – 0.74), indicating moderate agreement. The reclassification rates after consensus of the different readers (the percentage of cases whose classification was overruled by the consensus decision) were 9 % and 10 % (Table S8.1 [p. 225]). For the per-lobe ordinal scores (one score for each lobe and subtype), the inter-reader agreement was substantial (Krippendorff's alpha, 0.78, 95 % CI 0.68 – 0.87). The inter-reader agreement calculated separately for each lobe-subtype-combination can be found in Table S8.2 [p. 225]. Calculating a sum score from the ordinal scores can be expected to improve agreement and allows visualising comparisons. Based on the sum score, the agreement on the overall emphysema severity was excellent (Krippendorff's alpha 0.85, 95 % CI 0.82 – 0.87).

The results of the Bland-Altman analysis are shown in Figure 3. The mean difference (bias) between reader 1 and reader 2 was -0.1 . The upper and lower limit of agreement were 7.0 and -7.1 , respectively.

The intra-observer agreement shows substantial agreement for the per-lobe ordinal scores (Krippendorff's alpha 0.69, 95 % CI 0.43 – 0.87). For the emphysema sum score the intra-observer agreement was excellent (Krippendorff's alpha 0.85, 95 % CI 0.80 – 0.89).

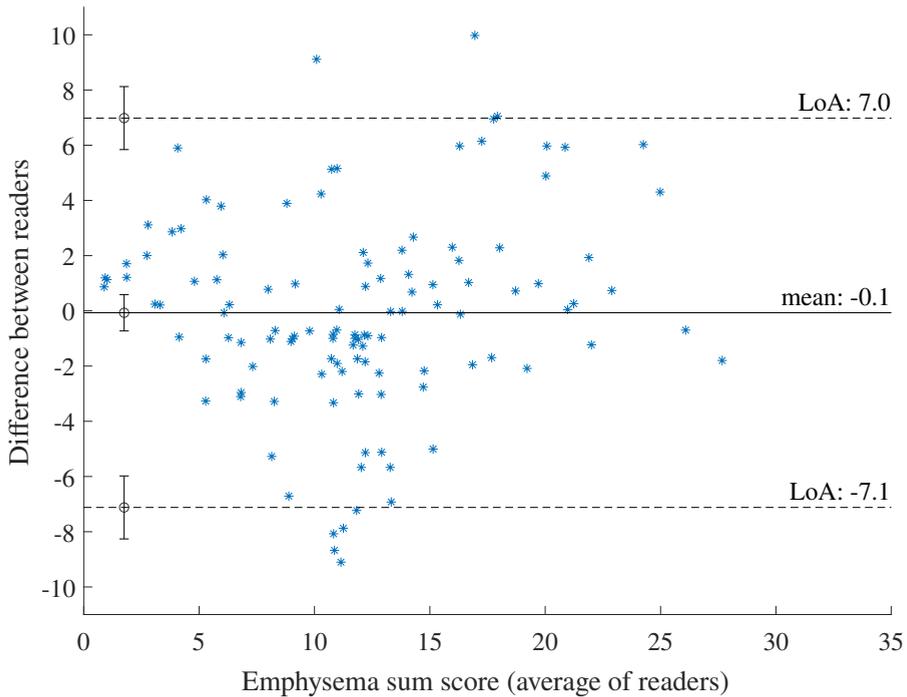


Figure 3: Inter-reader analysis. This figure shows a Bland-Altman plot. The solid line shows the mean difference, while the dotted lines represent limits of agreement. The indicators next to the y-axis show the confidence intervals for the mean difference and the limits of agreement (abbreviated as LoA in the figure). To decrease overlapping of multiple points at the same coordinate, the data was slightly jittered.

Emphysema distribution

Original Fleischner society criteria

According to the original Fleischner society criteria, 70 cases had CLE and 90 cases PSE, no cases had PLE. Most cases had either trace CLE (25 in ImaLife and 17 in NELCIN B3) or mild CLE (28+26). For PSE the most prevalent category was mild (50+37 cases). The 48 (30+18) cases with no or only trace CLE had mild or substantial PSE. Similarly, the 28 (15+13) cases with no PSE had at least mild CLE. When converting the separate criteria to a single category, there were 102 (58+44) cases with mild emphysema, while 16 cases had more severe emphysema.

The inter-reader agreement was fair to moderate (Krippendorff's alpha 0.41, 95% CI 0.30–0.51). There were 44 participants (38%) with a different Fleischner grade between the readers.

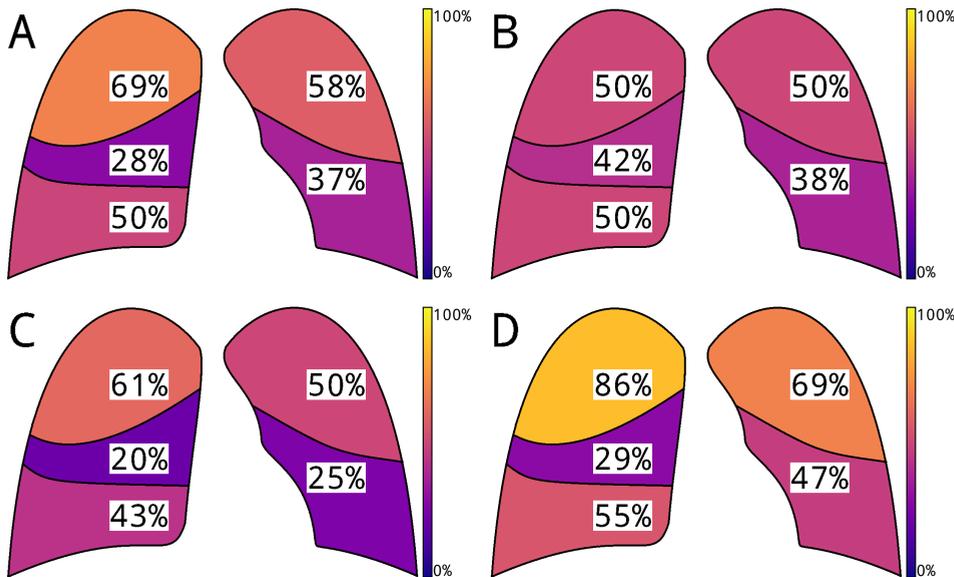


Figure 4: Lobar distribution of emphysema. These diagrams are a visual representation of the percentage of participants with more than trace emphysema of any subtype in each lobe.

A: all participants (N=117), B: never-smokers (N=24), C: former smokers (N=44), D: current smokers (N=49).

Proposed adapted method

When comparing the distribution of (more than trace) emphysema across the lungs, the upper lobes, the right upper lobe in particular, showed the highest emphysema frequency (right upper lobe 69%, left upper lobe 58%, middle lobe 28%, and lower lobes 37%–50%, Figure 4). The SD of the percentage of cases with emphysema in each lobe was 5.9%-point for never-smokers, versus 17 and 22%-point for ex-smokers and current smokers, respectively. This implies the degree of heterogeneity in terms of location and number of affected lobes is strongly related to the smoking status. Diagrams for all 12 combinations of cohort and smoking status are available in the supplementary materials, see Figure S8.3 [p. 229]. Mean emphysema sum score yields a pattern similar to emphysema frequency (data not shown).

There was a wide variety in emphysema distribution for cases with the same nominal severity of emphysema according to the Fleischner category, which is illustrated in Figure 2. To quantitatively assess this spread, the emphysema sum score was used. The violin plot in Figure 5 shows that for the original Fleischner society criteria, a wide range was seen of emphysema sum scores for the mild cases, and a more limited range for moderate and for severe emphysema cases. The spread of the emphysema sum score for mild cases (SD 4.3, range 2–22) was larger than the spread for moderate (3.1, 15–24) or severe cases (3.4, 17–28), although not significantly ($p = 0.481$). The absolute values were significantly dif-

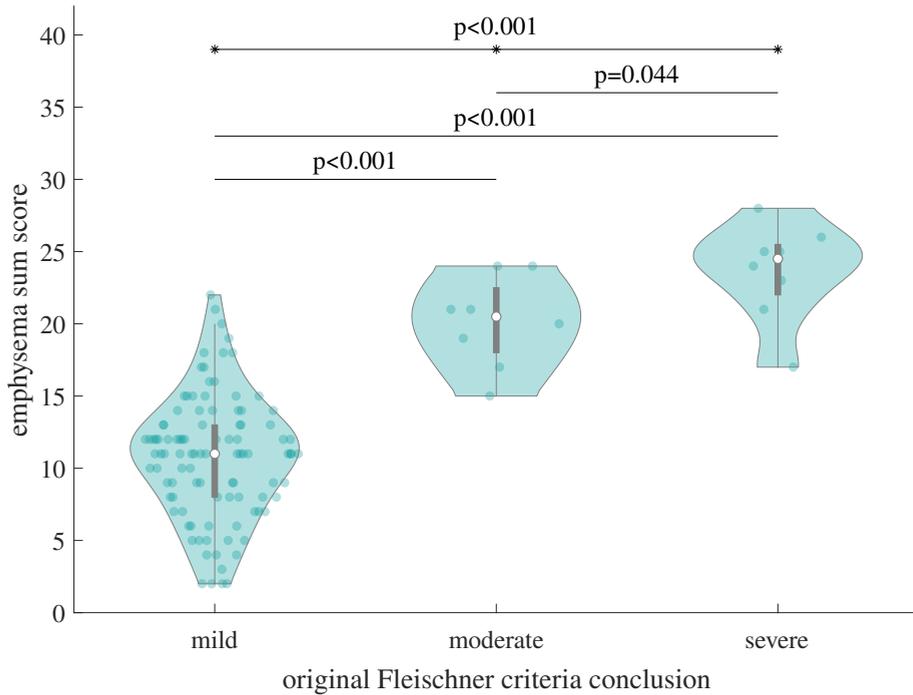


Figure 5: Violin plots of emphysema sum score by Fleischner category. The score on the y-axis is the sum of the emphysema grades of all lobes and subtypes. The category on the x-axis is the conclusion according to the original Fleischner society criteria. The severe category consists of confluent CLE, advanced destructive emphysema, and substantial PSE.

ferent between the mild, moderate, and severe emphysema groups (group means 10.7, 20.1, and 23.6; $p < 0.001$). This persisted when analysing both cohorts separately, see Figure S8.4 [p. 230]. The original Fleischner society criteria conclusion and the smoking status are compared in Table S8.3 [p. 226], showing similar results between the two cohorts.

The NELCIN B3 cohort sample had a lower proportion of women and ex-smokers. All analyses performed on the full dataset were repeated on the two cohorts separately, resulting in only minor differences (see the supplementary materials).

Figure S8.5 [p. 231] shows a violin plot and a heatmap plot comparing the GOLD stage with the emphysema sum score and the Fleischner society criteria. It also shows a correlation plot and a violin plot comparing the FEV₁ with the emphysema sum score and the Fleischner society criteria. Neither classification system showed a strong correlation with the GOLD stage (although both were significant). There is slight downward trend between the FEV₁ and increasing severity of visual emphysema (although neither was significant).

8.4 Discussion

In this proof-of-concept study, we present an extended visual emphysema evaluation method applied to low-dose CT scans. Participants within the same Fleischner category, in particular those with mild emphysema, showed a wide range of emphysema extent, as expressed in the emphysema sum score. This wide range implies that participants with mild emphysema form a heterogeneous group in terms of emphysema extent, with potentially clinically relevant differences. This more detailed analysis had a substantial to excellent inter-reader agreement (Krippendorff's alpha 0.79 and 0.85), which represents a substantial improvement compared to the original Fleischner categories (Krippendorff's alpha 0.41, with non-overlapping confidence intervals). The clinical value of this discrimination is not yet clear, as this is the first study to analyse visual CT emphysema in such detail. Given the link between emphysema severity and lung cancer risk, one of the possible clinical applications to be investigated is to use the proposed system to determine lung cancer screening frequency.

Prior literature – previous classification systems

Prior to the introduction of the Fleischner society criteria, there were other systems to characterise visually assessed emphysema on chest CT. In 1982, Goddard et al. published a scoring system assigning scores ranging from 0 (no emphysema) to 4 (<75 % emphysema) [218]. This system has been used in many other studies, sometimes including the computation of a sum score, and sometimes including 1–5 % as a separate score [219–221]. This system, however, does not distinguish between different subtypes or anatomic distribution of emphysema.

Prior literature – Fleischner society criteria

In the Fleischner society criteria PSE severity is divided into two subcategories, compared to five categories for CLE. However, the severity of each subtype of emphysema, not only CLE, in the different lung lobes can be relevant, e.g. the severity of PSE is one of the factors to determine eligibility for (endobronchial) lung volume reduction [37]. Therefore, in our extended method, there are the same number of severity levels for each emphysema subtype.

Prior literature – inter-reader variability

For subjective analyses like the visual review of CT-detected emphysema, it is important to ensure consistency across readers. An analysis of the COPDGene study (in which the Fleischner society criteria were used) reported the pair-wise kappa for the presence/absence of CLE, ranging from 0.79 to 0.85 [43]. In our study Krippendorff's alpha was 0.57 (95 % CI 0.39 – 0.74). The reason for the lower agreement may be attributed to the difference in methodology: our study did not assess the inter-reader variability for presence/absence of CLE, but for

trace of any emphysema subtype after cases without any emphysema had already been excluded. The same COPDGene study also reported the pair-wise kappa for emphysema grade (ranging from 0.71 to 0.80), which is comparable to the alpha for the emphysema sum score: 0.85 (95% CI 0.82 – 0.87; assuming an interval scale) [43]. An analysis in the SCAPIS trial used a three-point Goddard scale for classification of severity [225]. While they did not report the agreement for severity, the inter-reader agreement metrics they do report (Krippendorff's alpha for presence 0.8; for location 0.75; for subtype 0.73) are similar to Krippendorff's alpha for the emphysema sum score [225]. The lower inter-reader agreement for the original Fleischner society criteria found in this study (Krippendorff's alpha 0.41) is likely due to the calculation method. Since the original scores were not directly assessed, but inferred, this may have emphasised disagreement. The Bland-Altman analysis shows the mean difference (bias) between the readers is limited, whereas the limits of agreement are relatively broad.

Prior literature - lobar distribution

Wille et al. reported on progression of emphysema visible on chest CT in a lung cancer screening trial [221]. In their high-risk population, upper lung zones had relatively higher degrees of more severe emphysema, matching the results from this study. A COPDGene study by Park et al. reported the number of cases with moderate to severe CLE and their mortality rate separated by location predominance [38]. Upper lobe predominant emphysema was approximately six times more prevalent than lower lobe predominant (309 vs 53 cases), although a diffuse pattern was even more common (706 cases). Although specific comparative analyses were not performed, this distribution seems to mostly match the results from this study.

Strengths

One strength of this study is that the participants were selected from two very different population-based cohorts. One cohort is from a region with a mixed rural and semi-urban Western European population, while the other cohort is from a region with a highly urbanised Asian population. Because of the many differences both genetically and environmentally, this should ensure that any insights are applicable to most types of populations. Since the median number of pack-years is within one standard deviation from the mean found in the COPDGene cohort, our cohort should represent a middle ground between a general and high-risk population [43].

Limitations

One limitation of the proposed scoring method is that it requires more time to perform a more detailed analysis. Performing a per-lobe assessment approximately

doubled the time required for visual classification of emphysema on CT.

The inter-reader agreement was slightly lower for the lobar categorical scores (alpha of 0.79, classifying the 3 subtypes for the 5 lobes). The most likely explanation is that there is a large number of different categories, which may exaggerate apparent disagreement.

As shown in Table 2, there were several differences between the two population-based cohort samples used for this study. Despite the genetic differences, the different levels of urbanisation, and differences in sex distribution, only the smoking status seemed to substantially affect the results.

Another limitation concerns the panlobular emphysema subtype. Since the selection was performed after the design of the extension, the absence of PLE was unknown at that time. The current study cannot make inferences about the value of the proposed grading system for PLE.

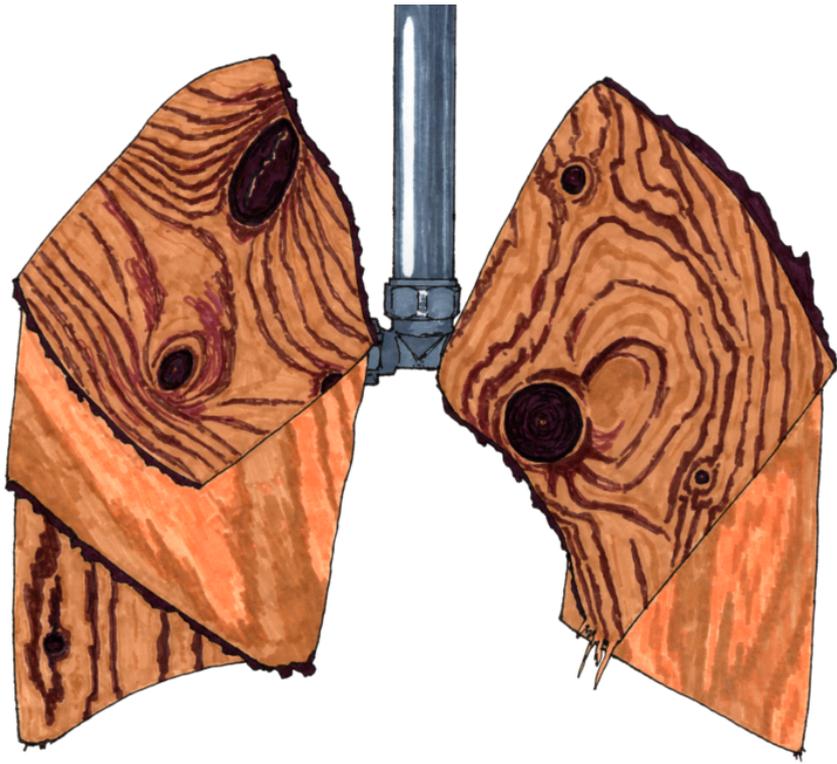
A third limitation concerns the comparison to clinical outcomes. Due to the limited cohort size the comparison between clinical parameters (PFT-based parameters) and visual emphysema classification (with the original Fleischner society criteria and the proposed classification system) did not show conclusive results.

Future research

This study showed there is substantial variation in emphysema sum score between cases marked as the same severity grade by the original Fleischner category. One of the avenues for future research is to determine the value of the additional level of detail. Given that this was only a proof-of-concept study, no large cohort, spirometric data, or long-term follow-up data are available to assess the clinical benefit. Using the emphysema sum score, the scores for each lobe were summed. That means that a participant with low emphysema severity in many lobes will have the same score as a participant with more severe emphysema in fewer lobes. It is not known whether this has a relation to clinical outcome. The best way to assess the clinical benefit will be to first determine a highly detailed score and then test it in large cohorts with a long-term follow-up. Additionally, future studies should use the same acquisition parameters for all CT scans. This would allow including quantitative parameters like the LAV% (the percentage of lung voxels with a density below a specific threshold, generally -950 HU). Since absolute values are highly dependent on acquisition parameters, such analyses were not possible in the present study.

8.5 Conclusion

In this paper, we propose an extended visual emphysema assessment method. The results in this study show that the proposed method is applicable in two different cohorts with different low-dose CT technology, suggesting a broad applicability. The results additionally show the proposed method has good reproducibility. The original Fleischner society criteria classify many cases as mild emphysema, while these cases have very different extents of emphysema, which becomes especially apparent in an emphysema sum score. In never-smokers, emphysema appears more homogeneous, whereas in smokers, there is heterogeneous lobar predominance. The added value of more detailed visual emphysema analysis for clinical outcomes needs to be determined in future studies.



Chapter 9

General discussion

9.1 Main findings

In this thesis, we explored CT-defined emphysema, both in the general population and in a high-risk population. First, the disease burden was assessed, then several technical considerations were discussed, and finally, strategies for improvement of measurements were presented.

In **chapter 2**, we analysed low-dose CT scans from a Chinese community-based lung cancer screening study and a Dutch population-based study. After visual assessment, approximately half of the participants had at least trace emphysema (58.8 % of Chinese and 39.7 % of Dutch participants). Never-smokers in both populations shared older age (adjusted odd ratio (aOR)=1.59 and 1.26) and male sex (aOR=1.50 and 1.93) as risk factors for emphysema presence. After adjusting for smoking, age, and sex, the Chinese participants still had higher odds of emphysema presence (especially centrilobular). This hints at an unmeasured risk factor underlying this difference.

To explore the association between lung cancer and CT-defined emphysema, we performed a systematic review and meta-analysis (**chapter 3**). This analysis comprised 107 082 patients from 21 studies. The analysis showed that presence of emphysema as assessed on CT by visual or quantitative evaluation is associated with higher odds of lung cancer. The overall pooled odds ratio (combining the two methods) was 2.3. The association is present for binary present/absent assessments, as well as for graded assessment based on severity. When considering the sub-types of emphysema, centrilobular emphysema was associated with increased risk for lung cancer, while paraseptal emphysema was not.

In pulmonology it is common to compare the measured pulmonary function to the expected value. In **chapter 4**, a comparison was performed between the gold standard model predicted lung volume and the CT-derived measured lung volume. From the ImaLife study (which is embedded in Lifelines), 173 participants without self-reported or diagnosed lung disease were selected. In this Dutch general population sample, there was a substantial mismatch (mean difference of 1.0 – 1.6L) and high variability (3.2 – 4.2 L) when comparing the predicted and measured lung volume. The predicted lung volume substantially overestimated the CT-derived lung volume, with low precision and accuracy.

There are several parameters that affect image quality in CT imaging. For our phantom study in **chapter 5**, we varied five different CT acquisition and reconstruction parameters to evaluate the effect of dose reduction while preserving image quality. With a newly developed image quality criterion, it was possible to determine the minimum radiation dose required to achieve sufficient image quality for emphysema quantification. The results from **chapter 5** showed that intermediate filtering with deep learning and iterative reconstruction allows reducing the

dose by 85 % (from 1.32 mGy) while maintaining sufficient image quality. These findings were validated in a patient study in **chapter 6** with pairs of ultra-low-dose CT scans and CT scans acquired with the standard-of-care, regular-dose protocol. Ultra-low-dose CT resulted in a slight underestimation of the quantified emphysema compared to the standard-of-care protocol. Intermediate noise reduction (either deep-learning based or with iterative reconstruction) reduced the measurement variability by 24–27 %.

To quantify image noise on CT, it is common to measure the standard deviation (SD) of a circle of a homogeneous region. In **chapter 7**, the air in the trachea was segmented to obtain the ground truth for the image noise. This was compared with measurements based on circular regions of interest (ROIs) and with spherical volumes of interest (VOIs). Using a VOI instead of an ROI reduced the variability by 40–53 %, without substantially affecting the difference to the ground truth or the measurement time.

The Fleischner society criteria are the current gold standard method for visual emphysema assessment. This classification system provides guidelines to assess emphysema using broad categories. We hypothesised a more detailed categorisation might uncover potentially clinically relevant differences. **Chapter 8** presents a proof of concept for an extended visual classification. The extended classification was applied to each lung lobe separately and an emphysema sum score was calculated. This analysis showed diverse emphysema severity in individuals in the same Fleischner category. Despite the additional granularity of the assessment, the inter-reader and intra-reader agreement were excellent (Krippendorff's alpha 0.85 for both).

9.2 Quantitative emphysema assessment

Within a few years of development of the CT scanner in 1973, physicians were exploring the possible application of quantifying emphysema severity on CT, and within fifteen years the gold standard method for quantification of emphysema was developed [42, 226–228]. The method developed by Müller et al. relies on the fact that air has a lower density than lung tissue, and emphysematous changes mean that healthy pulmonary tissue disappears and is replaced by air. By measuring the percentage of lung voxels below a specific threshold, a low attenuation value percentage (LAV%) can be calculated, which indicates the percentage of lung comprising emphysema. Because this method is based on the density of tissue, this method is sometimes called densitometry. During the following decade, emphysema as quantified on CT was shown to correlate with spirometric parameters. While modern scanners have improved in many ways, this correlation has remained approximately the same [229–231]. The initial threshold was –910 HU, but the current consensus is that a threshold of –950 HU is better to differentiate

emphysema from healthy lung tissue [42, 79].

This assumption, however, depends on the reliability of the measured HU value, which may vary between CT models and vendors, as well as the acquisition and reconstruction parameters [232]. While these factors may cause an absolute difference in HU value, image noise may also cause voxels to be misclassified as either belonging to healthy lung tissue or emphysematous tissue. Since image noise is mainly dependent on dose, the tube voltage (the kVp) and the tube current time product (the mAs) are important in determining the suitability of an image for quantitative analysis.

Another way to reduce image noise is to employ advanced reconstruction techniques or post-reconstruction filtering. After the introduction of iterative reconstruction, it became apparent that this would allow a substantial dose reduction, although iterative reconstruction may affect the mean density of the resulting CT image [172, 173]. Iterative reconstruction tends to have a local averaging (or ‘blurring’) effect, which will cause underestimation of low amounts of emphysema and overestimation of large amounts of emphysema.

In this thesis, five CT acquisition and reconstruction parameters relevant for image quality were identified and varied: tube voltage (kVp), tube current time product (mAs), reconstruction filter kernel, iterative reconstruction strength (ADMIRE), and post-hoc deep learning noise reduction using a non-iterative technique artificial neural network (DLNR or NiTANN). In **chapter 5**, a phantom was used to investigate the potential for dose reduction. In this chapter we found that ADMIRE 3 combined with applying the NiTANN allows reducing the $CTDI_{vol}$ by 85 %. Such a substantial reduction requires verification in human subjects, which was performed in **chapter 6**. The results in this thesis show that a drastic reduction in dose is possible. While there is a slight underestimation of the emphysema on ultra-low-dose CT, this was limited to 1.5 LAV%(ADMIRE 3) and 2.9 LAV%(DLNR 3). It is therefore fair to conclude these filtered scans achieve quantitative results similar to the clinical reference.

A previous study assessing the effect of modern iterative reconstruction methods concluded that quantitative determination of emphysema is possible on CT scans at dose levels equivalent to chest X-ray [157]. This and other studies caution that it is important to use the same parameters for consecutive scans if measurements are compared between different time-points [157, 233]. Later studies analysing the effect of reconstruction kernels echo this warning and suggest that the -950 HU threshold might need adjustment based on the reconstruction parameters [1, 178]. The optimal threshold would then have to be determined for each scanner and each set of acquisition and reconstruction parameters.

Instead of adjusting the threshold, changing the quantification method may also be considered. A promising method to reduce the impact of image noise on emphysema quantification was proposed by Heussel et al. [234]. This method (YACTA; ‘Yet Another CT Analyzer’) uses two thresholds; a low threshold to

determine which voxels are definitely emphysematous, and a high threshold to exclude voxels that are definitely not emphysematous. Any voxels between the two thresholds are labelled according to neighbouring voxels. This effectively adjusts the LAV% threshold dynamically to correct for image noise. YACTA is conceptually similar to the Canny edge detection method used in image processing. It is likely less sensitive to noise than the standard single-threshold method, but since the currently available software comparisons did not use scans with varying noise levels, future research is required to confirm this hypothesis [235, 236].

9.3 Visual emphysema assessment

The advantages of CT over chest radiographs for the assessment of emphysema, the higher sensitivity in particular, became apparent quickly after the introduction of CT [227, 228]. To remove part of the subjectivity of a visual assessment, Goddard et al. published a semi-quantitative structured method to assess emphysema severity [218]. This method uses a 5-point Likert scale based on the percentage of affected tissue, from 0 (no emphysema) to 4 (>75 % emphysema). This is then applied to several sections separately, adding up the scores. For the first CT scanners ‘sections’ referred to single slices acquired at regular intervals (e.g. at three or four levels), while for more modern scanners this term is meant more anatomically (e.g. dividing the lung in an apical, middle, and basal section). This system has been used often, either completely unmodified or with slight changes [219–221, 233]. The downside of such a semi-quantitative scale is that it does not leave room to distinguish different subtypes of emphysema. Since the subtype affects morbidity and mortality, the Fleischner society proposed a visual scoring system that has separate definitions for each subtype [40]. The Fleischner criteria captures additional information compared to the Goddard scale. However, it provides a relatively coarse grading mechanism that does not consider differences in severity and subtype between lung lobes.

As shown in the COPDGene study, the mortality hazard ratio increases with severity of Fleischner grade [43]. In addition to this, the lobar distribution of emphysema may impact the all-cause mortality, as a recent study showed that lower-lobe predominance had worse outcomes than upper-lobe predominance [38]. This suggests a detailed characterisation of emphysema may have implications beyond determining the severity of emphysema.

In **chapter 8**, we proposed an extended classification and applied it to each lobe separately. The upper lobe predominance of emphysema in smokers confirms previous findings [221]. We further calculated an emphysema sum score for each participant by adding up all per-lobe and per-subtype scores. We found that cases within the same Fleischner category have wide ranges of emphysema sum scores. This shows that our proposed criteria allow recording the expert opinion with a

high level of detail. Given the excellent inter-reader agreement and intra-reader agreement, it is likely these results reflect actual additional information, at the cost of more reading time per case.

9.4 Implications and future perspectives

Emphysema and lung cancer screening

Lung cancer screening is currently implemented in some European countries, and is under consideration in other countries and regions, with consensus recommendations for implementation [156, 237]. Such a screening programme involves performing periodic low-dose CT scans, on which early signs of lung cancer (lung nodules) may be apparent, in current or former heavy smokers. Since CT uses potentially harmful radiation, a balance must be struck between detecting early lung cancer and limiting the number of CT scans [238, 239]. An important part of this balance is to have an accurate lung cancer risk assessment to optimise both the selection criteria and the screening interval. Currently, the selection criteria are based on age and smoking behaviour, and the screening interval is either fixed or dependent on CT findings [237, 240].

As shown in this thesis and in prior literature, there is a link between emphysema and lung cancer risk, although only for one of the subtypes of emphysema. Even after adjusting for quantitative emphysema, the (all-cause) mortality is increased when high grades of visually assessed emphysema are present [43].

Combining the need for accurate lung cancer risk assessment and the association with emphysema, it may be worthwhile to go beyond the traditional risk factors of age and smoking when estimating the optimal lung cancer screening selection criteria and screening interval. Using a baseline low-dose CT scan to guide screening frequency has already been proposed, one such proposal going as far as suggesting excluding all participants without emphysema from further screening [241]. It is likely worthwhile assessing more nuanced criteria. The NELSON-POP study (a spin-off of the Dutch-Belgian lung cancer screening trial with a focus on personalised outcome prediction) will consider many factors, including genetic, environmental, and CT-based biomarkers, aiming to optimise personalised lung cancer screening [242]. Perhaps a machine learning approach will allow fine-tuning the selection criteria and screening interval. When emphysema is indeed included as a factor for fine-tuning, we suggest, based on our results in **chapter 3**, to stratify per subtype of emphysema for a personalised optimal screening interval.

Given the association between emphysema severity, lung cancer risk and mortality, future research should determine the value of using the extended criteria proposed in **chapter 8**, particularly the emphysema sum score. Since quantitative analyses depend on many parameters while visual analysis (under normal circumstances) does not, the emphysema sum score (or a metric like it) can be expected to provide

repeatable results. Since this thesis only presents a proof of concept, additional studies are required before strong recommendations can be made.

9.5 Visual or quantitative analysis

After considering the findings presented in this thesis, the question may arise which type of assessment should be used in future research and in clinical workup. From the results in this thesis and from prior literature, it is clear that both visual and quantitative analysis results are correlated with relevant outcomes like lung cancer risk and (all-cause) mortality. While visual analysis is less susceptible to image noise present in low-dose CT, it causes a higher workload than an automated quantitative analysis.

Structured reporting frameworks, like the Fleischner criteria and the extended criteria proposed in this thesis, may help reduce the impact of the higher workload in two ways. First, by providing a rigid framework, the visual assessment changes from assessing and describing (a cognitively intensive task) to categorising. Second, it allows training dedicated staff to perform the visual emphysema assessments.

The answer to the question which analysis should be preferred therefore depends on what workload is deemed acceptable, what resources are available and what output is preferred and how they relate to clinical outcomes.

Future technologies

In the rapidly evolving field of CT analysis, it is important to consider how new technologies affect current practices and methods, as well as proposed methods.

In **chapter 5**, we concluded that reducing the radiation dose in chest CT by 85 % is possible without loss of acceptable image quality for densitometry. Using advanced noise reduction techniques may impose limits on the usage of the resulting images. In **chapter 6**, we established that these settings yield similar results for quantitative emphysema analyses. Similar studies should confirm the suitability of these ultra-low-dose filtered images for other applications, structures and diseases. The resulting images may, for example, no longer be suitable for a visual review of emphysema or lung nodule quantification. If the results from these chapters are applicable to other applications and diseases, using such a low radiation dose will alleviate concerns about frequent screening using CT.

The studies in this thesis were largely based on the 7th and 8th generation of CT scanners (multidetector CT and dual source CT). The 9th generation of CT scanners (photon counting CT) are expected to allow a substantial further dose reduction while maintaining image quality. Combining the advanced noise reduction techniques discussed in this thesis with photon counting CT scanners may allow an even greater noise reduction.

In the current technological landscape, the role of artificial intelligence (AI) cannot be left undiscussed. Each application of AI can be placed on a sliding scale from measuring a traditional biomarker to the prediction of a clinical outcome.

An example of measuring a traditional biomarker is a study that showed good performance for emphysema detection based on minimum intensity projections [5]. The main benefit of AI on this end of the scale is either by requiring less computational power, or by requiring less human interaction. While lung segmentation is a computationally easy task in healthy subjects, it may require substantial manual interaction in patients due to patient-related factors like emphysema, consolidations or pleural effusions [243]. Therefore, even on ‘solved’ problems like emphysema detection and lung segmentation, AI can still provide a benefit.

An example towards the other end of the scale can be found in a study that used AI to predict the Fleischner category for centrilobular emphysema [244]. The AI developed for this study reached moderate agreement with the human readers ($\kappa=0.60$). Since human readers reach a better agreement ($\kappa=0.71-0.80$ in a study from the same consortium [43]) it is apparent the visual analysis cannot be replaced by the current AI systems. Disease-specific augmentation, like using minimum intensity projections, may improve the performance of AI systems [5].

Implementation of improved methods

In this thesis we presented two CT measurement methods that extend the currently used standard methods.

In **chapter 7**, we presented a method to measure image noise. In this chapter we compared a ground truth segmentation to the current standard method and an improved method. The current standard method is to draw a circular ROI in a uniform region (e.g. the airways) and record the standard deviation of the measured pixels. By using a VOI (a spherical region) instead, the number of included pixels can be greatly increased, improving the repeatability. This improved method does not require more time or extra training — it only requires implementation in clinical scan reading software. Some software suites already include this as an option. While further validation may strengthen the argument for using a VOI, the evidence presented in this thesis is already sufficient to recommend implementation.

The second method presented in this thesis is an extension of the Fleischner society criteria for the visual assessment of emphysema. Given the association between emphysema severity and lung cancer risk as well as all-cause mortality, an accurate assessment of emphysema has many potential applications. The Fleischner criteria provide a framework for a quick categorisation in broad categories. In **chapter 8** an extension is presented, allowing for a more detailed categorisation at the cost of an increase in reading time.

The results in this thesis are not sufficient to recommend immediate implementation of the extended classification in clinical practice. However, the results of this

proof of principle imply further research is warranted. The original Fleischner criteria provide broad categories, and combining visual and quantitative emphysema assessment is an improvement on using either separately. Taking both facts into consideration, the extended criteria may provide a middle ground, both providing a granular result and being relatively insensitive to technical CT acquisition and reconstruction parameters.

Future research should determine whether the time investment of an extended visual analysis results in a substantially improved assessment of both disease burden and comorbidity risk.

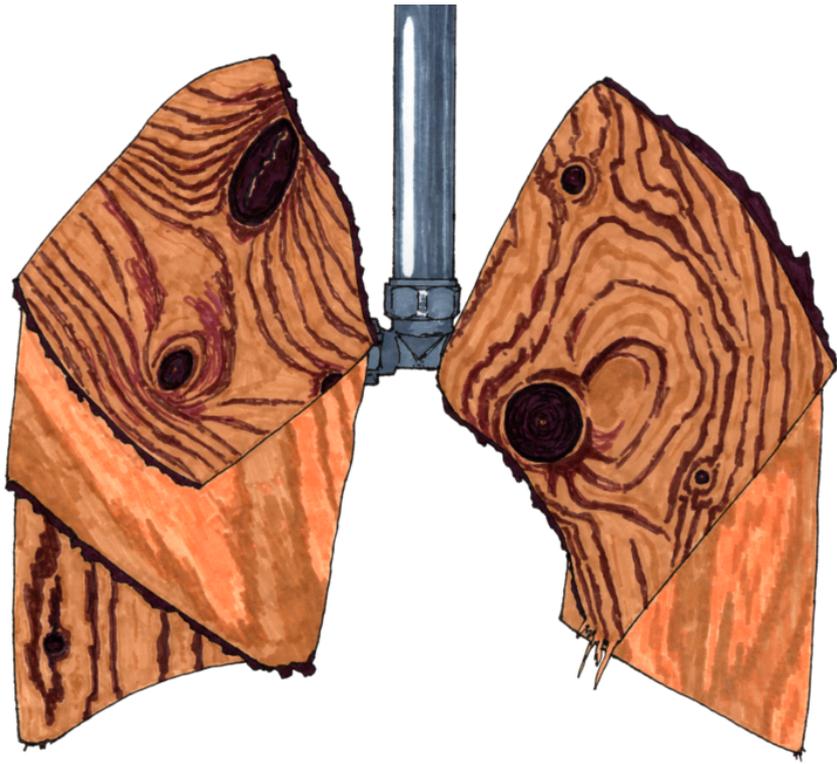
9.6 Conclusions

This thesis explores CT-defined emphysema in both the general and in a high-risk population. The presence as well as the severity of emphysema are risk factors for lung cancer. It is therefore important to assess the lung cancer risk when emphysema is detected on CT scans, both in clinical and screening settings.

Additionally, the effects of CT acquisition and reconstruction parameters on the quantification of emphysema were discussed. This thesis shows that it is possible to lower radiation dose by as much as 85 % when image noise filtering methods are used, without compromising image quality.

Finally, this thesis presents potential improvements for two separate types of measurements. Based on this, first, this thesis advocates the use of a simple yet effective volumetric assessment for noise measurements, which improves reliability without substantially affecting the measurement time. The second is an extension of the current method to visually assess emphysema. This extended classification system captures the expert opinion in more detail, potentially uncovering clinically relevant differences.

The work described in this thesis can be used to customise evaluation of emphysema on CT, potentially personalise lung cancer screening regimens, and to optimise clinical CT protocols to do more with less radiation.



Chapter 10

Bibliography

List of coauthored publications

Any updates in the list of publications after the printing of this thesis can be found in the ORCID record with the ID 0000-0002-1706-5076.

Full references including the DOI can be found in the list of references.

- [1] Effect of chest computed tomography kernel use on emphysema score in severe chronic obstructive pulmonary disease patients evaluated for lung volume reduction,
J.T. Bakker, K. Klooster, H.J. Wisselink, G.J. Pelgrim, R. Vliegenthart, and D.J. Slebos,
Respiration, 2023.
- [2] Predicted versus CT-derived total lung volume in a general population: The ImaLife study,
H.J. Wisselink, D.J. Steerenberg, M. Rook, G.J. Pelgrim, M.A. Heuvelmans, M. van den Berge, G.H. de Bock, and R. Vliegenthart,
PLoS ONE, 2023
- [3] CT-based emphysema characterization per lobe: a proof of concept,
H.J. Wisselink, X. Yang, M. Rook, M.A. Heuvelmans, W. Jiang, J. Zhang, Y. Du, M. Vonder, M.D. Dorrius, Z. Ye, G.H. de Bock, and R. Vliegenthart,
European Journal of Radiology, 2023
- [4] Association between chest CT-defined emphysema and lung cancer: a systematic review and meta-analysis,
X. Yang, H.J. Wisselink, R. Vliegenthart, M.A. Heuvelmans, H.J. Groen, M. Vonder, M.D. Dorrius, and G.H. de Bock,
Radiology, 2022
- [5] AI-driven model for automatic emphysema detection in low-dose computed tomography using disease-specific augmentation
Y. Nagaraj, H.J. Wisselink, M. Rook, J. Cai, S.B. Nagaraj, G. Sidorenkov, R. Veldhuis, M. Oudkerk, R. Vliegenthart, and P.M.A. van Ooijen,
Journal of Digital Imaging, 2022
- [6] COPD identification and grading based on deep learning of lung parenchyma and bronchial wall in chest CT images,
L. Zhang, B. Jiang, H.J. Wisselink, R. Vliegenthart, and X. Xie,
British Journal of Radiology, 2022
- [7] Improved precision of noise estimation in CT with a volume-based approach,
H.J. Wisselink, G.J. Pelgrim, M. Rook, I. Dudurych, M. van den Berge, G.H. de Bock, and R. Vliegenthart,
European Radiology Experimental, 2021

- [8] Ultra-low-dose CT combined with noise reduction techniques for quantification of emphysema in COPD patients: an intra-individual comparison study with standard-dose CT,
H.J. Wisselink, G.J. Pelgrim, M. Rook, K. Imkamp, P.M.A. van Ooijen, M. van den Berge, G.H. de Bock, and R. Vliegenthart,
European Journal of Radiology, 2021
- [9] Potential for dose reduction in CT emphysema densitometry with post-scan noise reduction: a phantom study,
H.J. Wisselink, G.J. Pelgrim, M. Rook, M. van den Berge, K. Slump, Y. Nagaraj, P.M.A. van Ooijen, M. Oudkerk, and R. Vliegenthart,
British Journal of Radiology, 2019
- [10] Early imaging biomarkers of lung cancer, COPD and coronary artery disease in the general population: rationale and design of the ImaLife (Imaging in Lifelines) study,
C. Xia, M. Rook, G.J. Pelgrim, G. Sidorenkov, H.J. Wisselink, J.N. van Bolhuis, P.M.A. van Ooijen, J. Guo, M. Oudkerk, H. Groen, M. van den Berge, P. van der Harst, H. Dijkstra, M. Vonder, M.A. Heuvelmans, M.D. Dorrius, P.P. De Deyn, G.H. de Bock, A. Dotinga, R. Vliegenthart,
European journal of epidemiology, 2020
- [11] Quantitative performance evaluation of ^{124}I PET/MRI lesion dosimetry in differentiated thyroid cancer,
R. Wierds, W. Jentzen, H.H. Quick, H.J. Wisselink, I.N.A. Pooters, J.E. Wildberger, K. Herrmann, G.J. Kemerink, W.H. Backes, and F.M. Mottaghy,
Physics in Medicine and Biology, 2018

References

- [1] J. T. Bakker, K. Klooster, H. J. Wisselink, G. J. Pelgrim, R. Vliegenthart, and D.-J. Slebos, “Effect of Chest Computed Tomography Kernel Use on Emphysema Score in Severe Chronic Obstructive Pulmonary Disease Patients Evaluated for Lung Volume Reduction,” *Respiration*, vol. 102, no. 2, pp. 164–172, 2023. [Online]. Available: <http://dx.doi.org/10.1159/000528628>
Cited on pages 2, 132, and 140.
- [2] H. J. Wisselink, D. J. Steerenberg, M. Rook, G.-J. Pelgrim, M. A. Heuvelmans, M. van den Berge, G. H. de Bock, and R. Vliegenthart, “Predicted versus CT-derived total lung volume in a general population: The ImaLife study,” *PLoS ONE*, vol. 18, no. 6, p. e0287383, 2023. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0287383>
Cited on pages 2, 49, and 140.
- [3] H. J. Wisselink, X. Yang, M. Rook, M. A. Heuvelmans, W. Jiang, J. Zhang, Y. Du, M. Vonder, M. D. Dorrius, Z. Ye, G. H. de Bock, and R. Vliegenthart, “CT-based emphysema characterization per lobe: a proof of concept,” *European Journal of Radiology*, 2023. [Online]. Available: <http://dx.doi.org/10.1016/j.ejrad.2023.110709>
Cited on pages 2, 111, and 140.
- [4] X. Yang, H. J. Wisselink, R. Vliegenthart, M. A. Heuvelmans, H. J. Groen, M. Vonder, M. D. Dorrius, and G. H. de Bock, “Association between chest CT-defined emphysema and lung cancer: a systematic review and meta-analysis,” *Radiology*, p. 212904, 2022. [Online]. Available: <http://dx.doi.org/10.1148/212904>
Cited on pages 2, 13, 29, 113, and 140.
- [5] Y. Nagaraj, H. J. Wisselink, M. Rook, J. Cai, S. B. Nagaraj, G. Sidorenkov, R. Veldhuis, M. Oudkerk, R. Vliegenthart, and P. M. A. van Ooijen, “AI-driven model for automatic emphysema detection in low-dose computed tomography using disease-specific augmentation,” *Journal of Digital Imaging*, 2022. [Online]. Available: <http://dx.doi.org/10.1007/s10278-022-00599-7>
Cited on pages 2, 136, and 140.
- [6] L. Zhang, B. Jiang, H. J. Wisselink, R. Vliegenthart, and X. Xie, “COPD identification and grading based on deep learning of lung parenchyma and bronchial wall in chest CT images,” *The British Journal of Radiology*, 2022. [Online]. Available: <https://doi.org/10.1259/bjr.20210637>
Cited on pages 2 and 140.
- [7] H. J. Wisselink, G. J. Pelgrim, M. Rook, I. Dudurych, M. van den Berge, G. H. de Bock, and R. Vliegenthart, “Improved precision

- of noise estimation in CT with a volume-based approach,” *European Radiology Experimental*, vol. 5, no. 1, pp. 1–7, 2021. [Online]. Available: <http://dx.doi.org/10.1186/s41747-021-00237-x>
Cited on pages 2, 99, and 140.
- [8] H. J. Wisselink, G. J. Pelgrim, M. Rook, K. Imkamp, P. M. A. van Ooijen, M. van den Berge, G. H. de Bock, and R. Vliegthart, “Ultra-low-dose CT combined with noise reduction techniques for quantification of emphysema in COPD patients: an intra-individual comparison study with standard-dose CT,” *European Journal of Radiology*, vol. 138, p. 109646, 2021. [Online]. Available: <http://dx.doi.org/10.1016/j.ejrad.2021.109646>
Cited on pages 2, 83, and 141.
- [9] H. J. Wisselink, G. J. Pelgrim, M. Rook, M. van den Berge, K. Slump, Y. Nagaraj, P. van Ooijen, M. Oudkerk, and R. Vliegthart, “Potential for dose reduction in CT emphysema densitometry with post-scan noise reduction: a phantom study,” *The British Journal of Radiology*, vol. 93, no. 1105, p. 20181019, 2019, pMID: 31724436. [Online]. Available: <http://dx.doi.org/10.1259/bjr.20181019>
Cited on pages 2, 65, 85, 96, 101, and 141.
- [10] C. Xia, M. Rook, G. J. Pelgrim, G. Sidorenkov, H. J. Wisselink, J. N. van Bolhuis, P. M. van Ooijen, J. Guo, M. Oudkerk, H. Groen *et al.*, “Early imaging biomarkers of lung cancer, COPD and coronary artery disease in the general population: rationale and design of the ImaLife (Imaging in Lifelines) Study,” *European journal of epidemiology*, vol. 35, no. 1, pp. 75–86, 2020. [Online]. Available: <http://dx.doi.org/10.1007/s10654-019-00519-0>
Cited on pages 2, 6, 13, 52, 53, 54, 114, and 141.
- [11] R. Wierds, W. Jentzen, H. H. Quick, H. J. Wisselink, I. N. A. Pooters, J. E. Wildberger, K. Herrmann, G. J. Kemmerink, W. H. Backes, and F. M. Mottaghy, “Quantitative performance evaluation of ¹²⁴I PET/MRI lesion dosimetry in differentiated thyroid cancer,” *Physics in Medicine and Biology*, vol. 63, no. 1, 2018.
Cited on pages 2 and 141.
- [12] M. A. Heuvelmans, M. Vonder, M. Rook, H. J. Groen, G. H. De Bock, X. Xie, M. J. Ijzerman, R. Vliegthart, and M. Oudkerk, “Screening for early lung cancer, chronic obstructive pulmonary disease, and cardiovascular disease (the Big-3) using low-dose chest computed tomography,” *Journal of thoracic imaging*, vol. 34, no. 3, pp. 160–169, 2019. [Online]. Available: <http://dx.doi.org/10.1097/RTI.0000000000000379>
Cited on page 2.

- [13] World Health Organization, “Global Health Estimates 2019: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019,” published 2020. Available at http://web.archive.org/web/20220416005128/https://cdn.who.int/media/docs/default-source/gho-documents/global-health-estimates/ghe2019_cod_global_2000_20194e572f53-509f-4578-b01e-6370c65d9fc5_3096f6a3-0f82-4c0c-94e2-623e802527c8.xlsx?sfvrsn=eaf8ca5_7.

Cited on pages 2, 31, and 113.

- [14] United States Surgeon General’s Advisory Committee on Smoking and Health, “Smoking and Health,” *Dept of Health, Education and Welfare, Washington, DC*, 1964. [Online]. Available: <http://web.archive.org/web/20220809044143/https://collections.nlm.nih.gov/ext/document/101584932X202/PDF/101584932X202.pdf>

Cited on page 2.

- [15] D. M. Mannino and A. S. Buist, “Global burden of COPD: risk factors, prevalence, and future trends,” *The Lancet*, vol. 370, no. 9589, pp. 765–773, 2007. [Online]. Available: [http://dx.doi.org/10.1016/S0140-6736\(07\)61380-4](http://dx.doi.org/10.1016/S0140-6736(07)61380-4)

Cited on pages 2 and 13.

- [16] J. Keto, H. Ventola, J. Jokelainen, K. Linden, S. Keinänen-Kiukaanniemi, M. Timonen, T. Ylisaukko-Oja, and J. Auvinen, “Cardiovascular disease risk factors in relation to smoking behaviour and history: a population-based cohort study,” *Open Heart*, vol. 3, no. 2, p. e000358, 2016. [Online]. Available: <http://dx.doi.org/10.1136/openhrt-2015-000358>

Cited on page 2.

- [17] A. Durawa, K. Dziadziuszko, M. Jelitto-Górska, and E. Szurowska, “Emphysema - The review of radiological presentation and its clinical impact in the LDCT screening era,” *Clinical Imaging*, vol. 64, pp. 85–91, 2020. [Online]. Available: <http://dx.doi.org/10.1016/j.clinimag.2020.04.018>

Cited on pages 2, 6, and 31.

- [18] A. D. Morgan, R. Zakeri, and J. K. Quint, “Defining the relationship between COPD and CVD: what are the implications for clinical practice?” *Therapeutic Advances in Respiratory Disease*, vol. 12, pp. 1–16, 2018. [Online]. Available: <http://dx.doi.org/10.1177/1753465817750524>

Cited on pages 2 and 6.

- [19] J. Wilson and G. Jungner, “Principles and practice of screening for disease,” on behalf of the World Health Organization. Published 1968. Available at <http://web.archive.org/web/20220922084925/https://apps.who.int/>

iris/bitstream/handle/10665/37650/%20WHO_PHP_34.pdf?sequence=17.
Cited on page 2.

- [20] Z. Wang, Y. Hu, Y. Wang, W. Han, L. Wang, F. Xue, X. Sui, W. Song, R. Shi, and J. Jiang, “Can CT screening give rise to a beneficial stage shift in lung cancer patients? Systematic review and meta-analysis,” *PLoS One*, vol. 11, no. 10, p. e0164416, 2016. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0164416>
Cited on page 2.
- [21] H. J. de Koning, C. M. van der Aalst, P. A. de Jong, E. T. Scholten, K. Nackaerts, M. A. Heuvelmans, J.-W. J. Lammers, C. Weenink, U. Yousaf-Khan, N. Horeweg *et al.*, “Reduced lung-cancer mortality with volume CT screening in a randomized trial,” *New England journal of medicine*, vol. 382, no. 6, pp. 503–513, 2020. [Online]. Available: <http://dx.doi.org/10.1056/NEJMoa1911793>
Cited on page 2.
- [22] A. Sadate, B. V. Océan, J.-P. Beregi, A. Hamard, T. Addala, H. de Forges, P. Fabbro-Peray, and J. Frandon, “Systematic review and meta-analysis on the impact of lung cancer screening by low-dose computed tomography,” *European Journal of Cancer*, vol. 134, pp. 107–114, 2020. [Online]. Available: <http://dx.doi.org/10.1016/j.ejca.2020.04.035>
Cited on page 2.
- [23] M. Vonder, C. M. van der Aalst, R. Vliegthart, P. M. van Ooijen, D. Kuijpers, J. W. Gratama, H. J. de Koning, and M. Oudkerk, “Coronary artery calcium imaging in the ROBINSCA trial: rationale, design, and technical background,” *Academic radiology*, vol. 25, no. 1, pp. 118–128, 2018. [Online]. Available: <http://dx.doi.org/10.1016/j.acra.2017.07.010>
Cited on page 2.
- [24] A. C. P. Diederichsen, L. M. Rasmussen, R. Sjøgaard, J. Lambrechtsen, F. H. Steffensen, L. Frost, K. Egstrup, G. Urbonaviciene, M. Busk, M. H. Olsen *et al.*, “The Danish Cardiovascular Screening Trial (DANCAVAS): study protocol for a randomized controlled trial,” *Trials*, vol. 16, no. 1, pp. 1–11, 2015. [Online]. Available: <http://dx.doi.org/10.1186/s13063-015-1082-6>
Cited on page 2.
- [25] R. Sjøgaard, A. C. P. Diederichsen, L. M. Rasmussen, J. Lambrechtsen, F. H. Steffensen, L. Frost, K. Egstrup, G. Urbonaviciene, M. Busk, and J. S. Lindholt, “Cost effectiveness of population screening vs. no screening for cardiovascular disease: the Danish Cardiovascular Screening trial (DANCAVAS),” *European Heart Journal*, 2022. [Online]. Available: <http://dx.doi.org/10.1093/eurheartj/ehac488>
Cited on pages 2 and 3.

- [26] S. J. Denissen, C. M. Van Der Aalst, M. Vonder, J. W. C. Gratama, H. J. Adriaansen, D. Kuijpers, J. E. Roeters van Lennep, R. Vliegenthart, P. Van Der Harst, R. L. Braam *et al.*, “Screening for coronary artery calcium in a high-risk population: the ROBINSCA trial,” *European journal of preventive cardiology*, vol. 28, no. 10, pp. 1155–1159, 2021. [Online]. Available: <http://dx.doi.org/10.1177/2047487320932263>
Cited on page 2.
- [27] F. A. Mohamed Hoesein, P. Zanen, P. A. de Jong, B. van Ginneken, H. M. Boezen, H. J. Groen, M. Oudkerk, H. J. de Koning, D. S. Postma, and J.-W. J. Lammers, “Rate of progression of CT-quantified emphysema in male current and ex-smokers: a follow-up study,” *Respiratory research*, vol. 14, no. 1, pp. 1–8, 2013. [Online]. Available: <http://dx.doi.org/10.1186/1465-9921-14-55>
Cited on page 2.
- [28] S. B. Shaker, T. Stavngaard, L. C. Laursen, B. C. Stoel, and A. Dirksen, “Rapid fall in lung density following smoking cessation in COPD,” *COPD: Journal of Chronic Obstructive Pulmonary Disease*, vol. 8, no. 1, pp. 2–7, 2011. [Online]. Available: <http://dx.doi.org/10.3109/15412555.2010.541306>
Cited on page 2.
- [29] K. Brain, B. Carter, K. J. Lifford, O. Burke, A. Devaraj, D. R. Baldwin, S. Duffy, and J. K. Field, “Impact of low dose computed tomography screening on smoking cessation among high-risk participants in the UK Lung Cancer Screening trial,” *Thorax*, 2017. [Online]. Available: <http://dx.doi.org/10.1136/thoraxjnl-2016-209690>
Cited on page 2.
- [30] US Preventive Services Task Force, “Screening for Lung Cancer: US Preventive Services Task Force Recommendation Statement,” *JAMA*, vol. 325, no. 10, pp. 962–970, 03 2021. [Online]. Available: <http://dx.doi.org/10.1001/jama.2021.1117>
Cited on page 3.
- [31] C. M. Behr, H. Koffijberg, K. Degeling, R. Vliegenthart, and M. J. IJzerman, “Can we increase efficiency of CT lung cancer screening by combining with CVD and COPD screening? Results of an early economic evaluation,” *European radiology*, vol. 32, no. 5, pp. 3067–3075, 2022. [Online]. Available: <http://dx.doi.org/10.1007/s00330-021-08422-7>
Cited on page 3.
- [32] RIVM, “Trends in het aantal CT-onderzoeken,” 2021. [Online]. Available: <https://web.archive.org/web/20220121024535/https://www.rivm.nl/medische-stralingstoepassingen/trends-en-stand-van->

zaken/diagnostiek/computer-tomografie/trends-in-aantal-ct-onderzoeken
Cited on page 3.

- [33] “Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease (2023 report).” [Online]. Available: http://web.archive.org/web/20230420055509/https://goldcopd.org/wp-content/uploads/2023/03/GOLD-2023-ver-1.3-17Feb2023_WMV.pdf
Cited on page 3.
- [34] F. Gaillard and D. Bell, “Thoracic HRCT terminology,” Reference article, Radiopaedia.org. (last revised on 2019/08/30) DOI: 10.53347/rID-8762. [Online]. Available: <https://web.archive.org/web/20220109211157/https://radiopaedia.org/articles/thoracic-hrct-terminology>
Cited on page 5.
- [35] W. Boron and E. Boulpaep, *Medical Physiology, Second Edition*. Saunders Elsevier, 2012, ch. 26, ISBN: 9781437717532.
Cited on page 5.
- [36] R. Pellegrino, G. Viegi, V. Brusasco, R. Crapo, F. Burgos, R. Casaburi, A. Coates, C. Van Der Grinten, P. Gustafsson, J. Hankinson *et al.*, “Interpretative strategies for lung function tests,” *European respiratory journal*, vol. 26, no. 5, pp. 948–968, 2005. [Online]. Available: <http://dx.doi.org/10.1183/09031936.05.00035205>
Cited on pages 3 and 51.
- [37] D.-J. Slebos, H. Nick, M. Hetzel, F. J. Herth, and P. L. Shah, “Endobronchial coils for endoscopic lung volume reduction: best practice recommendations from an expert panel,” *Respiration*, vol. 96, no. 1, pp. 1–11, 2018. [Online]. Available: <http://dx.doi.org/10.1159/000490193>
Cited on pages 3, 4, and 124.
- [38] J. Park, B. D. Hobbs, J. D. Crapo, B. J. Make, E. A. Regan, S. Humphries, V. J. Carey, D. A. Lynch, E. K. Silverman, C. Investigators *et al.*, “Subtyping COPD by using visual and quantitative CT imaging features,” *Chest*, vol. 157, no. 1, pp. 47–60, 2020. [Online]. Available: <http://dx.doi.org/10.1016/J.CHEST.2019.06.015>
Cited on pages 3, 113, 116, 125, and 133.
- [39] K. Lowe, E. Regan, A. Anzueto, E. Austin, J. Austin, T. Beaty, P. Benos, C. Benway, S. Bhatt, E. Bleecker *et al.*, “COPDGene® 2019: Redefining the Diagnosis of Chronic Obstructive Pulmonary Disease.” *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation*, vol. 6, no. 5, pp. 384–399, 2019. [Online]. Available: <http://dx.doi.org/10.15326/jcopdf.6.5.2019.0149>
Cited on pages 3, 97, and 113.

- [40] D. A. Lynch, J. H. Austin, J. C. Hogg, P. A. Grenier, H.-U. Kauczor, A. A. Bankier, R. G. Barr, T. V. Colby, J. R. Galvin, P. A. Gevenois *et al.*, “CT-definable subtypes of chronic obstructive pulmonary disease: a statement of the Fleischner Society,” *Radiology*, vol. 277, no. 1, pp. 192–205, 2015. [Online]. Available: <http://dx.doi.org/10.1148/radiol.2015141579>
Cited on pages 3, 4, 7, 16, 33, 45, 85, 87, 113, 116, 117, and 133.
- [41] “Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease (2020 report).” [Online]. Available: http://web.archive.org/web/20200311124055/https://goldcopd.org/wp-content/uploads/2019/12/GOLD-2020-FINAL-ver1.2-03Dec19_WMV.pdf
Cited on pages 3, 6, 51, 52, 85, 113, and 114.
- [42] N. L. Müller, C. A. Staples, R. R. Miller, and R. T. Abboud, ““Density mask”: an objective method to quantitate emphysema using computed tomography,” *Chest*, vol. 94, no. 4, pp. 782–787, 1988. [Online]. Available: <http://dx.doi.org/10.1378/chest.94.4.782>
Cited on pages 3, 4, 85, 113, 131, and 132.
- [43] D. A. Lynch, C. M. Moore, C. Wilson, D. Nevrekar, T. Jennermann, S. M. Humphries, J. H. Austin, P. A. Grenier, H.-U. Kauczor, M. K. Han *et al.*, “CT-based Visual Classification of Emphysema: Association with Mortality in the COPD Gene Study,” *Radiology*, vol. 288, no. 3, pp. 859–866, 2018. [Online]. Available: <http://dx.doi.org/10.1148/radiol.2018172294>
Cited on pages 4, 13, 16, 25, 67, 85, 97, 113, 124, 125, 133, 134, and 136.
- [44] W. Webb and C. Higgins, *Thoracic Imaging: Pulmonary and Cardiovascular Radiology*. Wolters Kluwer, 2017, ch. 24, ISBN: 9781496321046.
Cited on page 4.
- [45] R. A. Schmidt, R. W. Glenny, J. D. Godwin, and N. B. Hampson, “Panlobular Emphysema in Young Intravenous Ritalin® Abusers,” *Am Rev Respir Dis*, vol. 143, pp. 649–656, 1991. [Online]. Available: <http://dx.doi.org/10.1164/ajrccm/143.3.649>
Cited on page 4.
- [46] N. Tanabe, D. M. Vasilescu, C. J. Hague, K. Ikezoe, D. T. Murphy, M. Kirby, C. S. Stevenson, S. E. Verleden, B. M. Vanaudenaerde, G. Gayan-Ramirez *et al.*, “Pathological comparisons of paraseptal and centrilobular emphysema in chronic obstructive pulmonary disease,” *American Journal of Respiratory and Critical Care Medicine*, vol. 202, no. 6, pp. 803–811, 2020. [Online]. Available: <http://dx.doi.org/10.1164/rccm.201912-2327OC>
Cited on pages 6 and 113.

- [47] Y. Du, Y. Zhao, G. Sidorenkov, G. H. de Bock, X. Cui, Y. Huang, M. D. Dorrius, M. Rook, H. J. M. Groen, M. A. Heuvelmans *et al.*, “Methods of computed tomography screening and management of lung cancer in Tianjin: design of a population-based cohort study,” *Cancer biology & medicine*, vol. 16, no. 1, p. 181, 2019. [Online]. Available: <http://dx.doi.org/10.20892/j.issn.2095-3941.2018.0237>
Cited on pages 6, 13, 14, and 114.
- [48] G. L. Hall, N. Filipow, G. Ruppel, T. Okitika, B. Thompson, J. Kirkby, I. Steenbruggen, B. G. Cooper, S. Stanojevic *et al.*, “Official ERS technical standard: Global Lung Function Initiative reference values for static lung volumes in individuals of European ancestry,” *European Respiratory Journal*, vol. 57, no. 3, 2021. [Online]. Available: <http://dx.doi.org/10.1183/13993003.00289-2020>
Cited on pages 6, 51, 54, 55, 60, 61, and 62.
- [49] European Commission, *European Guidelines on Quality Criteria for CT*. Office for Official Publications of the European Communities, 2000, short link: <https://op.europa.eu/s/n8PM>. [Online]. Available: <http://web.archive.org/web/20210225144451/https://op.europa.eu/o/opportal-service/download-handler?identifier=d229c9e1-a967-49de-b169-59ee68605f1a&format=pdf&language=en&productionSystem=cellar>
Cited on pages 7 and 101.
- [50] S. M. May and J. T. Li, “Burden of chronic obstructive pulmonary disease: healthcare costs and beyond,” in *Allergy and asthma proceedings*, vol. 36, no. 1. OceanSide Publications, 2015, p. 4. [Online]. Available: <http://dx.doi.org/10.2500/aap.2015.36.3812>
Cited on page 13.
- [51] A. L. Friedlander, D. Lynch, L. A. Dyar, and R. P. Bowler, “Phenotypes of chronic obstructive pulmonary disease,” *COPD: Journal of Chronic Obstructive Pulmonary Disease*, vol. 4, no. 4, pp. 355–384, 2007. [Online]. Available: <http://dx.doi.org/10.1080/15412550701629663>
Cited on page 13.
- [52] L. N. Segal and F. J. Martinez, “Chronic obstructive pulmonary disease subpopulations and phenotyping,” *Journal of Allergy and Clinical Immunology*, vol. 141, no. 6, pp. 1961–1971, 2018. [Online]. Available: <http://dx.doi.org/10.1016/j.jaci.2018.02.035>
Cited on page 13.
- [53] D. Steiger, M. F. Siddiqi, R. Yip, D. F. Yankelevitz, C. I. Henschke, A. Jirapatnakul, R. Flores, A. Wolf, D. M. Libby, J. P. Smith *et al.*, “The importance of low-dose CT screening to identify emphysema in asymptomatic participants with and without a prior diagnosis of COPD,”

Clinical imaging, vol. 78, pp. 136–141, 2021. [Online]. Available: <http://dx.doi.org/10.1016/j.clinimag.2021.03.012>

Cited on pages 13 and 61.

- [54] W. C. Tan, “Trends in chronic obstructive pulmonary disease in the Asia-Pacific regions,” *Current Opinion in Pulmonary Medicine*, vol. 17, no. 2, pp. 56–61, 2011. [Online]. Available: <http://dx.doi.org/10.1097/MCP.0b013e32834316cd>

Cited on page 13.

- [55] E. Wachuła, S. Szabłowska-Siwik, D. Czyzewski, J. Kozielski, and M. Adamek, “Emphysema affects the number and appearance of solitary pulmonary nodules identified in chest low-dose computed tomography: A study on high risk lung cancer screenees recruited in Silesian District,” *Pol. Arch. Intern. Med.*, vol. 130, pp. 17–24, 2019. [Online]. Available: <http://dx.doi.org/10.20452/pamw.14985>

Cited on page 13.

- [56] M. C. Turner, Z. J. Andersen, A. Baccarelli, W. R. Diver, S. M. Gapstur, C. A. Pope III, D. Prada, J. Samet, G. Thurston, and A. Cohen, “Outdoor air pollution and cancer: An overview of the current evidence and public health recommendations,” *CA: a cancer journal for clinicians*, vol. 70, no. 6, pp. 460–479, 2020. [Online]. Available: <http://dx.doi.org/10.3322/caac.21632>

Cited on page 13.

- [57] Y. Zhang, L. Wang, G. M. Mutlu, and H. Cai, “More to Explore: Further Definition of Risk Factors for COPD–Differential Gender Difference, Modest Elevation in PM_{2.5}, and e-Cigarette Use,” *Frontiers in Physiology*, vol. 12, p. 669152, 2021. [Online]. Available: <http://dx.doi.org/10.3389/fphys.2021.669152>

Cited on page 13.

- [58] GBD 2015 Chronic Respiratory Disease Collaborators, “Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015,” *The Lancet. Respiratory Medicine*, vol. 5, no. 9, p. 691, 2017. [Online]. Available: [http://dx.doi.org/10.1016/S2213-2600\(17\)30293-X](http://dx.doi.org/10.1016/S2213-2600(17)30293-X)

Cited on page 13.

- [59] S. Scholtens, N. Smidt, M. A. Swertz, S. J. Bakker, A. Dotinga, J. M. Vonk, F. Van Dijk, S. K. van Zon, C. Wijmenga, B. H. Wolffenbuttel *et al.*, “Cohort Profile: LifeLines, a three-generation cohort study and biobank,” *International journal of epidemiology*, vol. 44, no. 4, pp. 1172–1180,

2015. [Online]. Available: <http://dx.doi.org/10.1093/ije/dyu229>
Cited on pages 14, 52, and 114.
- [60] K. Veldman, U. Bültmann, R. E. Stewart, J. Ormel, F. C. Verhulst, and S. A. Reijneveld, “Mental health problems and educational attainment in adolescence: 9-year follow-up of the TRAILS study,” *PloS one*, vol. 9, no. 7, p. e101751, 2014. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0101751>
Cited on page 14.
- [61] Y. Pan, S. Vayssettes, and E. Fordham, “Education in China: A snapshot,” *Organization for Economic Co-Operation and Development, Paris*, 2016. [Online]. Available: <https://web.archive.org/web/20221027151851/https://www.oecd.org/education/education-in-china-a-snapshot.pdf>
Cited on page 14.
- [62] A. E. Dijkstra, K. de Jong, H. M. Boezen, H. Kromhout, R. Vermeulen, H. J. Groen, D. S. Postma, and J. M. Vonk, “Risk factors for chronic mucus hypersecretion in individuals with and without COPD: influence of smoking and job exposure on CMH,” *Occupational and environmental medicine*, vol. 71, no. 5, pp. 346–352, 2014. [Online]. Available: <http://dx.doi.org/10.1136/oemed-2013-101654>
Cited on page 14.
- [63] K. Martini and T. Frauenfelder, “Advances in imaging for lung emphysema,” *Annals of Translational Medicine*, vol. 8, no. 21, 2020. [Online]. Available: <http://dx.doi.org/10.21037/atm.2020.04.44>
Cited on page 16.
- [64] Y. Li, Y. Du, Y. Huang, Y. Zhao, G. Sidorenkov, M. Vonder, X. Cui, S. Fan, M. D. Dorrius, R. Vliegenthart *et al.*, “Community-based lung cancer screening by low-dose computed tomography in China: First round results and a meta-analysis,” *European Journal of Radiology*, vol. 144, p. 109988, 2021. [Online]. Available: <http://dx.doi.org/10.1016/j.ejrad.2021.109988>
Cited on pages 24 and 114.
- [65] C. Xia, M. Vonder, G. Sidorenkov, R. Ma, M. Oudkerk, P. van der Harst, P. P. De Deyn, and R. Vliegenthart, “Coronary Artery Calcium and Cognitive Function in Dutch Adults: Cross-Sectional Results of the Population-Based ImaLife Study,” *Journal of the American Heart Association*, vol. 10, no. 4, p. e018172, 2021. [Online]. Available: <http://dx.doi.org/10.1161/JAHA.120.018172>
Cited on page 24.
- [66] W. C. Tan, C. J. Hague, J. Leipsic, J. Bourbeau, L. Zheng, P. Z. Li, D. D. Sin, H. O. Coxson, M. Kirby, J. C. Hogg *et al.*, “Findings on thoracic

computed tomography scans and respiratory outcomes in persons with and without chronic obstructive pulmonary disease: a population-based cohort study,” *PloS one*, vol. 11, no. 11, p. e0166745, 2016. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0166745>

Cited on page 24.

- [67] B. El Kaddouri, M. J. Strand, D. Baraghoshi, S. M. Humphries, J.-P. Charbonnier, E. M. van Rikxoort, and D. A. Lynch, “Fleischner Society visual emphysema CT patterns help predict progression of emphysema in current and former smokers: results from the COPDGene study,” *Radiology*, vol. 298, no. 2, p. 441, 2021. [Online]. Available: <http://dx.doi.org/10.1148/radiol.2020200563>

Cited on page 24.

- [68] V. Tejwani, A. Fawzy, N. Putcha, P. J. Castaldi, M. H. Cho, K. A. Pratte, S. P. Bhatt, D. A. Lynch, S. M. Humphries, G. L. Kinney *et al.*, “Emphysema progression and lung function decline among angiotensin converting enzyme inhibitors and angiotensin-receptor blockade users in the COPDGene Cohort,” *Chest*, vol. 160, no. 4, pp. 1245–1254, 2021. [Online]. Available: <http://dx.doi.org/10.1016/j.chest.2021.05.007>

Cited on page 24.

- [69] J. J. Zhang and J. M. Samet, “Chinese haze versus Western smog: lessons learned,” *Journal of thoracic disease*, vol. 7, no. 1, p. 3, 2015. [Online]. Available: <http://dx.doi.org/10.3978/j.issn.2072-1439.2014.12.06>

Cited on page 24.

- [70] A. Van Donkelaar, R. V. Martin, M. Brauer, R. Kahn, R. Levy, C. Verduzco, and P. J. Villeneuve, “Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application,” *Environmental health perspectives*, vol. 118, no. 6, pp. 847–855, 2010. [Online]. Available: <http://dx.doi.org/10.1289/ehp.0901623>

Cited on page 24.

- [71] Y. Liu, K. Lee, R. Perez-Padilla, N. Hudson, and D. Mannino, “Outdoor and indoor air pollution and COPD-related diseases in high-and low-income countries,” *The international journal of tuberculosis and lung disease*, vol. 12, no. 2, pp. 115–127, 2008. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/18230243/>

Cited on page 24.

- [72] H. Qiu, K. Tan, F. Long, L. Wang, H. Yu, R. Deng, H. Long, Y. Zhang, and J. Pan, “The burden of COPD morbidity attributable to the interaction between ambient air pollution and temperature in Chengdu, China,” *International journal of environmental research and public health*, vol. 15, no. 3, p. 492, 2018. [Online]. Available:

<http://dx.doi.org/10.3390/ijerph15030492>

Cited on page 24.

- [73] I. Buendia-Roldan, A. Palma-Lopez, D. Chan-Padilla, I. Herrera, M. Maldonado, R. Fernández, D. Martínez-Briseño, M. Mejia, and M. Selman, “Risk factors associated with the detection of pulmonary emphysema in older asymptomatic respiratory subjects,” *BMC Pulmonary Medicine*, vol. 20, no. 1, pp. 1–6, 2020. [Online]. Available: <http://dx.doi.org/10.1186/s12890-020-01204-9>
Cited on page 24.
- [74] D. Adeloye, P. Song, Y. Zhu, H. Campbell, A. Sheikh, I. Rudan, and N. R. G. R. H. Unit, “Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis,” *The Lancet Respiratory Medicine*, 2022. [Online]. Available: [http://dx.doi.org/10.1016/S2213-2600\(21\)00511-7](http://dx.doi.org/10.1016/S2213-2600(21)00511-7)
Cited on page 24.
- [75] R. E. Jordan, K. K. Cheng, M. R. Miller, and P. Adab, “Passive smoking and chronic obstructive pulmonary disease: cross-sectional analysis of data from the Health Survey for England,” *BMJ open*, vol. 1, no. 2, p. e000153, 2011. [Online]. Available: <http://dx.doi.org/10.1136/bmjopen-2011-000153>
Cited on page 25.
- [76] N. Zhong, C. Wang, W. Yao, P. Chen, J. Kang, S. Huang, B. Chen, C. Wang, D. Ni, Y. Zhou *et al.*, “Prevalence of chronic obstructive pulmonary disease in China: a large, population-based survey,” *American journal of respiratory and critical care medicine*, vol. 176, no. 8, pp. 753–760, 2007. [Online]. Available: <http://dx.doi.org/10.1164/rccm.200612-1749OC>
Cited on page 25.
- [77] S. H. Lee, E. D. Hwang, J. E. Lim, S. Moon, Y. Kang, J. Y. Jung, M. S. Park, S. K. Kim, J. Chang, Y. S. Kim *et al.*, “The risk factors and characteristics of COPD among nonsmokers in Korea: an analysis of KNHANES IV and V,” *Lung*, vol. 194, no. 3, pp. 353–361, 2016. [Online]. Available: <http://dx.doi.org/10.1007/s00408-016-9871-6>
Cited on page 25.
- [78] F. Xu, X. Yin, M. Zhang, H. Shen, L. Lu, and Y. Xu, “Prevalence of physician-diagnosed COPD and its association with smoking among urban and rural residents in regional mainland China,” *Chest*, vol. 128, no. 4, pp. 2818–2823, 2005. [Online]. Available: <http://dx.doi.org/10.1378/chest.128.4.2818>
Cited on page 25.

- [79] QIBA Lung Density Biomarker Committee. Computed Tomography: Lung Densitometry, Quantitative Imaging Biomarkers Alliance. Profile Stage: consensus. September 04, 2020. Available from: http://web.archive.org/web/20201216134027/https://qibawiki.rsna.org/images/a/a8/QIBA_CT_Lung_Density_Profile_090420-clean.pdf.
Cited on pages 25, 107, and 132.
- [80] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman, and F. Bray, “Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012,” *International journal of cancer*, vol. 136, no. 5, pp. E359–E386, 2015. [Online]. Available: <http://dx.doi.org/10.1002/ijc.29210>
Cited on page 31.
- [81] M. Gomes, A. L. Teixeira, A. Coelho, A. Araújo, and R. Medeiros, “The role of inflammation in lung cancer,” *Inflammation and cancer*, pp. 1–23, 2014. [Online]. Available: http://dx.doi.org/10.1007/978-3-0348-0837-8_1
Cited on page 31.
- [82] A. Sharafkhaneh, N. A. Hanania, and V. Kim, “Pathogenesis of emphysema: from the bench to the bedside,” *Proceedings of the American Thoracic Society*, vol. 5, no. 4, pp. 475–477, 2008. [Online]. Available: <http://dx.doi.org/10.1513/pats.200708-126ET>
Cited on page 31.
- [83] A. Biswas, H. J. Mehta, and E. E. Folch, “Chronic obstructive pulmonary disease and lung cancer: inter-relationships,” *Current opinion in pulmonary medicine*, vol. 24, no. 2, pp. 152–160, 2018. [Online]. Available: <http://dx.doi.org/10.1097/MCP.0000000000000451>
Cited on page 31.
- [84] C. Bergin, N. Müller, D. M. Nichols, G. Lillington, J. C. Hogg, B. Mullen, M. R. Grymaloski, S. Osborne, and P. D. Paré, “The diagnosis of emphysema: a computed tomographic-pathologic correlation,” *American Review of Respiratory Disease*, vol. 133, no. 4, pp. 541–546, 1986. [Online]. Available: <http://dx.doi.org/10.1164/arrd.1986.133.4.541>
Cited on page 31.
- [85] D. P. Naidich, F. P. Stitik, N. F. Khouri, P. B. Terry, and S. S. Siegelman, “Computed tomography of the bronchi. 2. Pathology.” *Journal of computer assisted tomography*, vol. 4, no. 6, pp. 754–762, 1980. [Online]. Available: <http://dx.doi.org/10.1097/00004728-198012000-00004>
Cited on page 31.
- [86] A. Madani, C. Keyzer, and P.-A. Gevenois, “Quantitative computed tomography assessment of lung structure and function in pulmonary

emphysema,” *European Respiratory Journal*, vol. 18, no. 4, pp. 720–730, 2001. [Online]. Available: <http://dx.doi.org/10.1183/09031936.01.00255701>

Cited on page 31.

- [87] F. Maldonado, B. J. Bartholmai, S. J. Swensen, D. E. Midthun, P. A. Decker, and J. R. Jett, “Are airflow obstruction and radiographic evidence of emphysema risk factors for lung cancer?: a nested case-control study using quantitative emphysema analysis,” *Chest*, vol. 138, no. 6, pp. 1295–1302, 2010. [Online]. Available: <http://dx.doi.org/10.1378/chest.09-2567>

Cited on pages 31, 34, 37, 45, 46, 197, and 208.

- [88] K. Kishi, J. Gurney, D. Schroeder, P. Scanlon, S. Swensen, and J. Jett, “The correlation of emphysema or airway obstruction with the risk of lung cancer: a matched case-controlled study,” *European Respiratory Journal*, vol. 19, no. 6, pp. 1093–1098, 2002. [Online]. Available: <http://dx.doi.org/10.1183/09031936.02.00264202>

Cited on pages 31, 34, 37, 197, and 208.

- [89] D. S. Gierada, P. Guniganti, B. J. Newman, M. T. Dransfield, P. A. Kvale, D. A. Lynch, and T. K. Pilgram, “Quantitative CT assessment of emphysema and airways in relation to lung cancer risk,” *Radiology*, vol. 261, no. 3, p. 950, 2011. [Online]. Available: <http://dx.doi.org/10.1148/radiol.11110542>

Cited on pages 31, 37, 198, and 209.

- [90] D. O. Wilson, J. L. Weissfeld, A. Balkan, J. G. Schragin, C. R. Fuhrman, S. N. Fisher, J. Wilson, J. K. Leader, J. M. Siegfried, S. D. Shapiro *et al.*, “Association of radiographic emphysema and airflow obstruction with lung cancer,” *American journal of respiratory and critical care medicine*, vol. 178, no. 7, pp. 738–744, 2008. [Online]. Available: <http://dx.doi.org/10.1164/rccm.200803-435OC>

Cited on pages 31, 36, 46, 192, and 203.

- [91] J. P. de Torres, G. Bastarrika, J. P. Wisnivesky, A. B. Alcaide, A. Campo, L. M. Seijo, J. C. Pueyo, A. Villanueva, M. D. Lozano, U. Montes *et al.*, “Assessing the relationship between lung cancer risk and emphysema detected on low-dose CT of the chest,” *Chest*, vol. 132, no. 6, pp. 1932–1938, 2007. [Online]. Available: <http://dx.doi.org/10.1378/chest.07-1490>

Cited on pages 31, 36, 192, and 203.

- [92] D. O. Wilson, J. K. Leader, C. R. Fuhrman, J. J. Reilly, F. C. Sciruba, and J. L. Weissfeld, “Quantitative computed tomography analysis, airflow obstruction, and lung cancer in the pittsburgh lung screening study,” *Journal of Thoracic Oncology*, vol. 6, no. 7, pp. 1200–1205, 2011.

- [Online]. Available: <http://dx.doi.org/10.1097/JTO.0b013e318219aa93>
Cited on pages 31 and 45.
- [93] B. M. Smith, L. Pinto, N. Ezer, N. Sverzellati, S. Muro, and K. Schwartzman, “Emphysema detected on computed tomography and risk of lung cancer: a systematic review and meta-analysis,” *Lung cancer*, vol. 77, no. 1, pp. 58–63, 2012. [Online]. Available: <http://dx.doi.org/10.1016/j.lungcan.2012.02.019>
Cited on pages 31 and 45.
- [94] C. Mouronte-Roibas, V. Leiro-Fernandez, A. Fernández-Villar, M. Botana-Rial, C. Ramos-Hernández, and A. Ruano-Ravina, “COPD, emphysema and the onset of lung cancer. A systematic review,” *Cancer letters*, vol. 382, no. 2, pp. 240–244, 2016. [Online]. Available: <http://dx.doi.org/10.1016/j.canlet.2016.09.002>
Cited on page 31.
- [95] S. Chubachi, S. Takahashi, A. Tsutsumi, N. Kameyama, M. Sasaki, K. Naoki, K. Soejima, H. Nakamura, K. Asano, and T. Betsuyaku, “Radiologic features of precancerous areas of the lungs in chronic obstructive pulmonary disease,” *International journal of chronic obstructive pulmonary disease*, vol. 12, p. 1613, 2017. [Online]. Available: <http://dx.doi.org/10.2147/COPD.S132709>
Cited on pages 31, 37, 45, 199, and 209.
- [96] A. A. Gagnat, M. Gjerdevik, F. Gallefoss, H. O. Coxson, A. Gulsvik, and P. Bakke, “Incidence of non-pulmonary cancer and lung cancer by amount of emphysema and airway wall thickness: a community-based cohort,” *European Respiratory Journal*, vol. 49, no. 5, 2017. [Online]. Available: <http://dx.doi.org/10.1183/13993003.01162-2016>
Cited on pages 31, 34, 37, 199, and 209.
- [97] C. Mouronte-Roibás, A. Fernández-Villar, A. Ruano-Raviña, C. Ramos-Hernández, A. Tilve-Gómez, P. Rodríguez-Fernández, A. C. C. Díaz, M. G. Vázquez-Noguerol, S. Fernández-García, and V. Leiro-Fernández, “Influence of the type of emphysema in the relationship between COPD and lung cancer,” *International Journal of Chronic Obstructive Pulmonary Disease*, vol. 13, p. 3563, 2018. [Online]. Available: <http://dx.doi.org/10.2147/COPD.S178109>
Cited on pages 31, 37, 46, 200, and 210.
- [98] J. González, C. I. Henschke, D. F. Yankelevitz, L. M. Seijo, A. P. Reeves, R. Yip, Y. Xie, M. Chung, P. Sánchez-Salcedo, A. B. Alcaide *et al.*, “Emphysema phenotypes and lung cancer risk,” *PLoS one*, vol. 14, no. 7, p. e0219187, 2019. [Online]. Available:

<http://dx.doi.org/10.1371/journal.pone.0219187>

Cited on pages 31, 36, 45, 46, 195, and 206.

- [99] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and t. PRISMA Group*, “Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement,” *Annals of internal medicine*, vol. 151, no. 4, pp. 264–269, 2009. [Online]. Available: <http://dx.doi.org/10.7326/0003-4819-151-4-200908180-00135>
Cited on page 31.
- [100] National Emphysema Treatment Trial Research Group, “Patients at high risk of death after lung-volume–reduction surgery,” *New England Journal of Medicine*, vol. 345, no. 15, pp. 1075–1083, 2001. [Online]. Available: <http://dx.doi.org/10.1056/NEJMoa11798>
Cited on page 33.
- [101] G. A. Wells, B. Shea, D. O’Connell, J. Peterson, V. Welch, M. Losos, P. Tugwell *et al.*, “The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses,” 2000. [Online]. Available: https://web.archive.org/web/20210716121605id_/http://www3.med.unipmn.it/dispense_ebm/2009-2010/Corso%20Perfezionamento%20EBM_Faggiano/NOS_oxford.pdf
Cited on page 33.
- [102] K. Boehm, F. Borrelli, and E. Ernst, “Green tea (*Camellia sinensis*) for the prevention of cancer,” *Cochrane Database of Systematic Reviews*, no. 3, p. CD005004, 2009. [Online]. Available: <http://dx.doi.org/10.1002/14651858.CD005004.pub2>
Cited on page 33.
- [103] J. Higgins and S. Green, *Cochrane handbook for systematic reviews of interventions*. Chichester, UK: Wiley-Blackwell, 2008. [Online]. Available: <http://dx.doi.org/10.1002/9780470712184>
Cited on pages 34 and 45.
- [104] J. S. Kuiper, M. Zuidersma, S. U. Zuidema, J. G. Burgerhof, R. P. Stolk, R. C. Oude Voshaar, and N. Smidt, “Social relationships and cognitive decline: a systematic review and meta-analysis of longitudinal cohort studies,” *International journal of epidemiology*, vol. 45, no. 4, pp. 1169–1206, 2016. [Online]. Available: <http://dx.doi.org/10.1093/ije/dyw089>
Cited on page 34.
- [105] J. P. Higgins, S. G. Thompson, J. J. Deeks, and D. G. Altman, “Measuring inconsistency in meta-analyses,” *Bmj*, vol. 327, no. 7414, pp. 557–560, 2003. [Online]. Available: <http://dx.doi.org/10.1136/bmj.327.7414.557>
Cited on page 34.

- [106] J. J. Deeks, J. P. Higgins, D. G. Altman, and C. S. M. Group, “Analysing data and undertaking meta-analyses,” *Cochrane handbook for systematic reviews of interventions*, pp. 241–284, 2019. [Online]. Available: <http://dx.doi.org/10.1002/9781119536604.ch10>
Cited on page 34.
- [107] L. L. Carr, S. Jacobson, D. A. Lynch, M. G. Foreman, E. L. Flanagan, C. P. Hersh, F. C. Sciurba, D. O. Wilson, J. C. Sieren, P. Mulhall *et al.*, “Features of COPD as predictors of lung cancer,” *Chest*, vol. 153, no. 6, pp. 1326–1335, 2018. [Online]. Available: <http://dx.doi.org/10.1016/j.chest.2018.01.049>
Cited on pages 34, 37, 45, 46, 196, and 212.
- [108] A. G. Schwartz, C. M. Lusk, A. S. Wenzlaff, D. Wata, S. Pandolfi, L. Mantha, M. L. Cote, A. O. Soubani, G. Walworth, A. Wozniak *et al.*, “Risk of Lung Cancer Associated with COPD Phenotype Based on Quantitative Image Analysis Lung Cancer and Image-Based COPD Phenotypes,” *Cancer Epidemiology, Biomarkers & Prevention*, vol. 25, no. 9, pp. 1341–1347, 2016. [Online]. Available: <http://dx.doi.org/10.1158/1055-9965.EPI-16-0176>
Cited on pages 34, 37, 198, and 211.
- [109] W. W. Labaki, M. Xia, S. Murray, C. R. Hatt, A. Al-Abcha, M. C. Ferrera, C. A. Meldrum, L. A. Keith, C. J. Galbán, D. A. Arenberg *et al.*, “Quantitative emphysema on low-dose CT imaging of the chest and risk of lung cancer and airflow obstruction: an analysis of the National Lung Screening Trial,” *Chest*, vol. 159, no. 5, pp. 1812–1820, 2021. [Online]. Available: <http://dx.doi.org/10.1016/j.chest.2020.12.004>
Cited on pages 34, 37, 201, and 211.
- [110] M. Nishio, T. Kubo, and K. Togashi, “Estimation of lung cancer risk using homology-based emphysema quantification in patients with lung nodules,” *PloS one*, vol. 14, no. 1, p. e0210720, 2019. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0210720>
Cited on pages 34, 37, 200, and 210.
- [111] Y. Li, S. J. Swensen, L. G. Karabekmez, R. S. Marks, S. M. Stoddard, R. Jiang, J. B. Worra, F. Zhang, D. E. Midthun, M. De Andrade *et al.*, “Effect of emphysema on lung cancer risk in smokers: a computed tomography-based assessment,” *Cancer prevention research*, vol. 4, no. 1, pp. 43–50, 2011. [Online]. Available: <http://dx.doi.org/10.1158/1940-6207.CAPR-10-0151>
Cited on pages 36, 193, and 204.
- [112] P. Maisonneuve, V. Bagnardi, M. Bellomi, L. Spaggiari, G. Pelosi, C. Rampinelli, R. Bertolotti, N. Rotmensz, J. K. Field, A. DeCensi

- et al.*, “Lung cancer risk prediction to select smokers for screening CT—a model based on the Italian COSMOS trial,” *Cancer Prevention Research*, vol. 4, no. 11, pp. 1778–1789, 2011. [Online]. Available: <http://dx.doi.org/10.1158/1940-6207.CAPR-11-0026>
Cited on pages 36, 193, and 204.
- [113] C. I. Henschke, R. Yip, P. Boffetta, S. Markowitz, A. Miller, T. Hanaoka, N. Wu, J. J. Zulueta, D. F. Yankelevitz, I.-E. Investigators *et al.*, “CT screening for lung cancer: Importance of emphysema for never smokers and smokers,” *Lung Cancer*, vol. 88, no. 1, pp. 42–47, 2015. [Online]. Available: <http://dx.doi.org/10.1016/j.lungcan.2015.01.014>
Cited on pages 36, 193, and 204.
- [114] P. Sanchez-Salcedo, J. Berto, J. P. de Torres, A. Campo, A. B. Alcaide, G. Bastarrika, J. C. Pueyo, A. Villanueva, J. I. Echeveste, M. D. Lozano *et al.*, “Lung cancer screening: fourteen year experience of the Pamplona early detection program (P-IELCAP),” *Archivos de Bronconeumología (English Edition)*, vol. 51, no. 4, pp. 169–176, 2015. [Online]. Available: <http://dx.doi.org/10.1016/j.arbr.2015.02.015>
Cited on pages 36, 194, and 205.
- [115] J. P. de Torres, D. O. Wilson, P. Sanchez-Salcedo, J. L. Weissfeld, J. Berto, A. Campo, A. B. Alcaide, M. García-Granero, B. R. Celli, and J. J. Zulueta, “Lung cancer in patients with chronic obstructive pulmonary disease. Development and validation of the COPD Lung Cancer Screening Score,” *American journal of respiratory and critical care medicine*, vol. 191, no. 3, pp. 285–291, 2015. [Online]. Available: <http://dx.doi.org/10.1164/rccm.201407-1210OC>
Cited on pages 36, 194, and 205.
- [116] Y. Liu, H. Wang, Q. Li, M. J. McGettigan, Y. Balagurunathan, A. L. Garcia, Z. J. Thompson, J. J. Heine, Z. Ye, R. J. Gillies *et al.*, “Radiologic features of small pulmonary nodules and lung cancer risk in the National Lung Screening Trial: a nested case-control study,” *Radiology*, vol. 286, no. 1, p. 298, 2018. [Online]. Available: <http://dx.doi.org/10.1148/radiol.2017161458>
Cited on pages 36, 195, and 206.
- [117] P. C. Yong, K. Sigel, J. P. de Torres, G. Mhango, M. Kale, C. Y. Kong, J. J. Zulueta, D. Wilson, S.-A. W. Brown, C. Slatore *et al.*, “The effect of radiographic emphysema in assessing lung cancer risk,” *Thorax*, vol. 74, no. 9, pp. 858–864, 2019. [Online]. Available: <http://dx.doi.org/10.1136/thoraxjnl-2018-212457>
Cited on pages 36, 197, and 207.
- [118] F. Maldonado, J. R. Jett, and P. A. Decker, “The Relationship Between Emphysema on CT Scan and Lung Cancer: Response,”

- Chest*, vol. 139, no. 5, pp. 1259–1260, 2011. [Online]. Available: <http://dx.doi.org/10.1378/chest.11-0011>
Cited on pages 37, 197, and 208.
- [119] G. R. Husebø, R. Nielsen, J. Hardie, P. S. Bakke, L. Lerner, C. D’Alessandro-Gabazza, J. Gyuris, E. Gabazza, P. Aukrust, and T. Eagan, “Risk factors for lung cancer in COPD—results from the Bergen COPD cohort study,” *Respiratory medicine*, vol. 152, pp. 81–88, 2019. [Online]. Available: <http://dx.doi.org/10.1016/j.rmed.2019.04.019>
Cited on pages 37, 201, and 210.
- [120] D. R. Brenner, J. R. McLaughlin, and R. J. Hung, “Previous lung diseases and lung cancer risk: a systematic review and meta-analysis,” *PLoS one*, vol. 6, no. 3, p. e17479, 2011. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0017479>
Cited on page 45.
- [121] X. Zhang, N. Jiang, L. Wang, H. Liu, and R. He, “Chronic obstructive pulmonary disease and risk of lung cancer: a meta-analysis of prospective cohort studies,” *Oncotarget*, vol. 8, no. 44, p. 78044, 2017. [Online]. Available: <http://dx.doi.org/10.18632/oncotarget.20351>
Cited on page 45.
- [122] B. C. Stoel and J. Stolk, “Optimization and standardization of lung densitometry in the assessment of pulmonary emphysema,” *Investigative radiology*, vol. 39, no. 11, pp. 681–688, 2004. [Online]. Available: <http://dx.doi.org/10.1097/00004424-200411000-00006>
Cited on page 45.
- [123] A. Hagiwara, S. Fujita, Y. Ohno, and S. Aoki, “Variability and standardization of quantitative imaging: monoparametric to multiparametric quantification, radiomics, and artificial intelligence,” *Investigative radiology*, vol. 55, no. 9, p. 601, 2020. [Online]. Available: <http://dx.doi.org/10.1097/RLI.0000000000000666>
Cited on page 45.
- [124] M. L. Wood, “Variability and standardization of quantitative imaging,” pp. 617–618, 2020. [Online]. Available: <http://dx.doi.org/10.1097/RLI.0000000000000667>
Cited on page 45.
- [125] S. Wasswa-Kintu, W. Gan, S. Man, P. Pare, and D. Sin, “Relationship between reduced forced expiratory volume in one second and the risk of lung cancer: a systematic review and meta-analysis,” *Thorax*, vol. 60, no. 7, pp. 570–575, 2005. [Online]. Available: <http://dx.doi.org/10.1136/thx.2004.037135>
Cited on page 46.

- [126] M. Hosseini, A. Almasi-Hashiani, M. Sepidarkish, and S. Maroufizadeh, "Global prevalence of asthma-COPD overlap (ACO) in the general population: a systematic review and meta-analysis," *Respiratory research*, vol. 20, no. 1, pp. 1–10, 2019. [Online]. Available: <http://dx.doi.org/10.1186/s12931-019-1198-4>
Cited on page 51.
- [127] R. Brown, D. Leith, and P. Enright, "Multiple breath helium dilution measurement of lung volumes in adults," *European Respiratory Journal*, vol. 11, no. 1, pp. 246–255, 1998. [Online]. Available: <http://dx.doi.org/10.1183/09031936.98.11010246>
Cited on page 51.
- [128] J. T. Bakker, K. Klooster, R. Vlieninghart, and D.-J. Slebos, "Measuring pulmonary function in COPD using quantitative chest computed tomography analysis," *European Respiratory Review*, vol. 30, no. 161, 2021. [Online]. Available: <http://dx.doi.org/10.1183/16000617.0031-2021>
Cited on pages 51, 59, and 60.
- [129] C. Criée, S. Sorichter, H. Smith, P. Kardos, R. Merget, D. Heise, D. Berdel, D. Köhler, H. Magnussen, W. Marek *et al.*, "Body plethysmography—its principles and clinical use," *Respiratory medicine*, vol. 105, no. 7, pp. 959–971, 2011. [Online]. Available: <http://dx.doi.org/10.1016/j.rmed.2011.02.006>
Cited on page 51.
- [130] M. Xu, S. Qi, Y. Yue, Y. Teng, L. Xu, Y. Yao, and W. Qian, "Segmentation of lung parenchyma in CT images using CNN trained with the clustering algorithm generated dataset," *Biomedical engineering online*, vol. 18, no. 1, pp. 1–21, 2019. [Online]. Available: <http://dx.doi.org/10.1186/s12938-018-0619-9>
Cited on page 51.
- [131] M. Shen, E. D. Tenda, W. McNulty, J. Garner, H. Robbie, V. Luzzi, A. M. Aboelhassan, W. H. Van Geffen, S. V. Kemp, C. Ridge *et al.*, "Quantitative evaluation of lobar pulmonary function of emphysema patients with endobronchial coils," *Respiration*, vol. 98, no. 1, pp. 70–81, 2019. [Online]. Available: <http://dx.doi.org/10.1159/000499622>
Cited on page 51.
- [132] A. J. Matsumoto, B. J. Bartholmai, and M. E. Wylam, "Comparison of total lung capacity determined by plethysmography with computed tomographic segmentation using CALIPER," *Journal of thoracic imaging*, vol. 32, no. 2, pp. 101–106, 2017. [Online]. Available: <http://dx.doi.org/10.1097/RTI.0000000000000249>
Cited on pages 51 and 60.

- [133] Y. Yamada, M. Yamada, Y. Yokoyama, A. Tanabe, S. Matsuoka, Y. Nijjima, K. Narita, T. Nakahara, M. Murata, K. Fukunaga *et al.*, “Differences in lung and lobe volumes between supine and standing positions scanned with conventional and newly developed 320-detector-row upright CT: intra-individual comparison,” *Respiration*, vol. 99, no. 7, pp. 598–605, 2020. [Online]. Available: <http://dx.doi.org/10.1159/000507265>
Cited on pages 51, 55, and 60.
- [134] J. L. Garfield, N. Marchetti, J. P. Gaughan, R. M. Steiner, and G. J. Criner, “Total lung capacity by plethysmography and high-resolution computed tomography in COPD,” *International journal of chronic obstructive pulmonary disease*, vol. 7, p. 119, 2012. [Online]. Available: <http://dx.doi.org/10.2147/COPD.S26419>
Cited on pages 51 and 60.
- [135] H. Coxson, P. N. Fauerbach, C. Storness-Bliss, N. Müller, S. Cogswell, D. Dillard, C. Finger, and S. Springmeyer, “Computed tomography assessment of lung volume changes after bronchial valve treatment,” *European Respiratory Journal*, vol. 32, no. 6, pp. 1443–1450, 2008. [Online]. Available: <http://dx.doi.org/10.1183/09031936.00056008>
Cited on pages 51 and 60.
- [136] J. Zaporozhan, S. Ley, R. Eberhardt, O. Weinheimer, S. Iliyushenko, F. Herth, and H.-U. Kauczor, “Paired inspiratory/expiratory volumetric thin-slice CT scan for emphysema analysis: comparison of different quantitative evaluations and pulmonary function test,” *Chest*, vol. 128, no. 5, pp. 3212–3220, 2005. [Online]. Available: <http://dx.doi.org/10.1378/chest.128.5.3212>
Cited on pages 51 and 60.
- [137] D. S. Gierada, S. Hakimian, R. M. Slone, and R. D. Yusen, “MR analysis of lung volume and thoracic dimensions in patients with emphysema before and after lung volume reduction surgery,” *AJR. American journal of roentgenology*, vol. 170, no. 3, pp. 707–714, 1998. [Online]. Available: <http://dx.doi.org/10.2214/ajr.170.3.9490958>
Cited on pages 51 and 60.
- [138] A. M. Ganapathi, M. S. Mulvihill, B. R. Englum, P. J. Speicher, B. C. Gulack, A. A. Osho, B. A. Yerokun, L. R. Snyder, D. Davis, and M. G. Hartwig, “Transplant size mismatch in restrictive lung disease,” *Transplant International*, vol. 30, no. 4, pp. 378–387, 2017. [Online]. Available: <http://dx.doi.org/10.1111/tri.12913>
Cited on page 51.
- [139] P. Riddell, J. Ma, B. Dunne, M. Binnie, M. Cypel, L. Donahoe, M. de Perrot, A. Pierre, T. K. Waddell, J. Yeung *et al.*, “A

simplified strategy for donor-recipient size-matching in lung transplant for interstitial lung disease,” *The Journal of Heart and Lung Transplantation*, vol. 40, no. 11, pp. 1422–1430, 2021. [Online]. Available: <http://dx.doi.org/10.1016/j.healun.2021.06.013>

Cited on page 51.

- [140] R. V. Guillamet, “Comments and opinions: Predicted total lung capacity equations: A barrier to the definition of safe lung size differences in lung transplantation,” *The Journal of Heart and Lung Transplantation*, 2022. [Online]. Available: <http://dx.doi.org/10.1016/J.HEALUN.2022.03.012>

Cited on pages 51 and 62.

- [141] A. Sijtsma, J. Rienks, P. van der Harst, G. Navis, J. G. M. Rosmalen, and A. Dotinga, “Cohort Profile Update: Lifelines, a three-generation cohort study and biobank,” *International Journal of Epidemiology*, 12 2021. [Online]. Available: <http://dx.doi.org/10.1093/ije/dyab257>

Cited on pages 52 and 114.

- [142] P. H. Quanjer, G. Tammeling, J. Cotes, O. Pedersen, R. Peslin, and J. Yernault, “Lung volumes and forced ventilatory flows,” *European Respiratory Journal*, vol. 6, no. Suppl 16, pp. 5–40, 1993. [Online]. Available: <http://dx.doi.org/10.1183/09041950.005s1693>

Cited on pages 54 and 61.

- [143] P. Quanjer, S. Stanojevic, T. Cole, X. Baur, G. Hall, B. Culver, P. Enright, J. Hankinson, M. Ip, J. Zheng *et al.*, “Multi-ethnic reference values for spirometry for the 3-95-yr age range: The global lung function 2012 equations,” *European Respiratory Journal*, vol. 40, no. 6, pp. 1324–1343, 2012. [Online]. Available: <http://dx.doi.org/10.1183/09031936.00080312>

Cited on pages 54 and 61.

- [144] P. A. de Jong, Y. Nakano, M. Lequin, J. Mayo, R. Woods, P. Pare, and H. Tiddens, “Progressive damage on high resolution computed tomography despite stable lung function in cystic fibrosis,” *European Respiratory Journal*, vol. 23, no. 1, pp. 93–97, 2004. [Online]. Available: <http://dx.doi.org/10.1183/09031936.03.00006603>

Cited on page 59.

- [145] H.-K. Koo, D. M. Vasilescu, S. Booth, A. Hsieh, O. L. Katsamenis, N. Fishbane, W. M. Elliott, M. Kirby, P. Lackie, I. Sinclair *et al.*, “Small airways disease in mild and moderate chronic obstructive pulmonary disease: a cross-sectional study,” *The Lancet Respiratory Medicine*, vol. 6, no. 8, pp. 591–602, 2018. [Online]. Available: [http://dx.doi.org/10.1016/S2213-2600\(18\)30196-6](http://dx.doi.org/10.1016/S2213-2600(18)30196-6)

Cited on page 59.

- [146] M. S. Brown, H. J. Kim, F. Abtin, I. Da Costa, R. Pais, S. Ahmad, E. Angel, C. Ni, E. C. Kleerup, D. W. Gjertson *et al.*, “Reproducibility of lung and lobar volume measurements using computed tomography,” *Academic radiology*, vol. 17, no. 3, pp. 316–322, 2010. [Online]. Available: <http://dx.doi.org/10.1016/j.acra.2009.10.005>
Cited on page 59.
- [147] J. K. Leader, R. M. Rogers, C. R. Fuhrman, F. C. Scirba, B. Zheng, P. F. Thompson, J. L. Weissfeld, S. K. Golla, and D. Gur, “Size and morphology of the trachea before and after lung volume reduction surgery,” *American Journal of Roentgenology*, vol. 183, no. 2, pp. 315–321, 2004. [Online]. Available: <http://dx.doi.org/10.2214/ajr.183.2.1830315>
Cited on page 60.
- [148] J. Pu, J. K. Leader, X. Meng, B. Whiting, D. Wilson, F. C. Scirba, J. J. Reilly, W. L. Bigbee, J. Siegfried, and D. Gur, “Three-dimensional airway tree architecture and pulmonary function,” *Academic radiology*, vol. 19, no. 11, pp. 1395–1401, 2012. [Online]. Available: <http://dx.doi.org/10.1016/j.acra.2012.06.007>
Cited on page 60.
- [149] R. L. Jones and M.-M. U. Nzekwu, “The effects of body mass index on lung volumes,” *Chest*, vol. 130, no. 3, pp. 827–833, 2006. [Online]. Available: <http://dx.doi.org/10.1378/chest.130.3.827>
Cited on page 61.
- [150] J. R. Bach and S.-W. Kang, “Disorders of ventilation: weakness, stiffness, and mobilization,” *Chest*, vol. 117, no. 2, pp. 301–303, 2000. [Online]. Available: <http://dx.doi.org/10.1378/chest.117.2.301>
Cited on page 61.
- [151] A. E. Moran, M. H. Forouzanfar, G. Roth, G. Mensah, M. Ezzati, C. J. Murray, and M. Naghavi, “Temporal trends in ischemic heart disease mortality in 21 world regions, 1980-2010: The Global Burden of Disease 2010 Study,” *Circulation*, vol. 129, no. 14, pp. 1483–1492, 2014. [Online]. Available: <http://dx.doi.org/10.1161/CIRCULATIONAHA.113.004042>
Cited on page 67.
- [152] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, “Global cancer statistics, 2012,” *CA: A Cancer Journal for Clinicians*, vol. 65, no. 2, pp. 87–108, 2015. [Online]. Available: <http://dx.doi.org/10.3322/caac.21262>
Cited on page 67.
- [153] P. G. Burney, J. Patel, R. Newson, C. Minelli, and M. Naghavi, “Global and regional trends in COPD mortality, 1990–2010,” *European Respiratory*

- Journal*, vol. 45, no. 5, pp. 1239–1247, 2015. [Online]. Available: <http://dx.doi.org/10.1183/09031936.00142414>
Cited on page 67.
- [154] I. S. Stone, N. C. Barnes, and S. E. Petersen, “Chronic obstructive pulmonary disease: a modifiable risk factor for cardiovascular disease?” *Heart*, vol. 98, no. 14, pp. 1055–1062, 2012. [Online]. Available: <http://dx.doi.org/10.1136/heartjnl-2012-301759>
Cited on page 67.
- [155] D. Aberle, A. Adams, C. Berg, W. Black, J. Clapp, R. Fagerstrom, I. Gareen, C. Gatsonis, P. Marcus, and J. Sicks, “Reduced lung-cancer mortality with low-dose computed tomographic screening,” *New England Journal of Medicine*, vol. 365, no. 5, pp. 395–409, 2011. [Online]. Available: <http://dx.doi.org/10.1056/NEJMoa1102873>
Cited on page 67.
- [156] M. Oudkerk, A. Devaraj, R. Vliegenthart, T. Henzler, H. Prosch, C. P. Heussel, G. Bastarrika, N. Sverzellati, M. Mascacchi, S. Delorme *et al.*, “European position statement on lung cancer screening,” *The Lancet Oncology*, vol. 18, no. 12, pp. e754–e766, 2017. [Online]. Available: [https://doi.org/10.1016/S1470-2045\(17\)30861-6](https://doi.org/10.1016/S1470-2045(17)30861-6)
Cited on pages 67, 81, and 134.
- [157] M. Messerli, T. Otilinger, R. Warschkow, S. Leschka, H. Alkadhi, S. Wildermuth, and R. W. Bauer, “Emphysema quantification and lung volumetry in chest X-ray equivalent ultralow dose CT—Intra-individual comparison with standard dose CT,” *European Journal of Radiology*, vol. 91, pp. 1–9, 2017. [Online]. Available: <http://dx.doi.org/10.1016/j.ejrad.2017.03.003>
Cited on pages 67, 80, 85, 86, 87, 95, 96, 101, and 132.
- [158] E. C. Oelsner, J. J. Carr, P. L. Enright, E. A. Hoffman, A. R. Folsom, S. M. Kawut, R. A. Kronmal, D. J. Lederer, J. A. Lima, G. S. Lovasi *et al.*, “Per cent emphysema is associated with respiratory and lung cancer mortality in the general population: a cohort study,” *Thorax*, vol. 71, no. 7, pp. 624–632, 2016. [Online]. Available: <http://dx.doi.org/10.1136/thoraxjnl-2015-207822>
Cited on pages 67, 85, 97, and 113.
- [159] H. Koyama, Y. Ohno, Y. Yamazaki, M. Nogami, K. Murase, Y. Onishi, K. Matsumoto, D. Takenaka, and K. Sugimura, “Quantitative and qualitative assessments of lung destruction and pulmonary functional loss from reduced-dose thin-section CT in pulmonary emphysema patients,” *Academic Radiology*, vol. 17, no. 2, pp. 163–168, 2010. [Online].

Available: <http://dx.doi.org/10.1016/j.acra.2009.08.009>

Cited on pages 67, 85, and 97.

- [160] A. Madani, J. Zanen, V. De Maertelaer, and P. A. Gevenois, “Pulmonary emphysema: objective quantification at multi-detector row CT—comparison with macroscopic and microscopic morphometry,” *Radiology*, vol. 238, no. 3, pp. 1036–1043, 2006. [Online]. Available: <http://dx.doi.org/10.1148/radiol.2382042196>
Cited on page 67.

- [161] J. D. Newell Jr, M. K. Fuld, T. Allmendinger, J. P. Sieren, K.-S. Chan, J. Guo, and E. A. Hoffman, “Very low-dose (0.15 mGy) chest CT protocols using the COPDGene 2 test object and a third-generation dual-source CT scanner with corresponding third-generation iterative reconstruction software,” *Investigative Radiology*, vol. 50, no. 1, p. 40, 2015. [Online]. Available: <http://dx.doi.org/10.1097/RLI.0000000000000093>
Cited on pages 67, 68, 71, 72, and 77.

- [162] QIBA Lung Density Biomarker Committee. Computed Tomography: Lung Densitometry, Quantitative Imaging Biomarkers Alliance. Profile Stage: internal BC review. QIBA, August 30, 2017. [Online]. Available: http://web.archive.org/web/20180122144040/http://qibawiki.rsna.org/images/7/70/QIBA_CT_Lung_Density_Profile_083017-Clean.docx
Cited on pages 67, 71, 72, 76, 77, 79, and 81.

- [163] U.S. Food and Drug Administration, *FDA clearance PixelShine*, 2016, retrieved 2017/02/12. [Online]. Available: http://web.archive.org/web/20170212231946/https://www.accessdata.fda.gov/cdrh_docs/pdf16/K161625.pdf
Cited on page 68.

- [164] S.-f. Tian, A.-l. Liu, J.-h. Liu, Y.-j. Liu, and J.-d. Pan, “Potential value of the PixelShine deep learning algorithm for increasing quality of 70 kVp+ ASiR-V reconstruction pelvic arterial phase CT images,” *Japanese Journal of Radiology*, vol. 37, no. 2, pp. 186–190, 2019. [Online]. Available: <http://dx.doi.org/10.1007/s11604-018-0798-0>
Cited on pages 69, 80, and 96.

- [165] *MATLAB version 9.5.0.944444 (R2018b)*, MathWorks, Natick, Massachusetts, 2018.
Cited on page 71.

- [166] *CTP698 and CCT162 Lung Phantom II Manual*, 2017. [Online]. Available: <http://web.archive.org/web/20171212092220/https://static1.squarespace.com/static/5367b059e4b05a1adcd295c2/t/5636b3d6e4b0507883462a67/1446425558230/CTP698+%26+CCT162+>

Lung+Phan+Manual+Nov14.pdf

Cited on page 71.

- [167] A. Rodriguez, F. N. Ranallo, P. F. Judy, and S. B. Fain, “The effects of iterative reconstruction and kernel selection on quantitative computed tomography measures of lung density,” *Medical Physics*, vol. 44, no. 6, pp. 2267–2280, 2017. [Online]. Available: <http://dx.doi.org/10.1002/mp.12255>
Cited on pages 79 and 80.
- [168] D. S. Gierada, A. J. Bierhals, C. K. Choong, S. T. Bartel, J. H. Ritter, N. A. Das, C. Hong, T. K. Pilgram, K. T. Bae, B. R. Whiting *et al.*, “Effects of CT section thickness and reconstruction kernel on emphysema quantification: relationship to the magnitude of the CT emphysema index,” *Academic Radiology*, vol. 17, no. 2, pp. 146–156, 2010. [Online]. Available: <http://dx.doi.org/10.1016/j.acra.2009.08.007>
Cited on page 79.
- [169] S. P. Martin, J. Gariani, A.-L. Hachulla, D. Botsikas, D. Adler, W. Karenovics, C. D. Becker, and X. Montet, “Impact of iterative reconstructions on objective and subjective emphysema assessment with computed tomography: a prospective study,” *European Radiology*, vol. 27, no. 7, pp. 2950–2956, 2017. [Online]. Available: <http://dx.doi.org/10.1007/s00330-016-4641-7>
Cited on page 80.
- [170] T. Yamashiro, T. Miyara, O. Honda, N. Tomiyama, Y. Ohno, S. Noma, and S. Murayama, “Iterative reconstruction for quantitative computed tomography analysis of emphysema: consistent results using different tube currents,” *International Journal of Chronic Obstructive Pulmonary Disease*, vol. 10, no. 1, p. 321, 2015. [Online]. Available: <http://dx.doi.org/10.2147/COPD.S74810>
Cited on pages 80 and 85.
- [171] O. M. Mets, M. J. Willeminck, F. P. de Kort, C. P. Mol, T. Leiner, M. Oudkerk, M. Prokop, and P. A. de Jong, “The effect of iterative reconstruction on computed tomography assessment of emphysema, air trapping and airway dimensions,” *European Radiology*, vol. 22, no. 10, pp. 2103–2109, 2012. [Online]. Available: <http://dx.doi.org/10.1007/s00330-012-2489-z>
Cited on pages 80 and 85.
- [172] M. Nishio, S. Matsumoto, Y. Ohno, N. Sugihara, H. Inokawa, T. Yoshikawa, and K. Sugimura, “Emphysema quantification by low-dose CT: potential impact of adaptive iterative dose reduction using 3D processing,” *American Journal of Roentgenology*, vol. 199, no. 3, pp. 595–601, 2012. [Online].

Available: <http://dx.doi.org/10.2214/AJR.11.8174>

Cited on pages 80, 85, and 132.

- [173] M. Nishio, H. Koyama, Y. Ohno, N. Negi, S. Seki, T. Yoshikawa, and K. Sugimura, “Emphysema quantification using ultralow-dose CT with iterative reconstruction and filtered back projection,” *American Journal of Roentgenology*, vol. 206, no. 6, pp. 1184–1192, 2016. [Online]. Available: <http://dx.doi.org/10.2214/AJR.15.15684>
Cited on pages 80, 86, 96, and 132.

- [174] N. M. Cross, J. DeBerry, D. Ortiz, J. Kemp, and J. Morey, “Diagnostic Quality of Machine Learning Algorithm for Optimization of Low-Dose Computed Tomography Data,” in *SIIM 2017 Scientific Session Posters & Demonstrations*, 2017, <https://web.archive.org/web/20171004113144/http://c.ymcdn.com/sites/siim.org/resource/resmgr/siim2017/abstracts/posters-Cross.pdf>.
Cited on pages 80 and 96.

- [175] C. de Margerie-Mellon, C. de Bazelaire, C. Montlahuc, J. Lambert, A. Martineau, P. Coulon, E. de Kerviler, and C. Beigelman, “Reducing radiation dose at chest CT: comparison among model-based type iterative reconstruction, hybrid iterative reconstruction, and filtered back projection,” *Academic Radiology*, vol. 23, no. 10, pp. 1246–1254, 2016. [Online]. Available: <http://dx.doi.org/10.1016/j.acra.2016.05.019>
Cited on page 80.

- [176] A. Madani, A. Van Muylem, and P. A. Gevenois, “Pulmonary emphysema: effect of lung volume on objective quantification at thin-section CT,” *Radiology*, vol. 257, no. 1, pp. 260–268, 2010. [Online]. Available: <http://dx.doi.org/10.1148/radiol.10091446>
Cited on pages 85 and 97.

- [177] H. Atta, G. S. Seifeldein, A. Rashad, and R. Elmorshidy, “Quantitative validation of the severity of emphysema by multi-detector CT,” *The Egyptian Journal of Radiology and Nuclear Medicine*, vol. 46, no. 2, pp. 355–361, 2015. [Online]. Available: <http://dx.doi.org/10.1016/j.ejrnm.2014.11.016>
Cited on pages 85 and 97.

- [178] A. Hata, M. Yanagawa, N. Kikuchi, O. Honda, and N. Tomiyama, “Pulmonary Emphysema Quantification on Ultra–Low-Dose Computed Tomography Using Model-Based Iterative Reconstruction With or Without Lung Setting,” *Journal of Computer Assisted Tomography*, vol. 42, no. 5, pp. 760–766, 2018. [Online]. Available: <http://dx.doi.org/10.1097/RCT.0000000000000755>
Cited on pages 85, 86, 96, and 132.

- [179] C. Kim, K. Y. Lee, C. Shin, E.-Y. Kang, Y.-W. Oh, M. Ha, C. S. Ko, and J. Cha, “Comparison of Filtered Back Projection, Hybrid Iterative Reconstruction, Model-Based Iterative Reconstruction, and Virtual Monoenergetic Reconstruction Images at Both Low-and Standard-Dose Settings in Measurement of Emphysema Volume and Airway Wall Thickness: A CT Phantom Study,” *Korean Journal of Radiology*, vol. 19, no. 4, pp. 809–817, 2018. [Online]. Available: <http://dx.doi.org/10.3348/kjr.2018.19.4.809>
Cited on page 85.
- [180] T. Pan, A. Hasegawa, D. Luo, C. C. Wu, and R. Vikram, “Technical Note: Impact on central frequency and noise magnitude ratios by advanced CT image reconstruction techniques,” *Medical Physics*, vol. 47, no. 2, pp. 480–487, 2020. [Online]. Available: <http://dx.doi.org/10.1002/mp.13937>
Cited on pages 85 and 107.
- [181] A. Steuwe, M. Weber, O. T. Bethge, C. Rademacher, M. Boschheidgen, L. M. Sawicki, G. Antoch, and J. Aissa, “Influence of a novel deep-learning based reconstruction software on the objective and subjective image quality in low-dose abdominal computed tomography,” *The British Journal of Radiology*, vol. 93, no. 1117, p. 20200677, 2020. [Online]. Available: <http://dx.doi.org/10.1259/bjr.20200677>
Cited on page 85.
- [182] J. Hsieh, E. Liu, B. Nett, J. Tang, J. Thibault, and S. Sahney, “A new era of image reconstruction: TrueFidelity™,” GE Healthcare, Technical white paper on deep learning image reconstruction, 2019. [Online]. Available: <http://web.archive.org/web/20220615235208/https://www.gehealthcare.com/-/jssmedia/040dd213fa89463287155151fdb01922.pdf>
Cited on page 85.
- [183] R. Singh, S. R. Digumarthy, V. V. Muse, A. R. Kambadakone, M. A. Blake, A. Tabari, Y. Hoi, N. Akino, E. Angel, R. Madan *et al.*, “Image Quality and Lesion Detection on Deep Learning Reconstruction and Iterative Reconstruction of Submillisievert Chest and Abdominal CT,” *American Journal of Roentgenology*, vol. 214, no. 3, pp. 566–573, 2020. [Online]. Available: <http://dx.doi.org/10.2214/AJR.19.21809>
Cited on page 85.
- [184] K. Park and J. Cho, “Quantification of emphysema using low-dose chest CT: Effect of iterative reconstruction,” *Respiratory Medicine*, vol. 132, p. 274, 2017. [Online]. Available: <http://dx.doi.org/10.1016/j.rmed.2017.07.036>
Cited on pages 86 and 96.
- [185] S. P. Martin, J. Gariani, G. Feutry, D. Adler, W. Karenovics, C. D. Becker, and X. Montet, “Emphysema quantification using hybrid versus

model-based generations of iterative reconstruction: SAFIRE versus ADMIRE,” *Medicine*, vol. 98, no. 7, p. e14450, 2019. [Online]. Available: <http://dx.doi.org/10.1097/MD.00000000000014450>

Cited on pages 86, 96, and 101.

- [186] C. F. Vogelmeier, G. J. Criner, F. J. Martinez, A. Anzueto, P. J. Barnes, J. Bourbeau, B. R. Celli, R. Chen, M. Decramer, L. M. Fabbri, P. Frith, D. M. Halpin, M. V. López Varela, M. Nishimura, N. Roche, R. Rodriguez-Roisin, D. D. Sin, D. Singh, R. Stockley, J. Vestbo, J. A. Wedzicha, and A. Agusti, “Global Strategy for the Diagnosis, Management and Prevention of Chronic Obstructive Lung Disease 2017 Report,” *Respirology*, vol. 22, no. 3, pp. 575–601, 2017. [Online]. Available: <http://dx.doi.org/10.1111/resp.13012>

Cited on page 86.

- [187] E. A. Regan, J. E. Hokanson, J. R. Murphy, B. Make, D. A. Lynch, T. H. Beaty, D. Curran-Everett, E. K. Silverman, and J. D. Crapo, “Genetic epidemiology of COPD (COPDGene) study design,” *COPD: Journal of Chronic Obstructive Pulmonary Disease*, vol. 7, no. 1, pp. 32–43, 2011. [Online]. Available: <http://dx.doi.org/10.3109/15412550903499522>

Cited on page 86.

- [188] J. P. Sieren, J. D. Newell Jr, R. G. Barr, E. R. Bleecker, N. Burnette, E. E. Carretta, D. Couper, J. Goldin, J. Guo, M. K. Han *et al.*, “SPIROMICS protocol for multicenter quantitative computed tomography to phenotype the lungs,” *American Journal of Respiratory and Critical Care Medicine*, vol. 194, no. 7, pp. 794–806, 2016. [Online]. Available: <http://dx.doi.org/10.1164/rccm.201506-1208PP>

Cited on page 86.

- [189] Z. Wang, S. Gu, J. K. Leader, S. Kundu, J. R. Tedrow, F. C. Sciruba, D. Gur, J. M. Siegfried, and J. Pu, “Optimal threshold in CT quantification of emphysema,” *European Radiology*, vol. 23, no. 4, pp. 975–984, 2013. [Online]. Available: <http://dx.doi.org/10.1007/s00330-012-2683-z>

Cited on page 87.

- [190] K. S. Iyer, R. W. Grout, G. K. Zamba, and E. A. Hoffman, “Repeatability and sample size assessment associated with computed tomography-based lung density metrics,” *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation*, vol. 1, no. 1, p. 97, 2014. [Online]. Available: <http://dx.doi.org/10.15326/jcopdf.1.1.2014.0111>

Cited on page 95.

- [191] H. A. Gietema, A. M. Schilham, B. van Ginneken, R. J. van Klaveren, J. W. J. Lammers, and M. Prokop, “Monitoring of smoking-induced emphysema with CT in a lung cancer screening setting: detection of real

- increase in extent of emphysema,” *Radiology*, vol. 244, no. 3, pp. 890–897, 2007. [Online]. Available: <http://dx.doi.org/10.1148/radiol.2443061330>
Cited on pages 95 and 96.
- [192] B. M. Keller, A. P. Reeves, C. I. Henschke, R. G. Barr, and D. F. Yankelevitz, “Variation of quantitative emphysema measurements from CT scans,” in *Medical Imaging 2008: Computer-Aided Diagnosis*, vol. 6915. International Society for Optics and Photonics, 2008, p. 69152I. [Online]. Available: <http://dx.doi.org/10.1117/12.770844>
Cited on page 95.
- [193] B. M. Keller, A. P. Reeves, C. I. Henschke, and D. F. Yankelevitz, “Multivariate compensation of quantitative pulmonary emphysema metric variation from low-dose, whole-lung CT scans,” *American Journal of Roentgenology*, vol. 197, no. 3, pp. W495–W502, 2011. [Online]. Available: <http://dx.doi.org/10.2214/AJR.11.6444>
Cited on page 95.
- [194] B. Hochegger, K. L. Irion, E. Marchiori, and J. S. Moreira, “Reconstruction algorithms influence the follow-up variability in the longitudinal CT emphysema index measurements,” *Korean Journal of Radiology*, vol. 12, no. 2, pp. 169–175, 2011. [Online]. Available: <http://dx.doi.org/10.3348/kjr.2011.12.2.169>
Cited on pages 95 and 96.
- [195] A. M. den Harder, E. de Boer, S. J. Lagerweij, M. F. Boomsma, A. M. Schilham, M. J. Willeminck, J. Milles, T. Leiner, R. P. Budde, and P. A. de Jong, “Emphysema quantification using chest CT: influence of radiation dose reduction and reconstruction technique,” *European Radiology Experimental*, vol. 2, no. 1, p. 30, 2018. [Online]. Available: <http://dx.doi.org/10.1186/s41747-018-0064-3>
Cited on pages 95, 96, 97, and 101.
- [196] E. de Boer, I. M. Nijholt, S. Jansen, M. Edens, S. Walen, J. van den Berg, and M. Boomsma, “Optimization of pulmonary emphysema quantification on CT scans of COPD patients using hybrid iterative and post processing techniques: correlation with pulmonary function tests,” *Insights into Imaging*, vol. 10, no. 1, p. 102, 2019. [Online]. Available: <http://dx.doi.org/10.1186/s13244-019-0776-9>
Cited on page 96.
- [197] P. A. Gevenois, P. De Vuyst, V. De Maertelaer, J. Zanen, D. Jacobovitz, M. G. Cosio, and J.-C. Yernault, “Comparison of computed density and microscopic morphometry in pulmonary emphysema.” *American Journal of Respiratory and Critical Care Medicine*, vol. 154, no. 1, pp. 187–192, 1996.

- [Online]. Available: <http://dx.doi.org/10.1164/ajrccm.154.1.8680679>
Cited on page 97.
- [198] B. Ganeshan, K. A. Miles, R. C. Young, and C. R. Chatwin, “In search of biologic correlates for liver texture on portal-phase CT,” *Academic Radiology*, vol. 14, no. 9, pp. 1058–1068, 2007. [Online]. Available: <http://dx.doi.org/10.1016/j.acra.2007.05.023>
Cited on page 101.
- [199] A. Bora, C. Alptekin, A. Yavuz, A. Batur, Z. Akdemir, and M. Berköz, “Assessment of liver volume with computed tomography and comparison of findings with ultrasonography,” *Abdominal Imaging*, vol. 39, no. 6, pp. 1153–1161, 2014. [Online]. Available: <http://dx.doi.org/10.1007/s00261-014-0146-5>
Cited on page 101.
- [200] M. K. Kalra, M. M. Maher, R. S. Kamath, T. Horiuchi, T. L. Toth, E. F. Halpern, and S. Saini, “Sixteen-detector row CT of abdomen and pelvis: study for optimization of Z-axis modulation technique performed in 153 patients,” *Radiology*, vol. 233, no. 1, pp. 241–249, 2004. [Online]. Available: <http://dx.doi.org/10.1148/radiol.2331031505>
Cited on pages 101, 107, and 108.
- [201] F. Pontana, J. Pagniez, T. Flohr, J.-B. Faivre, A. Duhamel, J. Remy, and M. Remy-Jardin, “Chest computed tomography using iterative reconstruction vs filtered back projection (Part 1): evaluation of image noise reduction in 32 patients,” *European Radiology*, vol. 21, no. 3, pp. 627–635, 2011. [Online]. Available: <http://dx.doi.org/10.1007/s00330-010-1990-5>
Cited on pages 101 and 107.
- [202] M. Wetzl, M. S. May, D. Weinmann, M. Hammon, C. Treutlein, M. Zeilinger, A. Kiefer, R. Trollmann, J. Woelfle, M. Uder *et al.*, “Dual-source computed tomography of the lung with spectral shaping and advanced iterative reconstruction: potential for maximum radiation dose reduction,” *Pediatric Radiology*, vol. 50, pp. 1240–1248, 2020. [Online]. Available: <http://dx.doi.org/10.1007/s00247-020-04714-0>
Cited on pages 101 and 107.
- [203] L. Lenga, M. Lange, S. S. Martin, M. H. Albrecht, C. Booz, I. Yel, C. T. Arendt, T. J. Vogl, and D. Leithner, “Head and neck single-and dual-energy CT: differences in radiation dose and image quality of 2nd and 3rd generation dual-source CT,” *The British Journal of Radiology*, vol. 94, p. 20210069, 2021. [Online]. Available: <http://dx.doi.org/10.1259/bjr.20210069>
Cited on pages 101, 103, 107, and 108.

- [204] Y. Noda, Y. Iritani, N. Kawai, T. Miyoshi, T. Ishihara, F. Hyodo, and M. Matsuo, “Deep learning image reconstruction for pancreatic low-dose computed tomography: comparison with hybrid iterative reconstruction,” *Abdominal Radiology*, pp. 1–7, 2021. [Online]. Available: <http://dx.doi.org/10.1007/s00261-021-03111-x>
Cited on pages 101 and 107.
- [205] J. M. Bland and D. G. Altman, “Statistical methods for assessing agreement between two methods of clinical measurement,” *The Lancet*, vol. 327, no. 8476, pp. 307–310, 1986. [Online]. Available: [http://dx.doi.org/10.1016/S0140-6736\(86\)90837-8](http://dx.doi.org/10.1016/S0140-6736(86)90837-8)
Cited on page 104.
- [206] S. J. Riederer, N. J. Pelc, and D. A. Chesler, “The noise power spectrum in computed X-ray tomography,” *Physics in Medicine & Biology*, vol. 23, no. 3, p. 446, 1978. [Online]. Available: <http://dx.doi.org/10.1088/0031-9155/23/3/008>
Cited on page 107.
- [207] F. L. Goerner and G. D. Clarke, “Measuring signal-to-noise ratio in partially parallel imaging MRI,” *Medical Physics*, vol. 38, no. 9, pp. 5049–5057, 2011. [Online]. Available: <http://dx.doi.org/10.1118/1.3618730>
Cited on page 107.
- [208] O. Christianson, J. Winslow, D. P. Frush, and E. Samei, “Automated technique to measure noise in clinical CT examinations,” *American Journal of Roentgenology*, vol. 205, no. 1, pp. W93–W99, 2015. [Online]. Available: <http://dx.doi.org/10.2214/AJR.14.13613>
Cited on page 107.
- [209] X. Tian and E. Samei, “Accurate assessment and prediction of noise in clinical CT images,” *Medical Physics*, vol. 43, no. 1, pp. 475–482, 2016. [Online]. Available: <http://dx.doi.org/10.1118/1.4938588>
Cited on page 107.
- [210] A. Malkus and T. P. Szczykutowicz, “A method to extract image noise level from patient images in CT,” *Medical Physics*, vol. 44, no. 6, pp. 2173–2184, 2017. [Online]. Available: <http://dx.doi.org/10.1002/mp.12240>
Cited on page 107.
- [211] X. Yi, X. Guan, C. Chen, Y. Zhang, Z. Zhang, M. Li, P. Liu, A. Yu, X. Long, L. Liu *et al.*, “Adrenal incidentaloma: machine learning-based quantitative texture analysis of unenhanced CT can effectively differentiate sPHEO from lipid-poor adrenal adenoma,” *Journal of Cancer*, vol. 9, no. 19, p. 3577, 2018. [Online]. Available: <http://dx.doi.org/10.7150/jca.26356>
Cited on page 107.

- [212] T. Masuda, T. Nakaura, Y. Funama, T. Okimoto, T. Sato, T. Higaki, N. Noda, N. Imada, Y. Baba, and K. Awai, “Machine-learning integration of CT histogram analysis to evaluate the composition of atherosclerotic plaques: validation with IB-IVUS,” *Journal of cardiovascular computed tomography*, vol. 13, no. 2, pp. 163–169, 2019. [Online]. Available: <http://dx.doi.org/10.1016/j.jcct.2018.10.018>
Cited on page 107.
- [213] D. Han, M. A. Heuvelmans, and M. Oudkerk, “Volume versus diameter assessment of small pulmonary nodules in CT lung cancer screening,” *Translational lung cancer research*, vol. 6, no. 1, p. 52, 2017. [Online]. Available: <http://dx.doi.org/10.21037/tlcr.2017.01.05>
Cited on page 107.
- [214] R. Boellaard, R. Delgado-Bolton, W. J. Oyen, F. Giammarile, K. Tatsch, W. Eschner, F. J. Verzijlbergen, S. F. Barrington, L. C. Pike, W. A. Weber *et al.*, “FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0,” *European journal of nuclear medicine and molecular imaging*, vol. 42, no. 2, pp. 328–354, 2015. [Online]. Available: <http://dx.doi.org/10.1007/s00259-014-2961-x>
Cited on page 107.
- [215] C. T. Jensen, X. Liu, E. P. Tamm, A. G. Chandler, J. Sun, A. C. Morani, S. Javadi, and N. A. Wagner-Bartak, “Image quality assessment of abdominal CT by use of new deep learning image reconstruction: initial experience,” *American Journal of Roentgenology*, vol. 215, no. 1, pp. 50–57, 2020. [Online]. Available: <http://dx.doi.org/10.2214/AJR.19.22332>
Cited on page 107.
- [216] T. Araki, M. Nishino, O. E. Zazueta, W. Gao, J. Dupuis, Y. Okajima, J. C. Latourelle, I. O. Rosas, T. Murakami, G. T. O’Connor *et al.*, “Paraseptal emphysema: prevalence and distribution on CT and association with interstitial lung abnormalities,” *European journal of radiology*, vol. 84, no. 7, pp. 1413–1418, 2015. [Online]. Available: <http://dx.doi.org/10.1016/j.ejrad.2015.03.010>
Cited on page 113.
- [217] M. Wilgus, F. Abtin, D. Markovic, D. Tashkin, J. Phillips, R. Buhr, M. Flynn, M. Dembek, C. Cooper, and I. Barjaktarevic, “Panlobular emphysema is associated with COPD disease severity: A study of emphysema subtype by computed tomography,” *Respiratory Medicine*, vol. 192, p. 106717, 2022. [Online]. Available: <http://dx.doi.org/10.1016/j.rmed.2021.106717>
Cited on page 113.
- [218] P. R. Goddard, E. Nicholson, G. Laszlo, and I. Watt, “Computed tomography in pulmonary emphysema,” *Clinical Radiology*, vol. 33, no. 4,

pp. 379–387, 1982. [Online]. Available: [http://dx.doi.org/10.1016/S0009-9260\(82\)80301-2](http://dx.doi.org/10.1016/S0009-9260(82)80301-2)

Cited on pages 113, 124, and 133.

- [219] G. R. Washko, G. J. Criner, Z. Mohsenifar, F. C. Sciurba, A. Sharafkhaneh, B. J. Make, E. A. Hoffman, J. J. Reilly, N. R. Group *et al.*, “Computed tomographic-based quantification of emphysema and correlation to pulmonary function and mechanics,” *COPD: Journal of Chronic Obstructive Pulmonary Disease*, vol. 5, no. 3, pp. 177–186, 2008. [Online]. Available: <http://dx.doi.org/10.1080/15412550802093025>

Cited on pages 113, 124, and 133.

- [220] H. Makita, Y. Nasuhara, K. Nagai, Y. Ito, M. Hasegawa, T. Betsuyaku, Y. Onodera, N. Hizawa, and M. Nishimura, “Characterisation of phenotypes based on severity of emphysema in chronic obstructive pulmonary disease,” *Thorax*, vol. 62, no. 11, pp. 932–937, 2007. [Online]. Available: <http://dx.doi.org/10.1136/thx.2006.072777>

Cited on pages 113, 124, and 133.

- [221] M. M. W. Wille, L. H. Thomsen, A. Dirksen, J. Petersen, J. H. Pedersen, and S. B. Shaker, “Emphysema progression is visually detectable in low-dose CT in continuous but not in former smokers,” *European radiology*, vol. 24, no. 11, pp. 2692–2699, 2014. [Online]. Available: <http://dx.doi.org/10.1007/s00330-014-3294-7>

Cited on pages 113, 124, 125, and 133.

- [222] M. O. Faruque, J. M. Vonk, U. Bültmann, and H. M. Boezen, “Airborne occupational exposures and inflammatory biomarkers in the lifelines cohort study,” *Occupational and Environmental Medicine*, vol. 78, no. 2, pp. 82–85, 2021. [Online]. Available: <http://dx.doi.org/10.1136/oemed-2020-106493>

Cited on page 114.

- [223] A. Zapf, S. Castell, L. Morawietz, and A. Karch, “Measuring inter-rater reliability for nominal data— which coefficients and confidence intervals are appropriate?” *BMC medical research methodology*, vol. 16, no. 1, pp. 1–10, 2016. [Online]. Available: <http://dx.doi.org/10.1186/s12874-016-0200-9>

Cited on page 118.

- [224] K. Krippendorff, “Estimating the reliability, systematic error and random error of interval data,” *Educational and Psychological Measurement*, vol. 30, no. 1, pp. 61–70, 1970. [Online]. Available: <http://dx.doi.org/10.1177/001316447003000105>

Cited on page 118.

- [225] J. Vikgren, M. Khalil, K. Cederlund, K. Sörensen, M. Boijesen, J. Brandberg, E. Lampa, M. C. Sköld, P. Wollmer, E. Lindberg *et al.*,

- “Visual and quantitative evaluation of emphysema: a case-control study of 1111 participants in the Pilot Swedish CARDioPulmonary BioImage Study (SCAPIS),” *Academic Radiology*, vol. 27, no. 5, pp. 636–643, 2019. [Online]. Available: <http://dx.doi.org/10.1016/j.acra.2019.06.019>
Cited on page 125.
- [226] G. N. Hounsfield, “Computerized transverse axial scanning (tomography): Part 1. Description of system,” *The British journal of radiology*, vol. 46, no. 552, pp. 1016–1022, 1973. [Online]. Available: <http://dx.doi.org/10.1259/0007-1285-46-552-1016>
Cited on page 131.
- [227] L. Kreel, “Computerized tomography using the EMI general purpose scanner,” *The British Journal of Radiology*, vol. 50, no. 589, pp. 2–14, 1977. [Online]. Available: <http://dx.doi.org/10.1259/0007-1285-50-589-2>
Cited on pages 131 and 133.
- [228] L. J. Rosenblum, R. A. Mauceri, D. E. Wellenstein, D. A. Bassano, W. N. Cohen, and E. R. Heitzman, “Computed tomography of the lung,” *Radiology*, vol. 129, no. 2, pp. 521–524, 1978. [Online]. Available: <http://dx.doi.org/10.1148/129.2.521>
Cited on pages 131 and 133.
- [229] K. J. Park, C. J. Bergin, and J. L. Clausen, “Quantitation of emphysema with three-dimensional CT densitometry: comparison with two-dimensional analysis, visual emphysema scores, and pulmonary function test results,” *Radiology*, vol. 211, no. 2, pp. 541–547, 1999. [Online]. Available: <http://dx.doi.org/10.1148/radiology.211.2.r99ma52541>
Cited on page 131.
- [230] J. D. Schroeder, A. S. McKenzie, J. A. Zach, C. G. Wilson, D. Curran-Everett, D. S. Stinson, J. D. Newell Jr, and D. A. Lynch, “Relationships between airflow obstruction and quantitative CT measurements of emphysema, air trapping, and airways in subjects with and without chronic obstructive pulmonary disease,” *AJR. American journal of roentgenology*, vol. 201, no. 3, p. W460, 2013. [Online]. Available: <http://dx.doi.org/10.2214/AJR.12.10102>
Cited on page 131.
- [231] A. M. Fischer, A. Varga-Szemes, M. van Assen, L. P. Griffith, P. Sahbaee, J. I. Sperl, J. W. Nance, and U. J. Schoepf, “Comparison of artificial intelligence-based fully automatic chest CT emphysema quantification to pulmonary function testing,” *AJR Am J Roentgenol*, vol. 214, no. 5, pp. 1065–1071, 2020. [Online]. Available: <http://dx.doi.org/10.2214/AJR.19.21572>
Cited on page 131.

- [232] R. J. Cropp, P. Seslija, D. Tso, and Y. Thakur, "Scanner and kVp dependence of measured CT numbers in the ACR CT phantom," *Journal of applied clinical medical physics*, vol. 14, no. 6, pp. 338–349, 2013. [Online]. Available: <http://dx.doi.org/10.1120/jacmp.v14i6.4417>
Cited on page 132.
- [233] M. Mascalchi, G. Camiciottoli, and S. Diciotti, "Lung densitometry: why, how and when," *Journal of Thoracic Disease*, vol. 9, no. 9, pp. 3319–3345, 2017. [Online]. Available: <http://dx.doi.org/10.21037/jtd.2017.08.17>
Cited on pages 132 and 133.
- [234] C. Heussel, F. Herth, J. Kappes, R. Hantusch, S. Hartlieb, O. Weinheimer, H. Kauczor, and R. Eberhardt, "Fully automatic quantitative assessment of emphysema in computed tomography: comparison with pulmonary function testing and normal values," *European radiology*, vol. 19, no. 10, pp. 2391–2402, 2009. [Online]. Available: <http://dx.doi.org/10.1007/s00330-009-1437-z>
Cited on page 132.
- [235] M. O. Wielpütz, D. Bardarova, O. Weinheimer, H.-U. Kauczor, M. Eichinger, B. J. Jobst, R. Eberhardt, M. Koenigkam-Santos, M. Puderbach, and C. P. Heussel, "Variation of densitometry on computed tomography in COPD—influence of different software tools," *PLOS One*, vol. 9, no. 11, p. e112898, 2014. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0112898>
Cited on page 133.
- [236] H.-j. Lim, O. Weinheimer, M. O. Wielpütz, J. Dinkel, T. Hielscher, D. Gompelmann, H.-U. Kauczor, and C. P. Heussel, "Fully automated pulmonary lobar segmentation: influence of different prototype software programs onto quantitative evaluation of chronic obstructive lung disease," *PLOS One*, vol. 11, no. 3, p. e0151498, 2016. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0151498>
Cited on page 133.
- [237] G. Veronesi, D. R. Baldwin, C. I. Henschke, S. Ghislandi, S. Iavicoli, M. Oudkerk, H. J. De Koning, J. Shemesh, J. K. Field, J. J. Zulueta *et al.*, "Recommendations for implementing lung cancer screening with low-dose computed tomography in Europe," *Cancers*, vol. 12, no. 6, p. 1672, 2020. [Online]. Available: <http://dx.doi.org/10.3390/cancers12061672>
Cited on page 134.
- [238] American Association of Physicists in Medicine, "AAPM Position Statement on Radiation Risks from Medical Imaging Procedures - PP25C," 2018. [Online]. Available: <http://web.archive.org/web/20200804223708/>

<https://www.aapm.org/org/policies/details.asp?id=439&type=PP>

Cited on page 134.

- [239] R. Meza, J. Jeon, I. Toumazis, K. Ten Haaf, P. Cao, M. Bastani, S. S. Han, E. F. Blom, D. E. Jonas, E. J. Feuer *et al.*, “Evaluation of the benefits and harms of lung cancer screening with low-dose computed tomography: modeling study for the US Preventive Services Task Force,” *Jama*, vol. 325, no. 10, pp. 988–997, 2021. [Online]. Available: <http://dx.doi.org/10.1001/jama.2021.1077>

Cited on page 134.

- [240] M. Oudkerk, S. Liu, M. A. Heuvelmans, J. E. Walter, and J. K. Field, “Lung cancer LDCT screening and mortality reduction—evidence, pitfalls and future perspectives,” *Nature reviews Clinical oncology*, vol. 18, no. 3, pp. 135–151, 2021. [Online]. Available: <http://dx.doi.org/10.1038/s41571-020-00432-6>

Cited on page 134.

- [241] J. H. Zurawska, R. Jen, S. Lam, H. O. Coxson, J. Leipsic, and D. D. Sin, “What to do when a smoker’s CT scan is “normal”?: implications for lung cancer screening,” *Chest*, vol. 141, no. 5, pp. 1147–1152, 2012. [Online]. Available: <http://dx.doi.org/10.1378/chest.11-1863>

Cited on page 134.

- [242] G. Sidorenkov, R. Stadhouders, C. Jacobs, F. A. Mohamed Hoessein, H. A. Gietema, K. Nackaerts, Z. Saghir, M. A. Heuvelmans, H. C. Donker, J. G. Aerts *et al.*, “Multi-source data approach for personalized outcome prediction in lung cancer screening: update from the NELSON trial,” *European Journal of Epidemiology*, pp. 1–10, 2023. [Online]. Available: <http://dx.doi.org/10.1007/s10654-023-00975-9>

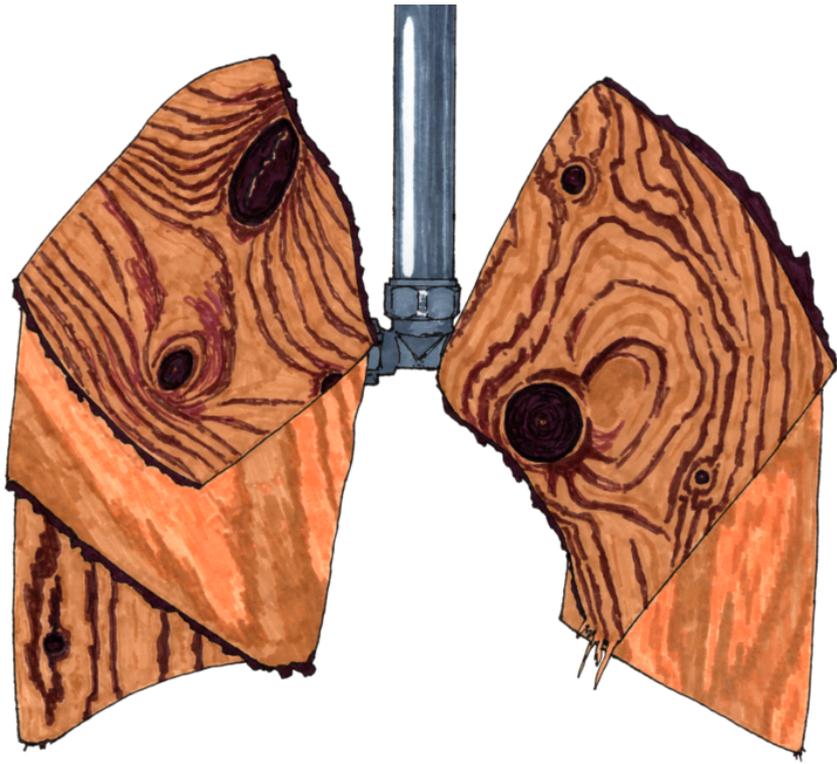
Cited on page 134.

- [243] A. Mansoor, U. Bagci, B. Foster, Z. Xu, G. Z. Papadakis, L. R. Folio, J. K. Udupa, and D. J. Mollura, “Segmentation and image analysis of abnormal lungs at CT: current approaches, challenges, and future trends,” *Radiographics*, vol. 35, no. 4, pp. 1056–1076, 2015. [Online]. Available: <http://dx.doi.org/10.1148/rg.2015140232>

Cited on page 136.

- [244] S. M. Humphries, A. M. Notary, J. P. Centeno, M. J. Strand, J. D. Crapo, E. K. Silverman, D. A. Lynch, and G. E. of COPD (COPDGene) Investigators, “Deep Learning Enables Automatic Classification of Emphysema Pattern at CT,” *Radiology*, vol. 294, no. 2, pp. 434–444, 2019. [Online]. Available: <http://dx.doi.org/10.1148/radiol.2019191022>

Cited on page 136.



Supplementary materials

Supplementary materials chapter 2

Table S2.1 can be found on page 183

Table S2.2 can be found on page 184

Table S2.3 can be found on page 185

Table S2.4 can be found on page 186

Table S2.5 can be found on page 187

Table S2.6 can be found on page 188

Table S2.7 can be found on page 189

Table S2.1: Definitions of collected variables in the two cohorts.

Variables	Chinese cohort	Dutch cohort
Smoking status		
Current smokers	Participants who smoked ≥ 1 cigarette a day for ≥ 6 months and did not quit before the interview.	Participants who smoked for ≥ 1 year and either were still smoking or had quit < 1 month before the assessment.
Never-smokers	Participants who had never smoked or smoked for < 6 months.	Participants who had never smoked or smoked for < 1 year.
Passive smoking	Exposure to smoke produced by others ≥ 1 day a week for ≥ 15 minutes indoors.	Regularly exposed to tobacco smoke from others in the past year.
Cooking/fireplace fumes exposure	The presence of at least moderate smoke during cooking.	Fireplace use ≥ 1 time/week.
Educational level		
Low	Finished at most lower secondary education.	
Moderate	Finished upper secondary and post-secondary.	
High	Finished at least bachelor or equivalent.	

Table S2.2: Characteristics of participants in the Chinese and Dutch cohorts, stratified by smoking status.

Characteristics	Chinese Cohort, n (%)				Dutch cohort, n (%)			
	Overall	Current smoker	Former smoker	Never-smoker	Overall	Current smoker	Former smoker	Never-smoker
Number of participants	1143 (48.8)	269 (23.5)	115 (10.1)	759 (66.4)	1200 (51.2)	170 (14.2)	571 (47.6)	459 (38.3)
Age (mean±SD)	61.7±6.3	62.1±5.6	65.2±5.0	61.1±6.5	59.8±8.1	57.6±7.8	61.8±7.4	58.1±8.5
Sex								
Women	627 (54.9)	20 (7.4)	0 (0)	607 (80.0)	707 (58.9)	88 (51.8)	337 (59.0)	282 (61.4)
Men	516 (45.1)	249 (92.6)	115 (100)	152 (20.0)	493 (41.1)	82 (48.2)	234 (41.0)	177 (38.6)
Passive smoking								
No	640 (56.0)	94 (34.9)	56 (48.7)	490 (64.6)	929 (77.4)	89 (52.4)	452 (79.2)	388 (84.5)
Yes	503 (44.0)	175 (65.1)	59 (51.3)	269 (35.4)	271 (22.6)	81 (47.6)	119 (20.8)	71 (15.5)
BMI (kg/m ²)								
<25	643 (56.3)	155 (57.6)	50 (43.5)	438 (57.7)	473 (39.4)	57 (33.5)	213 (37.3)	203 (44.2)
≥25	500 (43.7)	114 (42.4)	65 (56.5)	321 (42.3)	727 (60.6)	113 (66.5)	358 (62.7)	256 (55.8)
Educational level								
Low	431 (37.7)	121 (45.0)	49 (42.6)	261 (34.4)	242 (20.2)	38 (22.4)	126 (22.1)	78 (17.0)
Moderate	418 (36.6)	86 (32.0)	35 (30.4)	297 (39.1)	615 (51.2)	94 (55.3)	294 (51.5)	227 (49.5)
High	294 (25.7)	62 (23.0)	31 (27.0)	201 (26.5)	343 (28.6)	38 (22.4)	151 (26.4)	154 (33.6)
Fume exposure								
No	1066 (93.3)	238 (88.5)	105 (91.3)	723 (95.3)	1127 (93.9)	157 (92.4)	538 (94.2)	432 (94.1)
Yes	77 (6.7)	31 (11.5)	10 (8.7)	36 (4.7)	73 (6.1)	13 (7.6)	33 (5.8)	27 (5.9)

BMI: body mass index; SD: standard deviation.

Table S2.3: Distribution of subtypes and severity of emphysema (\geq mild) in participants with emphysema in the Chinese and Dutch cohorts.

	Chinese cohort n=235, n (%)	Dutch cohort n=188, n (%)	p-value
Predominant subtype of emphysema			<0.0001*#
CLE	189 (80.4)	105 (55.9)	
PSE	46 (19.6)	83 (44.1)	
Severity of CLE			0.5585#
Mild	132 (69.8)	70 (66.7)	
Moderate	32 (16.9)	23 (21.9)	
Confl-adv	25 (13.2)	12 (11.4)	
Severity of PSE			1.000§
Mild	44 (95.7)	79 (95.2)	
Substantial	2 (4.3)	4 (4.8)	

CLE centrilobular emphysema; PSE paraseptal emphysema;
Confl-adv confluent or advanced destructive.

* $p < 0.05$; # χ^2 test; § Fisher's Exact Test.

Table S2.4: Associations between participant characteristics and the presence of CLE (\geq trace) versus participants without emphysema.

Characteristics	Univariate logistical regression			Multivariable logistical regression		
	OR	95% CI	p-value	OR	95% CI	p-value
Dutch cohort	1.00			1.00		
Chinese cohort	2.45	2.06 – 2.91	<0.0001*	2.19	1.77 – 2.70	<0.0001*
Age (per 10 years increase)	1.68	1.49 – 1.89	<0.0001*	1.51	1.32 – 1.72	<0.0001*
Women	1.00			1.00		
Men	2.11	1.78 – 2.51	<0.0001*	1.55	1.28 – 1.89	<0.0001*
Smoking status						
Never	1.00		<0.0001*			<0.0001*
Former	1.13	0.93 – 1.37	0.22	1.44	1.13 – 1.82	0.0029*
Current	2.90	2.28 – 3.68	<0.0001*	2.31	1.75 – 3.05	<0.0001*
Control	1.00			1.00		
Passive smoking	1.72	1.44 – 2.06	<0.0001*	1.25	1.02 – 1.53	0.0312*
BMI < 25 kg/m ²	1.00			1.00		
BMI \geq 25 kg/m ²	0.69	0.58 – 0.81	<0.0001*	0.70	0.59 – 0.85	0.0002*
Educational level						
Low	1.00		<0.0001*			0.1300
Moderate	0.55	0.45 – 0.68	<0.0001*	0.80	0.64 – 0.99	0.0440*
High	0.66	0.52 – 0.82	<0.0002*	0.88	0.69 – 1.13	0.32
Control	1.00			1.00		
Fume exposure	1.63	1.16 – 2.29	0.0054*	1.37	0.94 – 2.00	0.10

95% CI: 95% confidence interval; BMI: body mass index; OR: odds ratio. * $p < 0.05$.

Table S2.5: Associations between participant characteristics and the presence of CLE (\geq mild) versus participants without emphysema.

Characteristics	Univariate logistical regression			Multivariable logistical regression		
	OR	95% CI	p-value	OR	95% CI	p-value
Dutch cohort	1.00			1.00		
Chinese cohort	2.01	1.56 – 2.59	<0.0001*	1.58	1.15 – 2.17	0.0048*
Age (per 10 years increase)	1.79	1.49 – 2.15	<0.0001*	1.67	1.35 – 2.05	<0.0001*
Women	1.00			1.00		
Men	3.30	2.54 – 4.28	<0.0001*	1.78	1.31 – 2.42	<0.0002*
Smoking status						
Never			<0.0001*			<0.0001*
Former	1.82	1.32 – 2.50	0.0003*	1.87	1.29 – 2.71	0.0010*
Current	6.04	4.46 – 8.18	<0.0001*	4.67	3.27 – 6.67	<0.0001*
Control	1.00			1.00		
Passive smoking	1.59	1.24 – 2.05	0.0003*	0.98	0.73 – 1.30	0.86
BMI < 25 kg/m ²	1.00			1.00		
BMI \geq 25 kg/m ²	0.63	0.49 – 0.80	0.0002*	0.62	0.47 – 0.81	0.0005*
Educational level						
Low	1.00		0.0013*			0.59
Moderate	0.59	0.44 – 0.79	0.0003*	0.85	0.62 – 1.16	0.31
High	0.70	0.51 – 0.96	0.0250*	0.90	0.64 – 1.27	0.56
Control	1.00			1.00		
Fume exposure	1.73	1.13 – 2.66	0.0117*	1.28	0.80 – 2.05	0.30

95% CI: 95% confidence interval; BMI: body mass index; OR: odds ratio. * $p < 0.05$.

Table S2.6: Multivariable associations between participant characteristics and the presence of CLE (\geq trace) versus participants without emphysema, stratified by smoking status.

Variables	Current smokers (n=388)		Former smokers (n=633)		Never-smokers (n=1193)	
	Adjusted OR (95% CI)	p-value	Adjusted OR (95% CI)	p-value	Adjusted OR (95% CI)	p-value
Dutch cohort	1.00		1.00		1.00	
Chinese cohort	1.43 (0.80 – 2.56)	0.22	1.92 (1.12 – 3.30)	0.0183*	2.61 (1.97 – 3.45)	<0.0001*
Age (per 10 years increase)	1.84 (1.28 – 2.64)	0.0009*	1.58 (1.21 – 2.05)	0.0008*	1.41 (1.18 – 1.68)	0.0001*
Women	1.00				1.00	
Men	0.99 (0.54 – 1.80)	0.96	1.60 (1.09 – 2.33)	0.0154*	1.76 (1.32 – 2.34)	0.0001*
Control	1.00		1.00		1.00	
Passive smoking	0.95 (0.59 – 1.55)	0.84	1.72 (1.16 – 2.56)	0.0074*	1.10 (0.84 – 1.45)	0.49
Quit smoking years	—		0.99 (0.97 – 1.01)	0.17	—	
Pack-years	1.04 (1.02 – 1.06)	0.0002*	—		—	
BMI < 25 kg/m ²	1.00		1.00		1.00	
BMI \geq 25 kg/m ²	0.47 (0.29 – 0.77)	0.0024*	0.72 (0.51 – 1.02)	0.07	0.75 (0.59 – 0.96)	0.0231*
Educational level						
Low	1.00	0.13	1.00	0.16	1.00	0.93
Moderate	0.67 (0.38 – 1.17)	0.16	0.67 (0.44 – 1.02)	0.06	0.96 (0.71 – 1.29)	0.77
High	1.19 (0.62 – 2.30)	0.60	0.72 (0.45 – 1.16)	0.18	0.94 (0.68 – 1.31)	0.73
Control	1.00		1.00		1.00	
Fume exposure	1.09 (0.50 – 2.40)	0.82	1.57 (0.79 – 3.12)	0.20	1.41 (0.81 – 2.45)	0.23

95% CI: 95% confidence interval; BMI: body mass index; OR: odds ratio. Pack-years or quit smoking years was adjusted among current and former smokers, respectively. * $p < 0.05$.

Table S2.7: Multivariable associations between participant characteristics and the presence of CLE (\geq mild) versus participants without emphysema, stratified by smoking status.

Variables	Current smokers (n=388)		Former smokers (n=633)		Never-smokers (n=1193)	
	Adjusted OR (95% CI)	p-value	Adjusted OR (95% CI)	p-value	Adjusted OR (95% CI)	p-value
Dutch cohort	1		1		1	
Chinese cohort	0.87 (0.46 – 1.64)	0.6734	1.49 (0.77 – 2.89)	0.2365	2.54 (1.47 – 4.40)	0.0009*
Age (per 10 years increase)	1.92 (1.30 – 2.82)	0.0010*	1.69 (1.12 – 2.54)	0.0116*	1.57 (1.13 – 2.20)	0.0081*
Women	1		1		1	
Men	1.45 (0.69 – 2.68)	0.3836	1.53 (0.84 – 2.78)	0.1685	2.44 (1.51 – 3.93)	0.0003*
Control	1		1		1	
Passive smoking	0.75 (0.46 – 1.20)	0.2309	1.31 (0.76 – 2.24)	0.3317	0.92 (0.55 – 1.53)	0.7494
Quit smoking years	—		0.97 (0.95 – 0.99)	0.0183*	—	
Pack-years	1.02 (1.01 – 1.04)	0.0060*	—		—	
BMI < 25 kg/m ²	1		1		1	
BMI \geq 25 kg/m ²	0.41 (0.25 – 0.66)	0.0002*	0.57 (0.35 – 0.94)	0.0261*	0.91 (0.58 – 1.43)	0.6755
Educational level						
Low	1	0.3473	1	0.0980	1	0.2478
Moderate	0.68 (0.39 – 1.19)	0.1771	0.60 (0.34 – 1.07)	0.0847	1.60 (0.91 – 2.82)	0.1058
High	0.96 (0.52 – 1.76)	0.8903	0.52 (0.27 – 1.01)	0.0536	1.49 (0.82 – 2.73)	0.1936
Control	1		1		1	
Fume exposure	1.44 (0.69 – 2.97)	0.3305	1.33 (0.54 – 3.28)	0.5382	1.00 (0.40 – 2.52)	0.9948

95% CI: 95% confidence interval; BMI: body mass index; OR: odds ratio. Pack-years or quit smoking years was adjusted among current and former smokers, respectively. * $p < 0.05$.

Supplementary materials chapter 3

Table S3.1 can be found on page 191

Table S3.2 can be found on page 192

Table S3.3 can be found on page 202

Table S3.4 can be found on page 203

Table S3.5 can be found on page 208

Table S3.6 can be found on page 213

Figure S3.1: can be found on page 214

Figure S3.2: can be found on page 214

Figure S3.3: can be found on page 215

Figure S3.4: can be found on page 215

Table S3.1: Search strategy by database.

The search strategy was optimised by a medical information specialist for terms that specified exposure and outcome. We also checked the references of included articles to identify any that were missed in the initial searches.

Database	Search strategy
PubMed	("Pulmonary Emphysema"[Mesh] OR pulmonary emphysema*[tiab]) AND ("Lung Neoplasms"[Mesh] OR "Solitary Pulmonary Nodule"[Mesh] OR lung nodule*[tiab] OR pulmonary nodule*[tiab] OR lung neoplasm*[tiab] OR lung cancer*[tiab] OR lung tumor*[tiab] OR lung tumour*[tiab] OR lung malignanc*[tiab]) NOT ("animals"[MeSH] NOT "humans"[MeSH])
Embase	("Lung Emphysema"/exp OR 'lung emphysema*':ti,ab) AND ('Lung cancer'/exp OR 'lung nodule'/exp OR 'lung nodule*':ab,ti OR 'pulmonary nodule*':ab,ti OR 'lung neoplasm*':ti,ab OR 'lung tumor*':ti,ab OR 'lung tumour*':ti,ab OR 'lung cancer*':ab,ti OR 'lung malignanc*':ab,ti) NOT ('animal'/exp NOT 'human'/exp)
Cochrane	"pulmonary emphysema*" AND ("lung nodule*" OR "pulmonary nodule*" OR "lung neoplasm*" OR "lung cancer*" OR "lung tumor*" OR "lung tumour*" OR "lung malignanc*")

Table S3.2: Definition of visual and quantitative emphysema and CT scan parameters across the included studies

Study	CT scan parameters	Definition of emphysema	Definition of severity of emphysema	Matched or adjusted factors
de Torres 2007 [91] (<i>visual assessment</i>)	CT scanner: Siemens (single-slice helical scanner, Somatom Volume Zoom); Scanning mode: Low-dose CT; Slice thickness: 1.25 mm; Reconstruction algorithm: High spatial frequency	Definition: ≥ 1 score where areas of vascular, lung disruption and low attenuation value occupy up to 25 % of any 3 apical-to-basal lung zones	NA	Adjusted: age, gender, pack-years, airflow obstruction
Wilson 2008 [90] (<i>visual assessment</i>)	CT scanner: GE (multidetector); Scanning mode: Low-dose CT; Slice thickness: Not specified; Reconstruction algorithm: High spatial frequency	Definition: ≥ 1 score where areas of vascular, lung disruption and low attenuation value occupy up to 10 % of any 3 apical-to-basal lung zones	Five-level semiquantitative scale, based on modified NETT. Trace: low trace attenuation value occupy 0–10 %; mild: 11–25 %; moderate: 26–50 %; severe emphysema: >50 %	Adjusted: sex, age, years of smoking, smoking dose intensity, airflow obstruction

Table S3.2: [continued]

Study	CT scan parameters	Definition of emphysema	Definition of severity of emphysema	Matched or adjusted factors
Li 2011 [111] <i>(visual assessment)</i>	CT scanner: Not specified; Scanning mode: Standard-dose CT; Slice thickness: 5.0 mm; Reconstruction algorithm: High spatial frequency	Definition: Estimate of the percentage of lung tissue destroyed by emphysema is more than 0 %	NA	Matched: age, gender, race, area, smoking status Adjusted: pack-years, Airflow obstruction, family history of lung cancer
Maisonneuve 2011 [112] <i>(visual assessment)</i>	CT scanner: GE (8-slice or 16-slice multidetector, High-Speed Advantage); Scanning mode: Low-dose CT; Reconstructed slice thickness: 1.2 mm; Reconstruction algorithm: Not specified	Definition: presence of subtle areas of low attenuation and loss of parenchymal structures that contrast with the surrounding lung parenchyma with normal attenuation	NA	Adjusted: age, gender, asbestos exposure, largest nodule size, nodule type, cigarettes per day, duration of smoking, and quitting
Henschke 2015 [113] <i>(visual assessment)</i>	CT scanner: Not specified; Scanning mode: Low-dose CT; Slice thickness: 1.25 mm; Reconstruction algorithm: Not specified	Definition: discrete areas of decreased attenuation could be identified anywhere in the lung parenchyma	NA	Adjusted: age, female gender, ethnicity

Table S3.2: [continued]

Study	CT scan parameters	Definition of emphysema	Definition of severity of emphysema	Matched or adjusted factors
Sanchez-Salcedo 2015 [114] (<i>visual assessment</i>)	CT scanner: Siemens (single-slice helical scanner, Somatom Volume Zoom); Scanning mode: Low-dose CT; Slice thickness: 1.25 mm; Reconstruction algorithm: Not specified	Definition: ≥ 1 score where areas of vascular, lung disruption and low attenuation value occupy up to 25 % of any 3 apical-to-basal lung zones	NA	Adjusted: age, sex, pack-years, airflow obstruction
de Torres 2015 [115] (<i>visual assessment</i>)	CT scanner: GE Systems scanner; Scanning mode: Low-dose CT; Slice thickness: Not specified; Reconstruction algorithm: High spatial frequency	Definition: <i>geq1</i> score where areas of vascular, lung disruption and low attenuation value occupy up to 10 % of any 3 apical-to-basal lung zones	NA	Adjusted: Age, BMI, pack-years

Table S3.2: [continued]

Study	CT scan parameters	Definition of emphysema	Definition of severity of emphysema	Matched or adjusted factors
Liu 2018 [116] (<i>visual assessment</i>)	CT scanner: Not specified; Scanning mode: Low-dose CT; Reconstruction algorithm: Soft tissue or thin section; Slice thickness: 1.0 – 3.2 mm	Definition: percentage of low attenuation and vascular disruption area in 3 levels (top of aortic arch, tracheal carina, and 2 cm above highest hemidiaphragm) is more than 0 %	NA	Matched: age, sex, race, smoking status Adjusted: age, sex, race, smoking status, pack-years, family history of lung cancer
Gonzalez 2019 [98] (<i>visual assessment</i>)	CT scanner: Siemens (64 detectors Somatom Plus 4) or Siemens Healthcare (Somatom Sensation 64, Somatom Definition); Scanning mode: Low-dose CT; Slice thickness: 1.0 mm; Reconstruction algorithm Kernel: B60	Definition: centrilobular emphysema (estimate of the percentage of centrilobular lucencies is more than 0 % of lung zone); Paraseptal emphysema: presence of a few well-demarcated rounded juxtapleural lucencies	Based on criteria of NETT: Mild: low attenuation value occupy 0–25 %; moderate: 26–50 %; severe emphysema: >50 %	Matched: sex, age, smoking status, pack-years Adjusted: smoking status, airflow obstruction

Table S3.2: [continued]

Study	CT scan parameters	Definition of emphysema	Definition of severity of emphysema	Matched or adjusted factors
Carr 2018 [107] (<i>visual assessment</i>)	CT scanner: Not specified; Scanning mode: Not specified; Slice thickness: 0.75 mm; Reconstruction algorithm: B35f	Definition: centrilobular emphysema (Estimate of the percentage of centrilobular lucencies is more than 0 % of lung zone); Paraseptal emphysema: presence of a few well-demarcated rounded juxtaleural lucencies	Based on criteria of Fleischner Society: Mild: scattered centrilobular lucencies, usually separated by large regions of normal lung, involving an estimated 0.5–5 % of a lung zone; Moderate: many well-defined centrilobular lucencies, occupying >5 % of any lung zone; Confluent: coalescent centrilobular or lobular lucencies, including multiple regions of lucencies that span several secondary pulmonary lobules; Advanced destructive emphysema: panlobular lucencies with hyperexpansion and distortion of pulmonary architecture.	Matched: age, race, sex, and smoking history Adjusted: age, sex, race, smoking status, pack-years, years since quitting, and airflow obstruction

Table S3.2: [continued]

Study	CT scan parameters	Definition of emphysema	Definition of severity of emphysema	Matched or adjusted factors
Yong 2019 [117] (<i>visual assessment</i>)	CT scanner: Not specified; Scanning mode: Low-dose CT; Slice thickness: 1.0 mm; Reconstruction algorithm: Soft tissue	Definition: no specific diagnostic criteria	NA	Adjusted: age, gender, smoking duration, family history of lung cancer, personal history of cancer, history of pneumonia, asbestos exposure
Kishi 2002 [88] (<i>quantitative assessment</i>)	CT scanner: GE (High Speed Advantage); Scanning mode: Low-dose CT; slice thickness: 5 mm; Reconstruction algorithm: Edge-enhancing	Definition: %LAA –900 HU > 5% (dichotomous and continuous)	NA	Matched: sex, age, pack-years Adjusted: pack-years
Maldonado 2010 [87, 118] (<i>quantitative assessment</i>)	CT scanner: GE (High-Speed Advantage); Scanning mode: Low-dose CT; Slice thickness: 5 mm; Reconstruction algorithm: standard	Definition: %LAA –900 HU > 5% (dichotomous and continuous)	%LAA trace: 5–9% Mild: 10–14% Moderate: > 15%	Matched: sex, age, smoking history

Table S3.2: [continued]

Study	CT scan parameters	Definition of emphysema	Definition of severity of emphysema	Matched or adjusted factors
Gierada 2011 [89] (<i>quantitative assessment</i>)	CT scanner: Toshiba (16 slice Aquilion 16), GE (4 slice, HiSpeed Qxi/i) and Siemens (16 slice, Sensation 16); Scanning mode: Low-dose CT; Slice thickness: 1.0 – 2.5 mm; Reconstruction algorithm: FC 51, B50f; standard, C and B30f	Definition: % upper lung LAA –950 HU ≥ 25 % (dichotomous)	NA	Matched: sex, age, and smoking history Adjusted: age, sex, pack-years, BMI, history of asthma
Schwartz 2016 [108] (<i>quantitative assessment</i>)	CT scanner: Not specified; Scanning mode: low-dose CT; Slice thickness: Not specified; Reconstruction algorithm: Not specified	Definition: %LAA –950 HU >4.8 % (dichotomous)	NA	Adjusted: age, race, gender, pack-years; total lung volume

Table S3.2: [continued]

Study	CT scan parameters	Definition of emphysema	Definition of severity of emphysema	Matched or adjusted factors
<p>Aamli Gagnat 2017 [96] <i>(quantitative assessment)</i></p>	<p>CT scanner: GE (8 slice, LightSpeed Ultra); Scanning mode: Standard-dose CT; Slice thickness: 1 mm; Reconstruction algorithm: Not specified</p>	<p>Definition: %LAA $\geq 3\%$ (dichotomous and continuous)</p>	<p>Mild: %LAA 3–10 % Moderate/severe: %LAA $\geq 10\%$</p>	<p>Adjusted: age, sex, smoking status, pack-years, age of smoking, airflow obstruction</p>
<p>Chubachi 2017 [95] <i>(quantitative assessment)</i></p>	<p>CT scanner: GE (64 detectors), LightSpeed VCT, and Discovery CT 750 HD; Toshiba (64 detectors, Aquilion 64), GE (256 detectors, Revolution CT) or Toshiba (320 detectors, Aquilion One Genesis); Scanning mode: Standard-dose CT; Slice thickness: 1.0 – 1.25 mm; Reconstruction algorithm: Chest and FC 50</p>	<p>Definition: %LAA $\geq 10\%$ (dichotomous)</p>	<p>Mild: $10\% \leq \%LAA < 20\%$ Moderate/severe: %LAA $\geq 20\%$</p>	<p>Adjusted: gender, age, and pack-years, and interstitial lung abnormality</p>

Table S3.2: [continued]

Study	CT scan parameters	Definition of emphysema	Definition of severity of emphysema	Matched or adjusted factors
Mouronte-Roibas 2018 [97] (<i>quantitative assessment</i>)	CT scanner: GE (64 detectors), Lightspeed VCT, or Siemens (6 detectors, Somatom Emotion); Scanning mode: Not specified; Slice thickness: Not specified; Reconstruction algorithm: Not specified	Definition: %LAA $\geq 1\%$ (dichotomous)	NA	Adjusted: sex, age, BMI, pack-years
Nishio 2019 [110] (<i>quantitative assessment</i>)	CT scanner: Toshiba (320 or 64 detectors, Aquilion ONE or Aquilion 64); Scanning mode: Standard-dose CT; slice thickness: 0.5 or 1.0 mm; Reconstruction algorithm: Not specified	Definition: %LAA ≥ 880 HU (continuous)	NA	Adjusted: sex, age, smoking history (Brinkman Index)

Table S3.2: [continued]

Study	CT scan parameters	Definition of emphysema	Definition of severity of emphysema	Matched or adjusted factors
Husebø 2019 [119] (<i>quantitative assessment</i>)	CT scanner: Not specified; Scanning mode: Not specified; Slice thickness: Not specified; Reconstruction algorithm: Not specified	Definition: %LAA -950 HU > 10 % (dichotomous)	NA	Adjusted: age, sex, smoking status, pack-years, BMI, use of inhaled steroids
Labaki 2021 [109] (<i>quantitative assessment</i>)	CT scanner: Not specified; Scanning mode: Low-dose CT; Reconstruction algorithm: Soft tissue or thin section; Slice thickness: 1.0 – 3.2 mm	Definition: %LAA -950 HU (continuous)	NA	Adjusted: age, BMI, race, education level, smoking intensity, and duration, time since smoking cessation, self-reported COPD, and a personal and family history of lung cancer

NA: not applicable; NETT: National Emphysema Treatment Trial; LAA: low attenuation area.



Table S8.3: Quality assessment of studies included in the meta-analysis

Author-Year	Selection (4 stars)	Comparability (2 stars)	Exposure/ outcome (3 stars)	Total (Degree) (9 stars)
Case-control studies				
Kishi (2002)	4	2	3	9 (High)
Maldonado (2010)	3	2	3	8 (High)
Gierada (2011)	4	2	3	9 (High)
Li (2011)	3	2	3	8 (High)
Schwartz (2016)	4	2	2	8 (High)
Mouronte-Roibas (2018)	2	2	2	6 (Medium)
Carr (2018)	4	2	2	8 (High)
Liu (2018)	4	2	2	8 (High)
Gonzalez (2019)	4	2	3	9 (High)
Nishio (2019)	3	2	2	7 (Medium)
Cohort studies				
de Torres (2007)	4	2	2	8 (High)
Wilson (2008)	4	2	2	8 (High)
Maisonneuve (2011)	4	2	2	8 (High)
Henschke (2015)	3	2	1	6 (Medium)
de Torres (2015)	4	1	2	7 (Medium)
Sanchez-Salcedo (2015)	4	2	2	8 (High)
Aamli Gagnat (2017)	4	2	3	9 (High)
Chubachi (2017)	4	2	1	7 (Medium)
Husebø (2019)	4	2	3	9 (High)
Yong (2019)	4	2	2	8 (High)
Labaki (2021)	4	1	2	7 (Medium)

Scoring was performed with the Newcastle–Ottawa Scale (NOS), with one star being awarded if the item was met.

Table S3.4: Characteristics of included studies that assessed emphysema visually on chest CT.

Study	Country	With/ without lung ca	Lung ca histologic type	Age (mean±SD)	Source	Study design	Feature of evaluation	Effect size (95% CI)
de Torres 2007 [91]	Spain	23/1166	13 (57%) Adenocarcinoma; 5 (22%) Squamous cell carcinoma; 4 (17%) Small cell carcinoma; 1 (4%) Large cell carcinoma;	Case: 54±8 Control: 54±8	PB	Cohort; Prospective study	Chest radiologist; Guideline: NETT	RR: 2.5 (95 % CI 1.0 – 6.2)
Wilson 2008 [90]	USA	99/3539	86 (87%) Non-small cell carcinoma (NSCLC); 13 (13%) Small cell carcinoma;	NS	PB	Cohort; Prospective study	Pulmonologist, general radiologist, chest radiologist. Guideline: NETT	OR: 3.1 (95 % CI 1.9 – 5.2)

Table S3.4: [continued]

Study	Country	With/ without lung ca	Lung ca histologic type	Age (mean±SD)	Source	Study design	Feature of evaluation	Effect size (95% CI)
Li 2011 [111]	USA	565/450	259 (46%) Adenocarcinoma; 159 (28%) Squamous cell carcinoma; 13 (2%) Large cell; 63 (11%) Other NSCLC; 71 (13%) Small cell carcinoma;	Case: 67±8 Control: 66±6	HB	Case- control; Retro- spective study	Chest radiologist. Guideline: NS	OR: 2.8 (95 % CI 2.1 – 3.8)
Maison neuve 2011 [112]	Italy	85/4511	NS	NS	PB	Cohort; Retro- spective study	Radiologist. Guideline: NS	HR: 1.8 (95 % CI 1.2 – 2.6)
Henschke 2015 [113]	USA	668/61456	NS	NS	PB	Cohort; Prospect- ive study	Radiologist. Guideline: NS	OR: 2.0 (95 % CI 1.4 – 2.9)

Table S3.4: [continued]

Study	Country	With/ without lung ca	Lung ca histologic type	Age (mean±SD)	Source	Study design	Feature of evaluation	Effect size (95% CI)
Sanchez-Salcedo 2015 [114]	Spain	53/2936	53 Participants (60 Lesions): 33 (55%) Adenocarcinoma; 13 (22%) Squamous carcinoma; 7 (12%) Large cell carcinoma; 5 (8%) Small cell carcinoma; 2 (3%) Others;	Case: 60 (55 – 65) [#] Control: 55 (49–62) [#]	HB	Cohort; Prospective study	Chest radiologist. Guideline: NS	HR: 3.3 (95 % CI 1.8 – 5.9)
de Torres 2015 [115]	USA	134/1419	NS	Overall: 61±7	PB	Cohort; Prospective study	Pulmonologist, general radiologist, chest radiologist. Guideline: NETT	HR: 2.7 (95 % CI 1.7 – 4.3)

Table S3.4: [continued]

Study	Country	With/ without lung ca	Lung ca histologic type	Age (mean±SD)	Source	Study design	Feature of evaluation	Effect size (95% CI)
Liu 2018 [116]	USA	73/157	33 (45%) Adenocarcinoma; 21 (29%) Squamous cell carcinoma; 3 (4%) Small cell carcinoma; 16 (22%) Other and not otherwise specified NSCLC;	Case: 64 (55 – 74) [#] Control: 63 (55–74) [#]	PB	Case- control; Prospect- ive study	Clinical radiologist. Guideline: Modified NETT	OR: 1.8 (95 % CI 1.4 – 1.9)
Gonzalez 2019 [98]	Spain	72/215	36 (50%) Adenocarcinoma; 15 (21%) Squamous cell carcinoma; 5 (7%) Small cell carcinoma; 7 (10%) Large cell carcinoma; 7 (10%) Others; 2 (3%) Unknown;	Case: 64±9 Control: 64±9	PB	Case- control; Prospect- ive study	Pulmono- logist; Guideline: Fleischner Society	OR: 5.4 (95 % CI 2.6 – 11.4)

Table S3.4: [continued]

Study	Country	With/ without lung ca	Lung ca histologic type	Age (mean±SD)	Source	Study design	Feature of evaluation	Effect size (95% CI)
Yong 2019 [117]	Norway	367/16 257	NS	Case: 62±6 Control: 61±5	PB	Cohort; Retro- spective study	Radiologist. Guideline: NS	HR: 2.0 (95 % CI 1.6 – 2.6)

HB: hospital-based; HR: hazard ratio; NETT: National Emphysema Treatment Trial; NS: not specified; OR: odds ratio;
PB: population-based; RR: risk ratio; SD: standard deviation.

All the effect sizes adjusted for smoking, except for study Henschke 2015; For specific adjusted factors, see Table S8.3 [p. 202].

Numbers are medians, with ranges in parentheses.



Table S3.5: Characteristics of included studies that assessed emphysema quantitatively on chest CT.

Study	Country	Lung ca yes/no	Lung ca histologic type	age (mean±SD)	Source	Study design	Feature of evaluation	Effect size (95% CI)
Kishi 2002 [88]	USA	24/96	14 (58 %) Adenocarcinoma; 6 (25 %) Squamous cell carcinoma; 3 (13 %) Small cell carcinoma; 1 (4 %) Large cell carcinoma;	Case: 64±7 Control: 63±6	HB	Case- control; Retro- spective study %LAA _{.900} ≥5 % (di- chotomous and con- tinuous)	OR: 1.1 (0.5 – 2.4) *OR: 1.1 (0.6 – 1.9)	
Maldonado 2010 [87, 118]	USA	64/377	34 (53 %); Adenocarcinomas; 14 (22 %) Small cell carcinoma; 5 (8 %) NSCLC without specified; 2 (3 %) Large cell neuroendocrine mixed large and small cell carcinomas; 1 (2 %) Unknown;	Case: 63±7 Control: 62±6	PB	Case- control; Pro- spective study	%LAA _{.900} >5 % (di- chotomous and continuous)	OR: 1.9 (1.1 – 3.3) *OR: 1.04 (0.8 – 1.3)

Table S3.5: [continued]

Study	Country	Lung ca yes/no	Lung ca histologic type	age (mean±SD)	Source	Study design	Feature of evaluation	Effect size (95% CI)
Gierada 2011 [89]	USA	279/279	NS	Case: 63±5 Control: 61±5	PB	Case- control; Retro- spective study	% Upper lung -950HU ≥25 % (di- chotomous) semiauto- matic assessment	OR: 2.0 (1.03 – 3.8)
Aamli Gagnat 2017 [96]	Norway	34/741	NS	Overall: 59±10	PB	Cohort; Pro- spective study	%LAA. ⁹⁵⁰ ≥3 % (di- chotomous and continuous)	HR: 2.4 (0.9 – 6.2) *HR: 1.03(0.7 – 1.5)
Chubachi 2017 [95]	Japan	21/219	9 (43 %) Adenocarcinoma; 4 (19 %) Squamous cell carcinoma; 3 (14 %) Small cell carcinoma; 5 (24 %) Unknown;	Case: 73±7 Control: 73±8	HB	Cohort; Pro- spective study	%LAA. ⁹⁵⁰ >10 % (di- chotomous)	OR: 4.2 (1.0 – 29.0)

Table S3.5: [continued]

Study	Country	Lung ca yes/no	Lung ca histologic type	age (mean±SD)	Source	Study design	Feature of evaluation	Effect size (95% CI)
Mouronte- Roibas 2018 [97]	Spain	139/56	70 (41%) Adenocarcinoma; 58 (35%) Squamous cell carcinoma; 28 (16%) Small cell carcinoma; 13 (8%) Unknown;	Case: 69±9 Control: 65±10	HB	Case- control; Retro- spective study	%LAA. ⁹⁵⁰ >1% (di- chotomous)	OR: 2.2 (1.1 – 4.3)
Nishio 2019 [110]	Japan	283/293	NS	Case: 69±10 Control: 65±14	HB	Case- control; Retro- spective study	%LAA. ⁸⁸⁰ (continu- ous)	*OR: 1.01 (1.00 – 1.02)
Husebø 2019 [119]	Norway	31/681	11 (36%) Adenocarcinoma; 9 (29%) Unspecified NSCLC; 5 (16%) Squamous cell carcinoma; 5 (16%) Unspecified cancer; 1 (3%) Small-cell carcinoma;	Case: 64±7 Overall: 58±10	HB	Cohort; Pro- spective study	%LAA. ⁹⁵⁰ >10% (di- chotomous)	HR: 4.4 (1.7 – 10.8)

Table S3.5: [continued]

Study	Country	Lung ca yes/no	Lung ca histologic type	age (mean±SD)	Source	Study design	Feature of evaluation	Effect size (95% CI)
Labaki 2021 [109]	USA	353/ 6 909	NS	Overall: 62±5	PB	Cohort; Pro- spective study	%LAA. ⁹⁵⁰ (contin- ous) automatic assessment	*HR: 1.02 (1.01 – 1.03)
Individual studies assessed emphysema by visual and quantitative CT								
Schwartz 2016 [108]	USA	341/752	183 (54 %) Adenocarcinoma; 91 (27 %) Squamous cell carcinoma; 30 (9 %) Small cell carcinoma; 14 (4.5 %) Other NSCLC; 21 (6 %) Other;	Case: 64±10 Control: 62±9	PB	Case- control; Retro- spective study	Radiologist read and %LAA. ⁹⁵⁰ >4.8 % (di- chotomous)	OR (visual): 1.8 (1.5 – 2.6) OR (quanti- tative): 2.7 (1.8 – 4.0)

Table S3.5: [continued]

Study	Country	Lung ca yes/no	Lung ca histologic type	age (mean±SD)	Source	Study design	Feature of evaluation	Effect size (95% CI)
Carr 2018 [107]	USA	169/671	61 (36%) Adenocarcinoma; 17 (10%) Squamous cell carcinoma; 68 (40%) Unknown; 18 (11%) Small cell carcinoma; 3 (2%) Large cell carcinoma; 2 (1%) Others;	Case: 66±8 Control: 64±8	PB	Case- control; Pro- spective study	Radiologist based on Fleischner society guideline and %LAA. ⁹⁵⁰ per 1 % increase	OR (visual)*: 2.3 (1.4 – 3.8) *OR (quanti- tative): 1.03 (0.6 – 1.8)

HB: hospital-based; HR: hazard ratio; LAA: low attenuation area; NS: not specified; OR: odds ratio; PB: population-based; RR: risk ratio; SD: standard deviation.

*: Effect size when emphysema was assessed as a continuous variable. All the effect sizes adjusted for smoking; For specific adjusted factors, see Table S3.2 [p. 192].

Table S3.6: Overall and stratified pooled odds ratios for lung cancer given emphysema.

Stratifications	Studies (n)	Pooled OR	95% CI	I ²	P _{heterogeneity}	P _{groups}
Overall (dichotomous)	19	2.3	2.0 – 2.6	34.6 %	0.07	—
Overall (continuous)	6	1.02	1.01 – 1.02	0.0	0.9	—
Cohort study design [#]	10	2.3	2.0 – 2.7	0 %	0.46	0.33
Case-control study design [#]	9	2.2	1.8 – 2.8	55 %	0.02	
Retrospective	8	2.2	1.8 – 2.6	25.0 %	0.24	0.83
Prospective	11	2.5	2.0 – 3.1	43.5 %	0.053	
Population-based [#]	13	2.2	1.9 – 2.5	27.0 %	0.17	0.06
Hospital-based [#]	6	2.6	1.9 – 3.6	32.7 %	0.19	
Low study quality [#]	0	—	—	—	—	
Medium study quality [#]	5	2.4	1.9 – 3.0	0 %	0.76	0.30
High study quality [#]	14	2.3	1.9 – 2.7	47.1 %	0.03	
Effect sizes: HR [#]	6	2.3	1.9 – 2.9	19.3 %	0.29	0.64
Effect sizes: OR [#]	12	2.3	1.9 – 2.8	47.5 %	0.03	
Normal slice thickness (≥ 5 mm)*	2	1.5	0.9 – 2.5	12.9 %	0.28	0.30
Thin slice thickness (0.5 – 1.25 mm)*	3	2.2	1.3 – 3.7	0 %	0.70	
Cut-off value –900HU*	2	1.5	0.9 – 2.5	12.9 %	0.28	0.06
Cut-off value –950HU*	6	2.6	2.0 – 3.4	0 %	0.76	

HR: hazard ratio; OR: odds ratio. *: Only within studies assessed emphysema quantitatively; #: Within studies assessed emphysema visually or quantitatively.

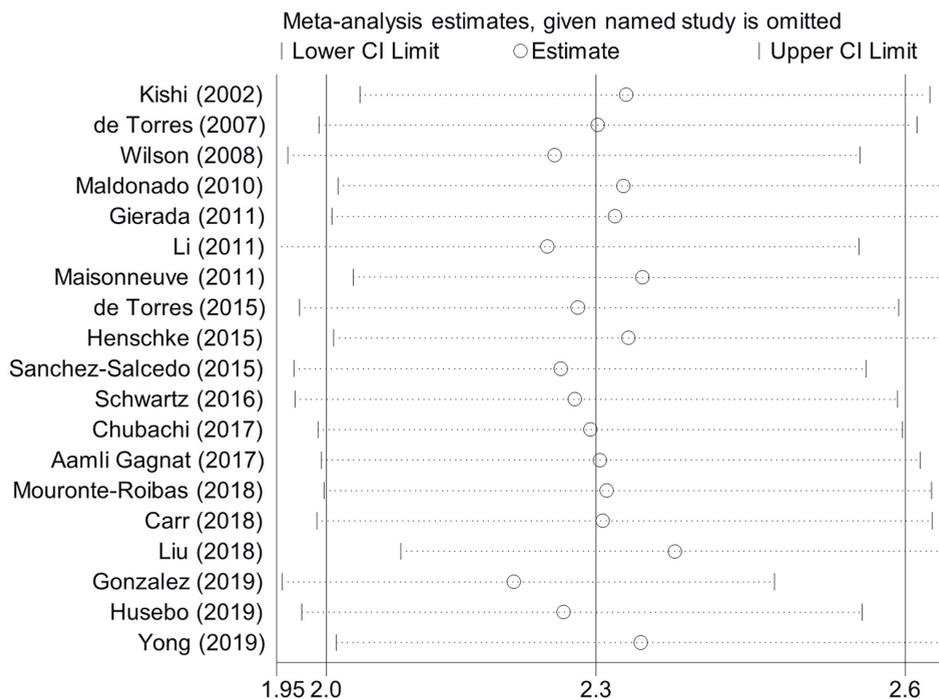


Figure S3.1: Sensitivity analysis for the overall association between emphysema (dichotomous variable, assessed visually and or quantitatively) and lung cancer within 19 studies.

Adjusted factors in these mixed effects models varied, as shown in Table S8.3 [p. 202]. Circles and horizontal lines represent the estimates and 95% CIs, respectively, for each study part.

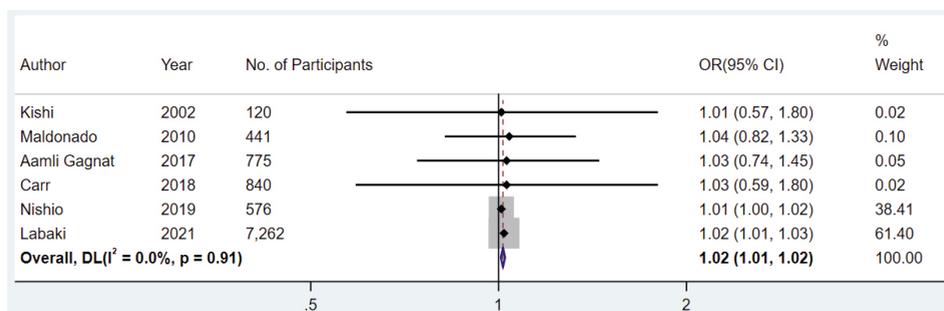


Figure S3.2: Forest plot of random-effects meta-analysis for the association between emphysema (continuous variable, assessed quantitatively only) and lung cancer. The pooled OR was 1.02 (95% CI 1.01 – 1.02; $p < 0.001$) per 1% increase in LAA. Adjusted factors in these mixed effects models varied, as shown in Table S8.3 [p. 202]. Squares and horizontal lines represent the estimates and 95% CIs, respectively, for each study part. Diamond indicates effect size and 95% CI.

DL: DerSimonian & Laird; LAA: low attenuation area; OR: odds ratio.

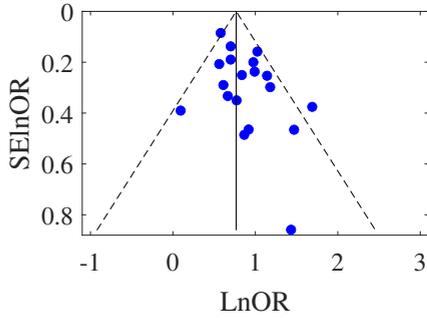


Figure S3.3: Funnel plot with pseudo 95% confidence limits to evaluate publication bias for the association between emphysema (assessed visually and or quantitatively) and lung cancer. The Y-axis shows the precision of the study (the inverse standard error), and the x-axis shows the emphysema effect. Studies with high precision will be near the average, and studies with low precision will spread evenly on both sides of average. Deviation from funnel-shaped indicates publication bias. In: natural logarithm; OR: odds ratio; SE: standard error.

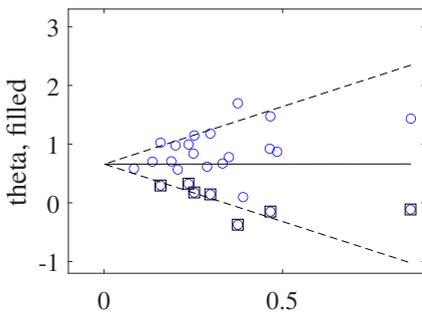


Figure S3.4: Trim and fill analysis for correction of overall publication bias in studies that evaluated the association between emphysema (assessed visually and or quantitatively) and lung cancer. Theta indicates true overall effect size.

Supplementary materials chapter 4

Table S4.1 can be found on page 217

Table S4.2 can be found on page 218

Table S4.1: Population characteristics stratified by sex. See Table 1 for abbreviation definitions. See Figure 1 for cohort definitions.

Variable	General population (cohort GP)		Healthy participants (cohort H)		Healthy never-smokers(cohort HNS)	
	Women (N=200)	Men (N=200)	Women (N=142)	Men (N=131)	Women (N=61)	Men (N=58)
Age (years)	54±5.5	54±5.4	54±5.4	53±5.5	53±5.4	53±5.5
Weight (kg)	74±12	86±10	74±12	87±11	73±13	83±11
Height (m)	1.70±0.07	1.83±0.07	1.70±0.07	1.84±0.07	1.70±0.08	1.84±0.07
BMI (kg/m ²)	25.6±4.1	25.7±2.9	25.6±3.8	25.7±2.9	25.4±4.0	24.7±3.0
Never-smokers	84 (42%)	85 (43%)	61 (43%)	58 (44%)	61 (100%)	58 (100%)
Former smokers	78 (39%)	73 (37%)	57 (40%)	51 (39%)	0 (0%)	0 (0%)
Current smokers	33 (17%)	35 (18%)	20 (14%)	17 (13%)	0 (0%)	0 (0%)
Missing	5 (3%)	7 (4%)	4 (3%)	5 (4%)	0 (0%)	0 (0%)
Pack-years	9.1±9.4	9.8±8.4	7.7±7.4	8.9±7.7	0.00	0.00
Emphysema score	4.2±3.0	7.0±4.8	4.0±3.3	6.2±4.0	3.8±3.3	5.7±4.1
No emphysema	131 (66%)	80 (40%)	98 (69%)	57 (44%)	43 (71%)	29 (50%)
Trace emphysema	68 (34%)	107 (54%)	43 (30%)	70 (53%)	18 (30%)	28 (48%)
Mild emphysema	1 (1%)	13 (7%)	1 (1%)	4 (3%)	0 (0%)	1 (2%)
FEV ₁ (L)	2.9±0.5	4.0±0.7	2.9±0.5	4.2±0.6	2.9±0.5	4.2±0.6
FVC (L)	3.9±0.7	5.4±0.8	3.8±0.6	5.3±0.8	3.8±0.7	5.4±0.7
GOLD stage	I: 20 (10%)	I: 32 (16%)	I: 0 (0%)	I: 0 (0%)	I: 0 (0%)	I: 0 (0%)
	II: 8 (4%)	II: 13 (7%)	II: 0 (0%)	II: 0 (0%)	II: 0 (0%)	II: 0 (0%)
	III: 0 (0%)	III: 1 (1%)	III: 0 (0%)	III: 0 (0%)	III: 0 (0%)	III: 0 (0%)
	IV: 0 (0%)	IV: 0 (0%)	IV: 0 (0%)	IV: 0 (0%)	IV: 0 (0%)	IV: 0 (0%)
Self-reported lung disease	19 (10%)	24 (12%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

Table S4.2: Outcome comparison.

	General population (N=200+200)	Healthy participants (N=142+131)	Healthy never-smokers (N=61+58)
Systematic bias (L)	F: 0.9 ($p = 0.388$) M: 1.4 ($p = 0.094$)	F: 1.0 (ref) M: 1.7 (ref)	F: 1.0 ($p = 0.556$) M: 1.8 ($p = 0.591$)
Δ LoA (L)	F: 3.1 ($p = 0.778$) M: 4.6 ($p = 0.259$)	F: 3.2 (ref) M: 4.2 (ref)	F: 3.4 ($p = 0.377$) M: 4.0 ($p = 0.784$)

F: women; M: men; Δ LoA: difference between the 95 % limits of agreement.

Supplementary materials chapter 6

Table S6.1 can be found on page 220

Figure S6.1: can be found on page 221

Figure S6.2: can be found on page 221

Figure S6.3: can be found on page 222



Table S6.1: Measured emphysema index for each reconstruction (N=49).

Reconstruction	mean (SD) LAV%	median (IQR) LAV%	Shapiro-Wilk test
Standard Dose	18.7 (11.9)	15.4 (9.2 – 28.7)	$p = 0.0026$
ULD FBP	20.2 (9.6)	18.6 (11.8 – 28.4)	$p = 0.0549$
ULD ADMIRE 1	19.3 (10.0)	17.3 (10.6 – 27.8)	$p = 0.0303$
ULD ADMIRE 3	17.2 (10.8)	14.4 (7.7 – 26.1)	$p = 0.0052$
ULD ADMIRE 5	13.9 (11.3)	10.1 (4.2 – 22.9)	$p = 0.0002$
ULD DLNR 1	18.8 (10.0)	16.6 (10.1 – 27.1)	$p = 0.0248$
ULD DLNR 3	15.8 (10.6)	12.8 (6.8 – 24.6)	$p = 0.0022$
ULD DLNR 5	13.2 (11.1)	9.6 (3.9 – 22.3)	$p = 0.0001$
ULD DLNR 9	9.9 (11.0)	5.3 (1.4 – 18.0)	$p < 0.0001$

LAV%: low attenuation volume percentage; IQR: 25th-75th percentile values; ULD: ultra-low-dose CT; FBP: filtered backprojection; ADMIRE: advanced modelled iterative reconstruction; DLNR: deep learning-based noise reduction.



Figure S6.1: Screenshot of the Pulmo3D software suite. Identifying interface text has been removed.

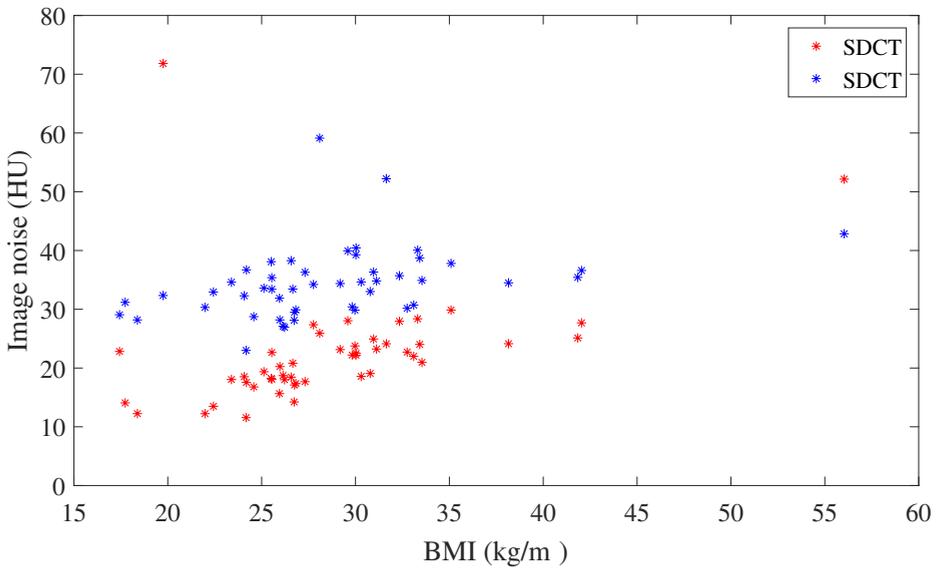


Figure S6.2: Image noise plotted against BMI. Red markers describe the standard dose CT, while blue markers describe the ultra-low dose CT. Both are FBP-reconstructions.

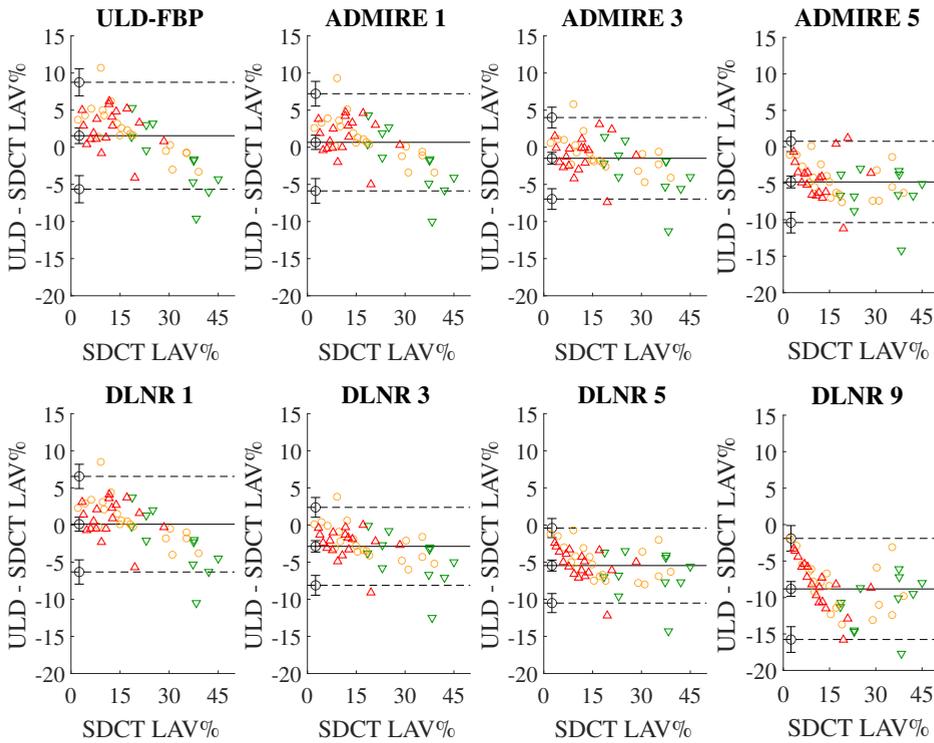


Figure S6.3: Residual plots showing the results of the Bland-Altman analyses of the LAV%, including the confidence intervals for the mean and limits of agreement. Each subplot compares a different reconstruction to SDCT.

Green downward arrows show results of normal/low BMI ($\leq 25 \text{ kg/m}^2$), orange circles for medium BMI ($25 - 30 \text{ kg/m}^2$), and red upward arrows show high BMI ($\geq 30 \text{ kg/m}^2$) cases.

The mean and limits of agreement are for the aggregated data.

Supplementary materials chapter 8

A calculation example for how to map the extended scores to the original criteria is available on page 224. Table S8.1 can be found on page 225

Table S8.2 can be found on page 225

Table S8.3 can be found on page 226

Figure S8.1: can be found on page 227

Figure S8.2: can be found on page 228

Figure S8.3: can be found on page 229

Figure S8.4: can be found on page 230

Figure S8.5: can be found on page 231



Calculation example

This is a calculation example for mapping the extended scores to Fleischner society criteria.

The original score can be inferred as follows:

First change the paraseptal scores: change 1 to 0 and change 3 to 2.

Next, change the panlobular scores: change 2 and 4 to 3.

Now we take the highest value in the table, which represents the Fleischner grade.

If the grade is required per subtype, we only take the highest value in each row.

Example 1

Presume these are the lobar scores:

No=0, trace=1, mild=2, moderate=3, severe=4, ADE=5					
	RUL	RML	RLL	LUL	LLL
CLE	1	1	1	1	1
PLE	0	0	0	0	0
PSE	1	2	2	1	1
Sum	2	3	3	2	2
Emphysema sum score = 12					

The paraseptal scores need to be changed to [0 2 2 0 0], and there is no panlobular emphysema, so the zeros remain zeros.

The inferred Fleischner grade is therefore trace CLE and mild PSE, leading to a total grade of mild emphysema.

Example 2

Presume these are the lobar scores:

No=0, trace=1, mild=2, moderate=3, severe=4, ADE=5					
	RUL	RML	RLL	LUL	LLL
CLE	2	2	0	1	0
PLE	0	0	3	0	2
PSE	1	0	1	3	0
Sum	3	2	4	4	2
Emphysema sum score = 15					

The paraseptal scores need to be changed to [0 0 0 3 0], and the panlobular to [0 0 3 0 3].

The inferred Fleischner grade is therefore mild CLE, PLE and moderate PSE, leading to a total grade of moderate emphysema.

Table S8.1: Agreement between reader and consensus decision for inclusion.

Each row in this table is a contingency table, comparing the classification of each reader (separated by cohort) to the classification determined after the consensus read.

Reader (cohort)	Trace	Reclassified as trace	Reclassified as >trace	>trace	Reclassification rate
ImaLife					
1	209	8	22	43	11 %
2	175	42	0	65	15 %
NELCIN B3					
1	86	5	2	48	5 %
2	91	0	1	49	1 %

Table S8.2: This table shows the inter-reader agreement for each subtype and lobe separately. Krippendorff’s alpha was calculated for an ordinal scale.

	CLE	PSE
RUL	0.76 (0.72 – 0.80)	0.78 (0.74 – 0.82)
RML	0.58 (0.52 – 0.63)	0.41 (0.31 – 0.52)
RLL	0.62 (0.55 – 0.69)	0.61 (0.54 – 0.68)
LUL	0.68 (0.62 – 0.73)	0.82 (0.78 – 0.85)
LLL	0.62 (0.55 – 0.69)	0.48 (0.39 – 0.57)

Values are shown as Krippendorff’s alpha (95 % confidence interval).

CLE: centrilobular emphysema; PSE: paraseptal emphysema; RUL: right upper lobe; RML: right middle lobe; RLL: right lower lobe; LUL: left upper lobe; LLL: left lower lobe

Table S8.3: This table shows the relative frequencies of the original Fleischer society criteria, separated by smoking status.
 χ^2 p-values: 0.388/0.143/0.262

	Total	ImaLife	NELCIN B3
Never-smokers			
Mild	21	8	13
Moderate	2	0	2
Severe	1	0	1
Ex-smokers			
Mild	41	32	9
Moderate	1	1	0
Severe	2	1	1
Current smokers			
Mild	39	18	21
Moderate	5	5	0
Severe	5	1	4

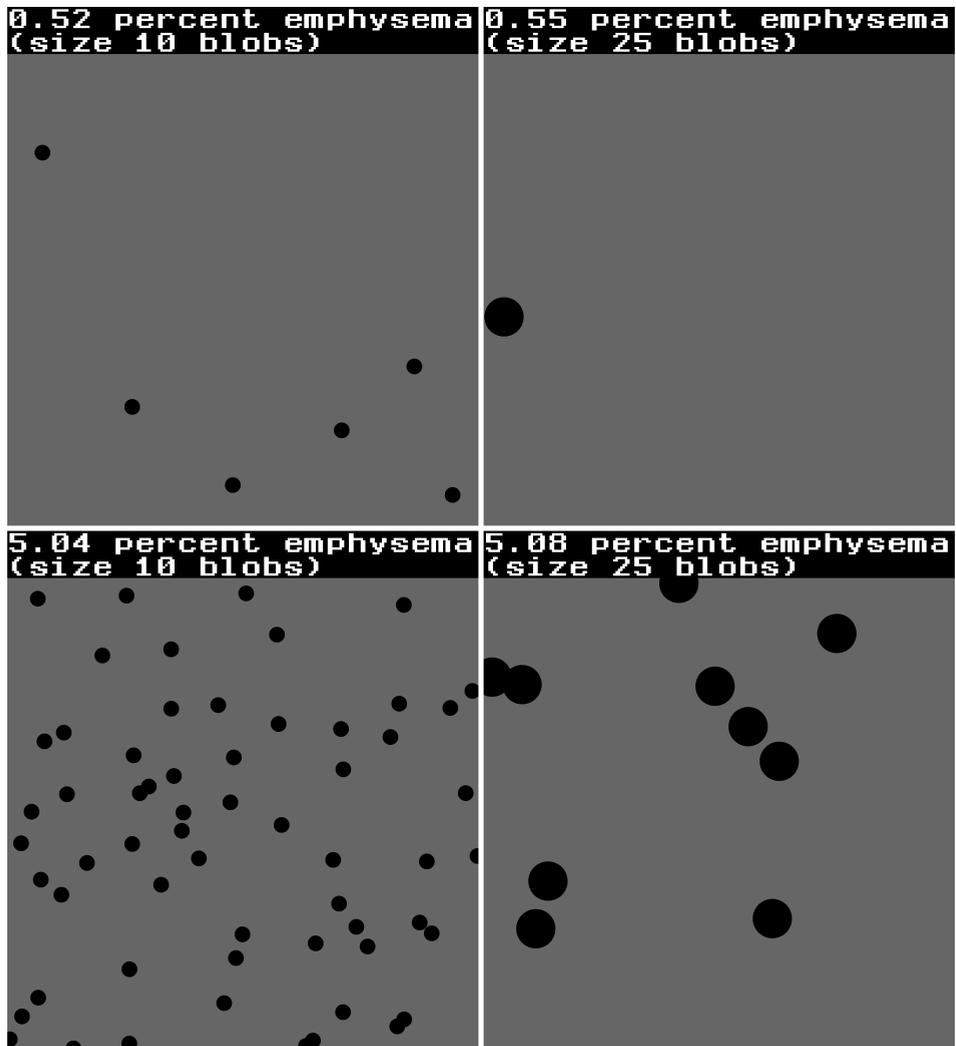
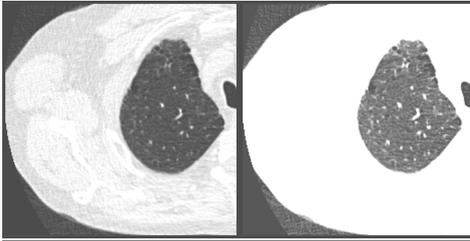
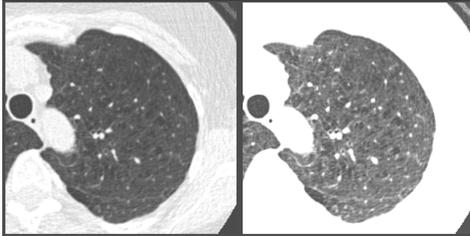


Figure S8.1: Example distributions. These examples provide a visual reference for what coverage constitutes 0.5% or 5% emphysematous area. These images were available for the readers in this study.

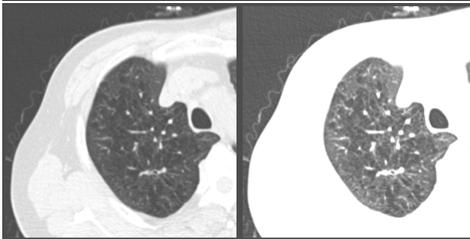
Figure S8.2: Each of the sections below contains one slice shown in two window levels (W1600L-700 (left) and W800L-900 (right), the initial settings for the readings). The images below were cropped to centre around the area of interest.



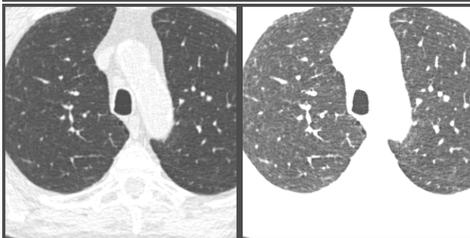
Case 1: Trace CLE and mild PSE, leading to a sum score of 3 for the right upper lobe (total sum score was 16 for this participant).



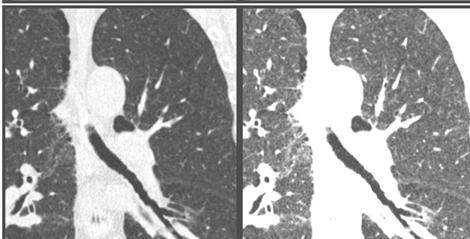
Case 2: moderate CLE, mild PSE, leading to a sum score of 5 for the left upper lobe (25 in total).



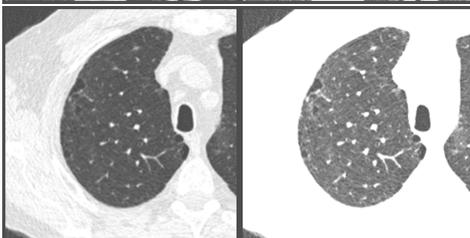
Case 3: confluent CLE, leading to a sum score of 4 for the right upper lobe (24 for this participant).



Case 4: trace PSE in the right upper lobe contributing 1 to the sum score of 4 for this lobe (16 for this participant).



Case 5: mild PSE in the left upper lobe (shown in coronal plane), contributing 2 to the sum score of 4 for this lobe (sum score was 16 for this participant).



Case 6: moderate PSE in the right upper lobe, contributing 3 to a sum score of 4 for this lobe (16 for this participant).

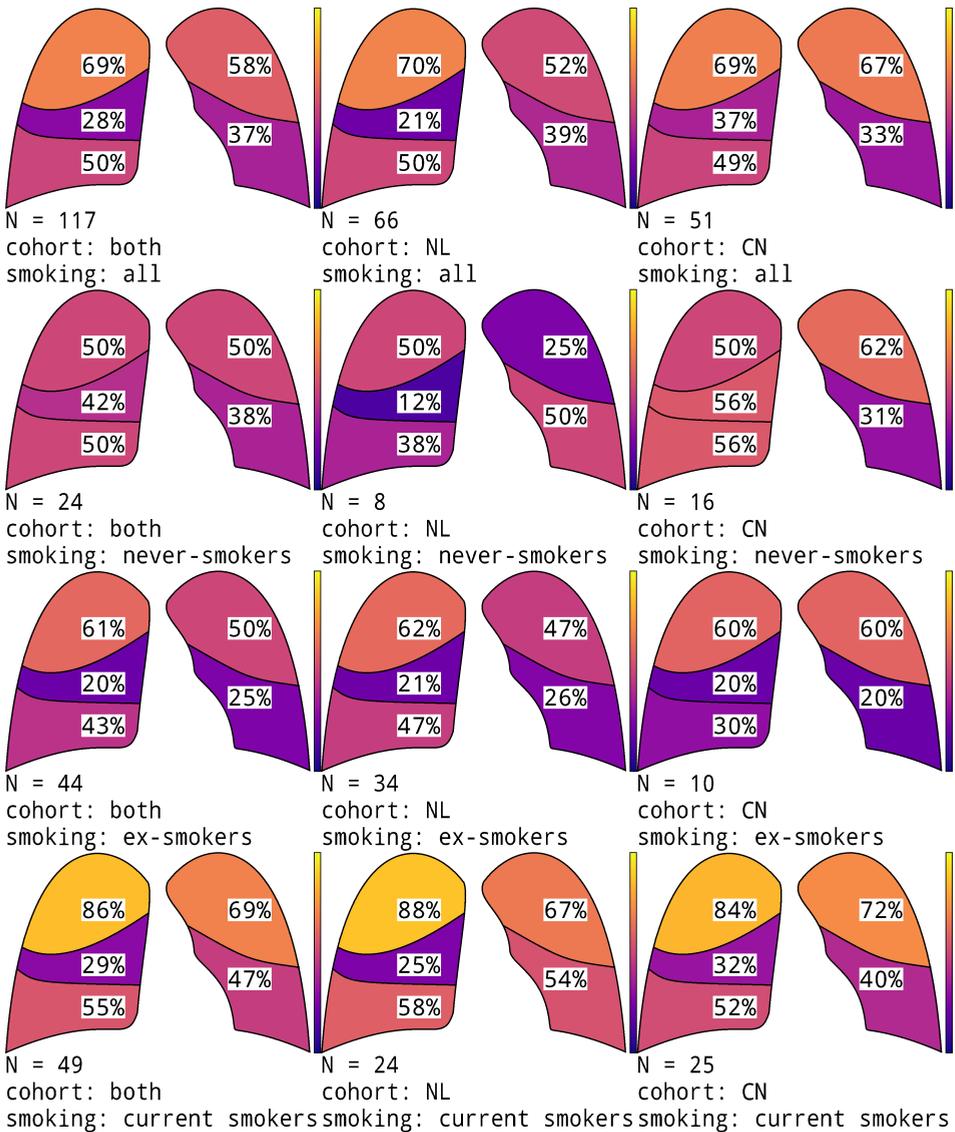


Figure S8.3: Lobar distribution diagrams.

Each percentage shows the number of participants in the given cohort having more than trace emphysema in that lobe. The number of participants is shown below each diagram, as well as the smoking status and the cohort (both cohorts, or only the Dutch ImaLife cohort, or the Chinese NELCIN B3 cohort).

The colour bar at the right edge of each diagram shows the range from 0% to 100%.

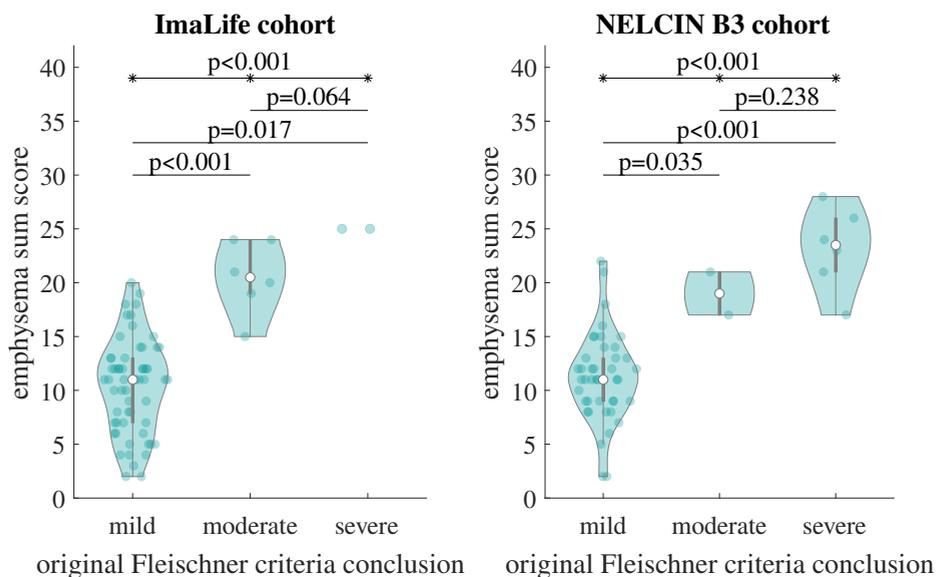


Figure S8.4: Violin plots of emphysema sum score by cohort.

The score on the y-axis is the sum of the emphysema grades of all lobes and subtypes. The category on the x-axis is the original Fleischner category. The left panel shows the results for only the ImaLife cohort, while the right panel shows the results for only the NELCIN B3 cohort.

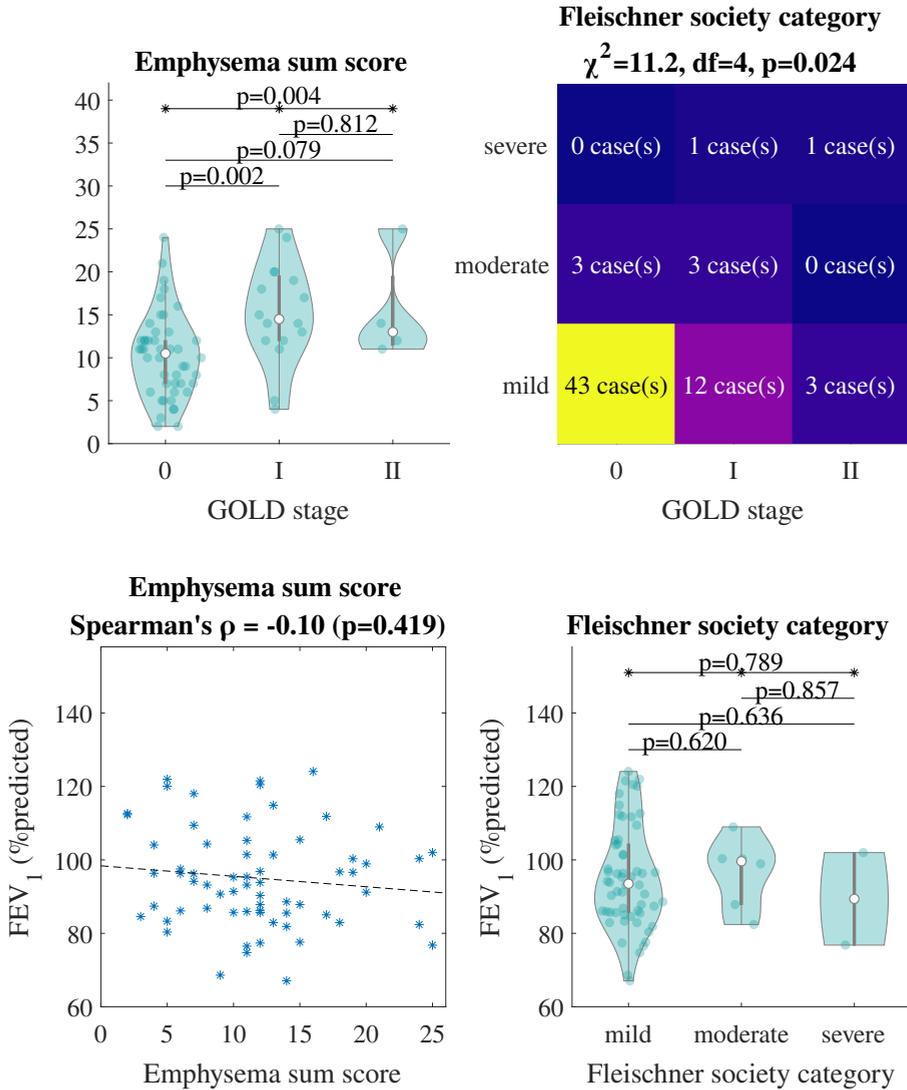


Figure S8.5: Comparison of pulmonary function with emphysema classification.

The upper left section shows violin plots of the emphysema sum score, separated by the GOLD stage, shown on the horizontal axis. The upper right section is a heatmap plot, showing the GOLD stage on the horizontal axis as well. The vertical axis shows the original Fleischner society grade.

The lower left section shows a correlation plot comparing the emphysema sum score and the FEV₁. The lower right section shows violin plots of the FEV₁, separated by the original Fleischner society grade.



Summary

The Big Three diseases (chronic obstructive pulmonary disease [COPD], cardiovascular disease, and lung cancer) are responsible for a large portion of the global mortality. In 2000 they were responsible for 36 % of deaths, and in 2019 this increased to 41 %. These diseases are strongly related, since they have shared risk factors and mostly exist in the chest.

This thesis primarily discusses pulmonary emphysema, which is a component of COPD. Emphysema is the destruction of pulmonary tissue, or more precisely, alveolar walls, which reduces the surface available for gas exchange in the lungs. Emphysema can manifest on thoracic computer tomography (CT) scan in different ways. It can be categorised as three different subtypes, each with their own aetiology and disease burden.

In **chapter 2**, 2343 participants from two general population-based cohorts, a Dutch and a Chinese one, were reviewed in order to better understand prevalence and risk factors of pulmonary emphysema identified on a chest computed tomography (CT) scan. Presence, subtype, and severity of CT-based emphysema were all recorded for each case and logistic regression analyses were performed to determine the odds of emphysema for the two cohorts. Since smoking is one of the main risk factors for the development of emphysema, the analyses considered smoking status as well. Despite a higher proportion of never-smokers, emphysema prevalence was higher in the Chinese cohort, in particular centrilobular emphysema. There was no difference in emphysema severity. The never-smokers in both cohorts shared older age and male sex as risk factors for emphysema. While emphysema prevalence was higher in the Chinese cohort compared to the Dutch, stratification showed the higher odds was only present in never-smokers. This indicates that factors other than smoking, age and sex contribute to emphysema formation. While we did not investigate which factors these might be in this chapter, genetic differences and different levels of air pollution exposure are reasonable candidates.

As mentioned, the B3 are strongly related. Not only do these diseases share risk factors, they may also be risk factors for each other. While a link between CT-defined emphysema and lung cancer was expected, the precise relation was until recently not fully understood. The emphysema assessment method can affect how much emphysema is detected, both in terms of severity and subtype. Because of this, it may be expected that different emphysema assessment methods (visual or quantitative) might influence the precise association of emphysema with lung cancer. In **chapter 3** we performed a systematic review and meta-analysis. In this chapter, we analysed 21 studies (with a total of 107 082 patients) reporting the association between lung cancer and emphysema, either assessed visually or quantified. The overall pooled odds ratio (OR) for lung cancer given the presence of visual emphysema was 2.3, and for quantified emphysema the OR was 1.02 per 1 % increase of the amount of emphysema (measured as LAV%, the percentage of lung tissue with a low density). Increased emphysema severity was associated with

higher odds of lung cancer presence. A sub-analysis showed only centrilobular emphysema (and not paraseptal emphysema) to be associated with an increased risk of lung cancer.

In general, the correct interpretation of numeric values depends on the context. This truism also holds in pulmonary medicine, where it is common to predict the total lung capacity (TLC) for a comparable healthy person based on a person's sex, age and height. This is then used to express the measured TLC as a percentage of predicted. In 2021, the Global Lung Function Initiative (GLI) published a model for use in pulmonary disease detection and monitoring. It is unknown how well the predicted lung volume corresponds with the volume as measured on CT, (the CT-derived total lung volume, TLV). In **chapter 4** we compared the GLI-2021 model predictions of total lung capacity (TLC) with CT-derived TLV. This analysis was performed with data from 142 female and 131 male healthy participants from a Dutch general population cohort. The mean \pm SD of TLV was 4.7 \pm 0.9 L in women and 6.1 \pm 1.2 L in men. The predicted TLC was 5.7 L in women and 7.8 L in men, which was a substantial overestimation compared to the measured CT-derived TLV, as the difference was 24 % (1.0 L) in women and 32 % (1.7 L) in men. In addition to this high systematic bias, there was also a high variability: the difference between 95 % limits of agreement was 3.2 L for women and 4.2 L for men. This means that in a clinical context where an accurate or precise lung volume is required, measurement of lung volume should be considered.

Since CT scanners use potentially harmful X-rays, there are on-going efforts to reduce the radiation dose. In **chapter 5** we explored the effects of five different acquisition, reconstruction, and noise reduction parameters. To objectively assess the resulting image quality, we used the COPDGene phantom and developed a quality criterion. This criterion relies on the fact that a homogeneous material will produce a range of values. For two materials to be distinguishable, these ranges must not overlap too much. A theoretical analysis showed how much overlap is allowed for the simulated lung material and the simulated emphysema material in this phantom.

Using iterative reconstruction and noise suppression software can help reduce radiation dose by 85 % while maintaining an acceptable image quality. Because an 85 % reduction in radiation dose is substantial and this finding is based on a phantom, this requires confirmation in human subjects. This replication was performed in **chapter 6**. In this chapter, forty-nine COPD patients underwent a standard clinical protocol CT (SDCT) scan as well as an ultra-low-dose CT scan (ULDCT). The median dose for ULDCT was 84 % lower than for SDCT. Bland-Altman analyses were used to determine the systematic bias and the variability between the ground truth (SDCT) and emphysema measured on ULDCT. The use of intermediate iterative reconstruction (ADMIRE 3) or noise suppression software (DLNR 3) resulted in a slight underestimation of the amount of emphysema compared to regular dose (-1.5 % and -2.9 %, respectively) and reduced the variability by 24 % and 27 %

compared to ULDCT without noise reduction. This shows state-of-the-art noise reduction techniques allow a substantial dose reduction for both phantom studies and for studies with human participants.

One important step in assessing image quality is measuring image noise. This is generally measured by calculating the standard deviation of a reasonably homogeneous circular region of interest (ROI). In **chapter 7** we explored the effects of extending the two-dimensional circular ROI to a three-dimensional spherical volume of interest (VOI). We used CT scans of forty-nine COPD patient who underwent both the reference regular dose clinical protocol CT scan (RDCT) and an ultra-low-dose CT (ULDCT). In each scan we measured the noise in the distal trachea and proximal main bronchi as the ground truth. This was then compared to the noise measured with a 1 cm^2 ROI and a VOI with the same radius (i.e., 0.75 cm^3). To simulate manual measurements, each centre point was moved 1 pixel in each direction, resulting in 27 measurements. The systematic bias of the ROI and VOI methods was similar: -1.6 HU and -0.9 HU . The variability was measured as the distance between the 95 % limits of agreement. Switching from the circular to the volumetric method reduced the variability by 40-53 %.

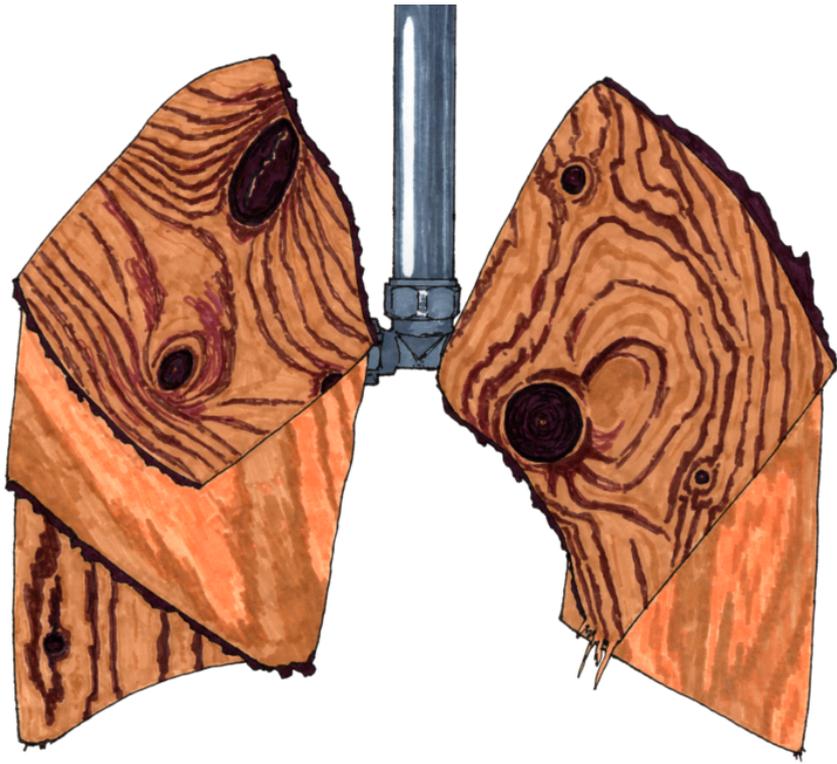
In the final chapter (**chapter 8**) we presented an expansion of the current method for visual assessment of emphysema on CT. The current method (the Fleischner criteria) provides a general overview of the presence and severity of emphysema. This might cause potentially clinically relevant differences to be missed. In the extended method, additional categories for severity were added for the paraseptal and panlobular subtypes of emphysema. This extended method was then applied to each lung lobe separately. For this study we selected 117 participants with more than trace emphysema out of 480 Dutch and 308 Chinese consecutive participants from population cohort studies. The results between readers were compared directly with the per-lobe scores, as well as with the emphysema sum score. Despite the larger number of parameters, the inter- and intra-observer variability were substantial to excellent (Krippendorff's alpha 0.78/0.85 and 0.69/0.85, inter/intra observer, and per-lobe and sum score respectively). When categorising participants with the Fleischner society criteria, the groups showed a wide range in the emphysema sum scores. While there was significant overlap between the ranges, the emphysema severity was significantly different between the different Fleischner scores. The lobar analysis allowed analysis of the patterns of emphysema distribution stratified by smoking status. This showed the emphysema was distributed homogeneously for never-smokers, but was upper-lobe predominant in current smokers. The results from this study suggest show this extended method has excellent reproducibility and captures the expert opinion with a high level of detail.

This thesis explores CT-defined emphysema in both the general and in a high-risk population. The presence as well as the severity of emphysema are risk factors for lung cancer. It is therefore important to assess lung cancer risk when emphysema is detected on CT scans, both in clinical and screening settings.

Additionally, the effects of CT acquisition and reconstruction parameters on the quantification of emphysema were discussed. This thesis shows that it is possible to lower radiation dose by as much as 85 % when image noise filtering methods are used, without compromising image quality.

Finally, this thesis presents potential improvements for two separate types of measurements. Based on this, first, this thesis advocates for the use of a simple yet effective volumetric assessment for noise measurements, which improves reliability without substantially affecting the measurement time. The second is an extension of the current method to visually assess emphysema. This extended classification system captures the expert opinion in more detail, potentially uncovering clinically relevant differences.

The work described in this thesis can be used to customise evaluation of emphysema on CT, potentially personalise lung cancer screening regimens, and to optimise clinical CT protocols to do more with less radiation dose.



Samenvatting

De Grote Drie ziektes (chronisch obstructieve longziekte [COPD], hart- en vaatziekte, en longkanker) zijn verantwoordelijk voor een groot deel van de wereldwijde sterftecijfers. In het jaar 2000 zorgden zij voor 36 % van de overlijdens, en in 2019 was dit toegenomen tot 41 %. Deze drie ziektes — die ook wel B3-ziektes worden genoemd — zijn met elkaar verweven, aangezien zij risicofactoren delen en met name in de borstkas voorkomen.

In dit proefschrift gaat het voornamelijk over emfyseem, dat een onderdeel is van COPD. Emfyseem is de afbraak van longweefsel (of preciezer: de wanden van de alveoli), wat het gaswisselend oppervlak in de longen doet afnemen. Emfyseem kan zich op verschillende manieren uiten op een computertomografie-scan (CT-scan) van de borstkas. Het is te categoriseren in drie verschillende subtypes (centrilobulair, paraseptaal, panlobulair), ieder met hun eigen wijze van ontstaan en ieder met hun eigen ziektelast.

In **hoofdstuk 2** werden de gegevens van 2 343 deelnemers van een Nederlands en een Chinees bevolkingsonderzoek onderzocht. Dit onderzoek was gericht op het voorkomen en beter begrijpen van de risicofactoren voor longemfyseem, en dan specifiek voor longemfyseem zoals dat vast te stellen is op een borstkas-CT. Een logistische-regressie-analyse maakt een model waarmee de kans op een bepaalde gebeurtenis berekend kan worden op basis van de gekozen variabelen. De aanwezigheid, het subtype en de ernst van het CT-gebaseerde emfyseem werden vastgesteld voor alle deelnemers en een logistische-regressie-analyse werd gebruikt om de kans op emfyseem voor beide cohorten te berekenen.

Aangezien roken een van de belangrijkste risicofactoren is voor het ontstaan van emfyseem, is de rookstatus ook meegenomen in de analyses. Ondanks een hoger aandeel nooit-rokers, kwam er meer emfyseem voor in het Chinese cohort, en dan met name centrilobulair emfyseem. Er was geen verschil in de ernst van het emfyseem. De nooit-rokers in beide cohorten hadden de eigenschappen hogere leeftijd en mannelijk geslacht als risicofactoren voor het voorkomen van emfyseem. Bij stratificatie wordt een analyse herhaald voor iedere subgroep om te filteren op het effect van een specifieke risicofactor. Stratificatie op rook-status toonde aan dat er alleen een toegenomen kans op emfyseem te zien was in nooit-rokers, ondanks het vaker voorkomen van emfyseem in het Chinese cohort dan in het Nederlandse cohort. Dit wijst erop dat er ook andere factoren dan rook-status, leeftijd, en geslacht van belang zijn voor het ontstaan van emfyseem. Hoewel in dit hoofdstuk geen verder onderzoek is gedaan naar wat deze factoren kunnen zijn, zijn genetische verschillen en verschillende blootstelling aan luchtvervuiling voor de hand liggende kandidaten.

Zoals eerder genoemd, zijn de B3-ziektes sterk verweven. Niet alleen delen zij risicofactoren, maar zij kunnen ook een risicofactor voor elkaar zijn. Hoewel een verband tussen CT-gediagnosticeerd emfyseem en longkanker te verwachten was, was het precieze verband tot voor kort nog niet volledig bekend. De manier

van vaststellen van emfyseem kan invloed hebben op hoeveel emfyseem gevonden wordt, zowel ernst als subtype. Het is daarom mogelijk dat de methode (visueel of kwantitatief) invloed zou kunnen hebben op het precieze verband tussen emfyseem en longkanker. Een visuele beoordeling houdt in dat een mens (meestal een radio-loog) naar een scan kijkt om een oordeel te geven over de aanwezigheid en ernst van emfyseem. Een kwantitatieve beoordeling houdt in dat een programma meet welk percentage van de longen een dichtheid onder een bepaalde drempelwaarde heeft. In **hoofdstuk 3** hebben wij een systematische review en meta-analyse uitgevoerd. In dit hoofdstuk hebben we gekeken naar 21 onderzoeken (met een totaal van 107 082 patiënten) die het verband tussen emfyseem (visueel of kwantitatief vastgesteld) en longkanker rapporteerden. Bij de analyses in dit hoofdstuk neemt de odds ratio (OR) een belangrijke plaats in. De OR is de verhouding tussen twee kansen en kan in de regel worden gebruikt om aan te geven hoeveel waarschijnlijker een optie is ten opzichte van een andere optie. Een OR boven de 1 geeft aan dat iets waarschijnlijker is dan het alternatief; een OR van 3 geeft bijvoorbeeld aan dat iets driemaal zo waarschijnlijk is. Een OR tussen 0 en 1 geeft aan dat het alternatief vaker optreedt: een OR van 0,5 geeft bijvoorbeeld dat het alternatief tweemaal zo vaak optreedt.

De samengenomen OR van longkanker, gegeven de aanwezigheid van visueel vastgesteld emfyseem, was 2,3. Voor kwantitatief vastgesteld emfyseem was dit 1,02 per 1 % toename van emfyseem (gemeten als LAV%; het percentage longweefsel met een lage dichtheid). Meer emfyseem staat dus in verband met een hogere kans op de aanwezigheid van longkanker. Een sub-analyse liet zien dat alleen centrilobulair emfyseem (en niet paraseptaal emfyseem) in verband gebracht kan worden met een grotere kans op longkanker.

In het algemeen is de juiste interpretatie van getallen afhankelijk van de context. Deze vanzelfsprekendheid geldt ook in de longgeneeskunde, waar het gebruikelijk is om gemeten waarden uit te drukken als percentage van voorspeld. Eerst wordt dus de totale longcapaciteit (TLC) gemeten bij een patiënt. Daarna wordt de TLC voorspeld voor een vergelijkbare gezonde persoon op basis van geslacht, leeftijd, en lichaamslengte. Dit wordt vervolgens gebruikt om de gemeten TLC uit te drukken als percentage van voorspeld. In 2021 heeft de Global Lung Function Initiative (GLI) een model gepubliceerd voor het gebruik in de diagnose en het door de tijd volgen van longziektes. Het was niet bekend hoe goed het voorspelde longvolume overeenkomt met het volume dat gemeten kan worden op een CT-scan (het CT-afgeleide totale longvolume, TLV). In **hoofdstuk 4** hebben we de volumina voorspeld met het GLI-2021-model vergeleken met CT-afgeleide TLV. Deze analyse is uitgevoerd met data van 142 gezonde vrouwen en 131 gezonde mannen uit een Nederlands bevolkingsonderzoekscohort. Het gemiddelde \pm standaarddeviatie van de TLV was $4,7 \pm 0,9$ L voor vrouwen en $6,1 \pm 1,2$ L voor mannen. De voorspelde TLC was een substantiële overschatting in vergelijking met de CT-afgeleide TLV: $5,7$ L voor vrouwen (24 % overschatting) en $7,8$ L voor mannen (32 % overschatting). Naast dit grote systematische verschil, was er ook een grote

variatie. De variatie wordt in een Bland-Altman-analyse uitgedrukt in de afstand tussen de 95 % grenzen van overeenstemming (ΔLoA). Dit is de bandbreedte die 95 % van de verschillen omvat, dus het verschil tussen de grootste en de kleinste (als de extremen genegeerd worden). De ΔLoA was 3,2 L voor vrouwen, en 4,2 L voor mannen. Dit betekent dat in een context waar een precies (klein systematisch verschil) of accuraat (kleine variatie) longvolume nodig is, een daadwerkelijke meting van het longvolume overwogen moet worden.

Aangezien CT-scanners gebruikmaken van mogelijk schadelijke röntgenstraling, wordt er voortdurend onderzoek gedaan naar hoe de stralingsdosis verlaagd kan worden. In **hoofdstuk 5** hebben we de effecten bestudeerd van vijf verschillende acquisitie-, reconstructie- en ruisonderdrukingsparameters (de buisspanning, de buisstroom, de reconstructiekern, iteratieve reconstructie, en ruisonderdrukingssoftware). Om tot een objectief oordeel te komen over de daaruit voortvloeiende beeldkwaliteit, hebben we het COPDGene-fantoom (een testobject) gebruikt en een kwaliteitscriterium ontwikkeld op basis van het nagemaakte longmateriaal en het nagemaakte emfysemateuze materiaal. Dit kwaliteitscriterium maakt gebruik van het feit dat een homogeen materiaal een bandbreedte aan waardes zal opleveren op een CT-scan. Om twee materialen van elkaar te kunnen onderscheiden, moeten deze bandbreedtes niet te veel overlappen. Met een analytisch onderzoek is de drempelwaarde voor de twee materialen in dit fantoom bepaald.

Het gebruik van iteratieve reconstructie en ruisonderdrukingssoftware zorgt ervoor dat de stralingsdosis 85 % lager kan worden gemaakt, zonder dat de beeldkwaliteit hierdoor ontoereikend wordt. Omdat een vermindering van 85 % substantieel is en dit onderzoek gebaseerd is op een fantoom, moet deze bevinding gestaafd worden in een onderzoek met (menselijke) proefpersonen. Dit onderzoek is uitgevoerd in **hoofdstuk 6**. Voor dit hoofdstuk ondergingen 49 COPD-patiënten een CT-scan met het reguliere klinische protocol (SDCT) en ook een CT-scan met ultra-lage stralingsdosis (ULDCT). Het emfyseem werd op iedere scan volledig automatisch kwantitatief bepaald. De mediaan van de dosis van de ULDCT was 84 % lager dan die van de SDCT. Het systemische verschil en de variabiliteit tussen het emfyseem op SDCT en op ULDCT werden bepaald met Bland-Altman-analyses. Hieruit bleek dat middelsterke iteratieve reconstructie (ADMIRE 3) of een middelsterke instelling van de ruisonderdrukingssoftware (DLNR 3) zorgen voor een kleine onderschatting van het emfyseem ten opzichte van het standaard protocol (respectievelijk $-1,5\%$ en $-2,9\%$). Daarnaast verminderen deze instellingen de variatie (de ΔLoA , zie de uitleg over hoofdstuk 4) met 24 % en 27 % ten opzichte van ULDCT zonder ruisonderdrukking. Dit toont aan dat geavanceerde technieken voor ruisonderdrukking een substantiële vermindering van de stralingsdosis mogelijk maken, zowel voor fantoomonderzoek als voor onderzoek met mensen.

Een van de stappen in het beoordelen van beeldkwaliteit is het meten van de ruis. Dit gebeurt normaliter door de standaarddeviatie te meten in een cirkelvormig gebied dat een homogene dichtheid heeft. In **hoofdstuk 7** hebben we het effect onderzocht van het uitbreiden van deze tweedimensionale ROI-methode naar een driedimensionale VOI-methode. Deze VOI-methode gebruikt een bolvormig gebied in plaats van een cirkelvormig gebied. Voor dit onderzoek zijn scans van 49 COPD-patiënten gebruikt. Ieder van hen onderging zowel een CT-scan met het gebruikelijke klinische protocol (RDCT), als een scan met ultra-lage stralingsdosis (ULDCT). In iedere scan hebben we de ruis gemeten in het laatste deel van de luchtpijp en het eerste deel van de aftakkingen van de luchtpijp. Deze ruismeting geeft de daadwerkelijke ruis weer. Deze waarden zijn vervolgens vergeleken met een 1 cm^2 ROI en met een VOI met de zelfde straal (dus $0,75 \text{ cm}^3$). Om een handmatige meting na te bootsen, is het middelpunt van de cirkel en het bolletje 1 pixel in alle richtingen verplaatst, waardoor er 27 metingen zijn per scan en per methode. De systematische verschillen (ten opzichte van de daadwerkelijke ruis) van de ROI-methode en de VOI-methode waren vergelijkbaar: $-1,6 \text{ HU}$ en $-0,9 \text{ HU}$. De variatie tussen de metingen is (net als in hoofdstuk 4) uitgerekend met de afstand tussen de 95 % grenzen van overeenstemming. Overstappen van een cirkel naar een bolletje vermindert de variatie met 40-53 %.

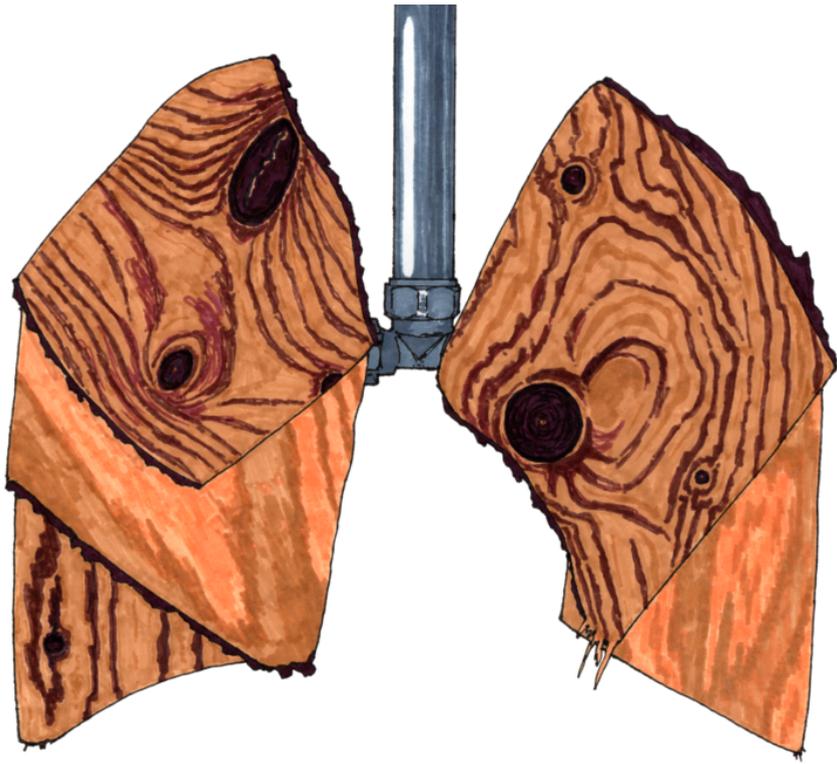
In het laatste hoofdstuk (**hoofdstuk 8**) presenteren we een uitbreiding van de huidige methode voor het visueel beoordelen van emfyseem op een CT-scan. De huidige methode (de Fleischner-criteria) geeft een algemeen beeld van de aanwezigheid en ernst van emfyseem. Dit kan ervoor zorgen dat potentieel klinisch relevante verschillen gemist worden. In deze uitgebreide methode zijn er categorieën voor ernst toegevoegd voor de paraseptale en panlobulaire subtypes van emfyseem. Dit is vervolgens toegepast op iedere afzonderlijke longkwab. Voor deze studie hebben we 117 deelnemers met meer dan sporen van emfyseem geselecteerd uit 480 Nederlandse en 308 Chinese deelnemers aan bevolkingsonderzoeken. Iedere scan is beoordeeld door twee verschillende beoordelaars. De verschillen tussen deze beoordelingen zijn zowel vergeleken per kwab, als met de emfyseem-som-score. Ondanks het grotere aantal parameters bleken de inter- en intrabeoordelaar-overeenstemming substantieel tot uitstekend (Krippendorff's α $0,78/0,85$ en $0,69/0,85$, voor inter-/intrabeoordelaar en respectievelijk per kwab en som-score). Bij het categoriseren van de deelnemers met de originele Fleischnercriteria, bleek er een grote bandbreedte aan som-scores te bestaan binnen iedere Fleischnercategorie. Hoewel er een significante overlap bestaat tussen de bandbreedtes, zijn de verschillen tussen de groepen ook significant. Een analyse per longkwab maakte een analyse van de verdeling van emfyseem binnen de longen mogelijk. Dit liet zien dat emfyseem homogeen verdeeld is voor nooit-rokers, en dat emfyseem met name in de bovenkwabben voorkomt bij huidige rokers. De resultaten van dit onderzoek suggereren dat deze methode een uitstekende reproduceerbaarheid heeft en de mening van de deskundige met veel detail vastlegt.

In dit proefschrift is er onderzoek gedaan naar emfyseem, zowel in de algemene bevolking, als in de hoog-risico-bevolking. De aanwezigheid en de ernst van emfyseem zijn risicofactoren voor longkanker. Het is daarom van belang om het longkankerrisico te overwegen als longemfyseem gevonden wordt op een CT-scan, zowel in een klinische context als bij screening.

Daarnaast zijn de effecten van acquisitie- en reconstructieparameters op de kwantificatie van emfyseem op CT-scans onderzocht. Dit proefschrift laat zien dat de stralingsdosis tot wel 85 % verlaagd kan worden wanneer ruisonderdrukking toegepast wordt, zonder dat dit ten koste gaat van de beeldkwaliteit.

Ten slotte worden in dit proefschrift mogelijke verbeterde methodes gepresenteerd voor twee verschillende metingen. Allereerst wordt in dit proefschrift een lans gebroken voor het gebruik van een eenvoudige doch effectieve volumetrische meting van het ruisniveau, dat de betrouwbaarheid vergroot zonder noemenswaardige invloed op de benodigde tijd. De tweede is een uitbreiding van de huidige methode om emfyseem visueel te beoordelen. Dit uitgebreide classificatiesysteem legt de mening van de expert in meer detail vast, waardoor mogelijk klinisch relevante verschillen te zien zijn.

De onderzoeken uit dit proefschrift kunnen worden gebruikt om de beoordeling van emfyseem specifieker te maken, waardoor longkankerscreening wellicht gepersonaliseerd kan worden en om klinische CT-protocollen te optimaliseren om meer te doen met minder stralingsdosis.



Acknowledgements

While a PhD is a solitary journey, those who make it to the end are never alone. I could of course write a page for each of you, but I will (uncharacteristically for me) keep it brief.

I am absolutely certain I will forget to mention someone very important; please forgive me.

In de afgelopen 5,5 jaar heb ik **Rozemarijn Vliegthart** als mijn begeleider mee mogen maken. Beste Rozemarijn, in die jaren heb ik erg veel van je mogen leren. Niet alleen van jou zelf, maar ook van alle mensen die je om je heen verzamelde. Hoewel onze stijlen van werken en communiceren niet altijd op elkaar aansloten, heb ik het als een groot voorrecht ervaren om bij jou te mogen promoveren. Menig promendus mag blij zijn met een paar zinnen commentaar van de promotor. Zo zit jij niet in elkaar. Iedere keer als ik een nieuwe versie had voorbereid met hulp van mijn copromotor(es), nam jij uitgebreid de tijd om gedetailleerde feedback te leveren. Het was niet altijd leuk om een regenboog aan track changes terug te krijgen, maar het werd er wel iedere ronde beter van. Het is dan ook mede dankzij jou dat ik nu een proefschrift mag verdedigen waarvan de hoofdstukken staan als een huis.

Toen ik begon bij het UMCG voor mijn masterstage, zwaaide professor Oudkerk de scepter over het CMI. Net nadat ik begon aan mijn promotie, kwam daar verandering in, en mocht ik uiteindelijk mijn PhD voortzetten met **Truuske de Bock** als tweede promotor. Truuske, ik heb je mogen leren kennen als een warm, betrokken, en enthousiast iemand. Dit was niet alleen het geval bij de reguliere begeleiding, maar trad ook op de voorgrond tijdens de reis naar Tianjin en Shanghai voor de NELCIN B3 annual meeting in november 2019. Truuske, ik wil je ook via deze weg bedanken voor je geduldige commentaar als ik weer eens te ~~technisch~~ enthousiast was in mijn methode, of als mijn inleiding te kort door de bocht was. Mede dankzij jou zijn de teksten in dit proefschrift ook voor niet-coauteurs nog te volgen.

Dan wil ik nu het woord richten tot iemand met wie ik zonder twijfel de meeste uren op Teams heb doorgebracht. Dan heb ik het natuurlijk over **Mieneke Rook**. Het zit namelijk zo, dat Mieneke eerst part-time bij het UMCG werkte (met name aan het ImaLife-onderzoek). Daar kwam op een gegeven moment verandering in, en sindsdien kon ik niet meer de gang oversteken om Mieneke om advies te vragen of om koffie te brengen. Normale mensen zeggen dan ook hun promovendi vaarwel. Gelukkig voor mij is Mieneke uit ander hout gesneden en bleef ze mijn copromotor en zetten we onze overleggen online voort. De onderwerpen die ik met je besprak op Teams liepen nogal uiteen. Van de avonturen van (of met) je kinderen, tot de fundamentele problemen van de huidige wetenschapsfinanciering, van racefietsen, tot vakanties op Ameland. Gelukkig hadden we ook tijd voor persoonlijke gesprekken en kon ik altijd bij je terecht voor een hart onder de riem. Als kers op de taart heb je mijn ook nog

inhoudelijk veel bijgebracht. Ik hoop dat ik genoeg van deze domme radioloog geleerd heb, dat ik mij met recht een domme doctor mag noemen.

Met zoveel papers in mijn proefschrift met een longgeneeskundig tintje, was het eigenlijk gek dat **Marjolein Heuvelmans** niet eerder ten tonele verscheen. Ik heb je alleen maar tijdens de eindsprint mee mogen maken, maar in die (stiekem toch niet zo) korte tijd heb ik je leren kennen als een gedreven onderzoeker. Jouw comments waren er niet alleen maar op gericht om mijn papers beter te maken, maar ook om peer-reviewers bij voorbaat het gras voor de voeten weg te maaien. Die vaardigheid is benijdenswaardig, en ik heb er alle vertrouwen in dat je die nog vaak toe zult kunnen passen.

I would like to take this opportunity to thank the members of the assessment committee, professor **Dirk-Jan Slebos**, professor **Pim de Jong**, and professor **Thomas Flohr**. Thank you for taking the time to read my thesis with such attention to detail and providing your valuable comments. I would also like to thank the other members of the corona for their time and allowing me to defend this thesis.

During the last phase of my PhD research there was much work to do, and not much time to do it in. Luckily for me, I could count on **Xiaofei Yang**. He is not just the first author of two chapters in this thesis, but he has also been a curious and intelligent colleague. Xiaofei, I want to thank you very much for all the help and guidance you have given me. I trust all current and future colleagues will enjoy working with you. I know I have.

感谢您成为我值得信赖的同事和朋友。也祝您未来前程似锦，一切顺利。

De reis die vandaag tot een einde komt, werd aan het begin begeleid door **Gert Jan Pelgrim**. Mede dankzij jouw inzet en begeleiding heb ik mijn masterstage tot een goed einde mogen brengen en kon ik vol goede moed beginnen aan mijn promotieonderzoek. Dat jij daarbij niet optrad als mijn copromotor mocht jouw inzet niet deren. Ik wil je hartelijk bedanken voor alles wat je in die jaren voor mij hebt gedaan, en ik hoop dat je in Twente vindt wat je nog miste in Groningen.

Tijdens een PhD is het belangrijk om mensen te vinden die niet alleen maar collega's zijn. **Stella Noach** is hier een uitstekend voorbeeld van. De rol van onderzoekscoördinator vormde zij om tot moederkloek van de afdeling. Stella, het was fijn om jou in de buurt te hebben, als orakel en als vertrouwenspersoon. Ik wil je hartelijk bedanken voor alle morele en praktische ondersteuning die je mij in de loop der jaren hebt geboden (en natuurlijk voor het organiseren van de pot-luck-diners).

Over the years I have had many colleagues. Many of them played an important role in my PhD, and I have learned a lot from all those smart people surrounding me. I want to thank you all for your kindness, collegiality, and friendship. A PhD is never easy, but with COVID-19 providing an extra twist, it was a great comfort to have you all around as colleagues.

In no particular order, I would like to thank:

Marleen Vonder, Monique Dorrius, Moniek Koopman, Ludo Cornelissen, Mirjam Wielema, Daiwei Han, Wout Heerink, Merel Kaatee, Yeshu Nagaraj, Congying Xia, Xiaonan Cui, Jiali Cai, Marly van Assen, Daan Ties, Randy van Dijk, Ricardo Rivas Loya, Iris Hamelink, Alessia de Biase, Luis de la O Arevalo, Sunyi Zheng, Runlei Ma, Xueping Jing, Nikos Sourlos, Jingxuan Wang, Zhenhui Nie, Nils van der Velden, and Erfan Darzi Dehkalani. Thank you for making the G2/Bloemsingel group what it was for me.

Karin Klooster, Jens Bakker, Marlies van Dijk, Maarten van den Berge, Kai Imkamp, en de andere collega's van de afdeling longziekten. Bedankt dat ik jullie mocht af luisteren om bij te leren over longgeneeskunde.

Yihui Du, Iris Stroot, Keris Poelhekken, Pato Cortés, Petra Vinke, Erick Suazo Zepeda, Melissa Castañeda Vanegas, Grigory Sidorenkov, and all other colleagues from epi. Thank you for your openness and solidarity during the coaching group meetings, and for teaching me how much I have to learn about epidemiology during the journal clubs.

Jurjen van Bolhuis, Seija Jansen, Marloes Smit, Trix Reinders, Brenda Strijk, Trynke de Jong en de andere mensen van Lifelines. Bedankt voor al jullie harde werk en al het regel. Zonder jullie inzet was het niet zo soepel verlopen. En uiteraard alle deelnemers van ImaLife, zonder wie een groot deel van dit proefschrift onmogelijk was geweest.

Jan Braaksma, Erik Ebbers, Bernadette Karssen, Robin Theunissen, en alle andere laboranten. Het was altijd fijn om bij jullie te komen buurten, en jullie inzet voor de acquisitie van ImaLife is van onschatbare waarde.

Of course my colleagues from the student attic cannot be left unmentioned. Some of you I've worked with for longer, others a bit shorter. Know that you made working (and not working) always a pleasure. I can't mention everyone, but special thanks to **Aniek Zwart, Lisa van Smoorenburg, Harriet Lancaster, Elsje Nieboer, Ilse van Bruggen, Lisanne Zwart, Abel Bregman, Herbert Kruitbosch, Ruben Hijlkema, Najod Alsabaan, and Sarah Rügger.**

A special mention for **Daan van den Oever** and **Ivan Dudurych**. They graciously agreed to be my paranympths and have had to endure many a pun over the past years. I hope to call them my friends for many years to come and share many eye-rolling moments with them. A special mention also for **Yifei Mao** (who was kind enough to correct my Chinese), and **Anke Marije Huisman** and **Annika Schoonbeek** (who helped to make my Dutch summary understandable for normal people as well).

Verder wil ik al mijn vrienden, familie, en kennissen bedanken voor de steun die ik de afgelopen jaren van jullie gehad heb. Ik heb het enorm gewaardeerd dat jullie geduldig knikten als ik weer eens iets probeerde uit te leggen over mijn onderzoeken.

Ik zal niet iedereen noemen, maar ik wil toch een kleine greep doen uit de mensen die mijn leven hebben opgevrolijkt (wederom in willekeurige volgorde): **Erwin Jongsma, Emmy Plas, Jaap Plas, Lysanne Douma, Marelien Oudman, Manon ten Have, Christiaan de Haas, Christiaan de Vries, Lotte Cruiming, Anne-Maaïke Pathuis, Elize Boerema, Judith Antonides, Peter Kloosterziel, Sabrina Schreur, Sterre Bondt-Kasper, Eline Nuninga** (a.k.a. **Eliane Haan**), **Danielle Koppenaar, Corine Sneep, Jan-Willem Dijkshoorn, Froukje Johanna Basoski, Theo Basoski, Christine Dirkse, en Erik Dirkse.**

Papa, mama, Wim, Petra, Anke, Eelco, Mark. Ons gezin is altijd een belangrijke basis voor mij geweest en jullie stonden altijd om mij heen (samen met **René, Margriet, en Jorèl**). Hoewel we een voor een uit huis gingen en ieder een eigen hoekje in het land gingen opzoeken, bleven we er toch voor elkaar. Ik hoop dat we ook de komende jaren om elkaar heen kunnen staan bij de mooie en de minder mooie momenten in onze levens. Diezelfde hoop heb ik uiteraard ook voor ‘ons volkje’: **Lucas, Hanneke, Nynke, Marlies, en Joanne** (en natuurlijk **Lorenzo en Luca**).

Mijn lieve **Lydia**, ik ben zonder jou aan mijn PhD begonnen, maar zonder jou had ik het niet af kunnen maken. Ik ben heel erg blij dat ik jou heb ontmoet. Ik had niet zo lang nodig om erachter te komen dat ik jou niet moest laten lopen, en ik ben blij dat jij je leven met mij wil delen. Mijn lieve schat, ik hoop dat wij altijd op elkaars steun kunnen blijven rekenen.