The Impact of Chronic Obstructive Pulmonary Disease in Lung Cancer and its Outcomes in High Risk Smokers

Raewyn Janice Scott

A thesis submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Medical and Health Sciences
University of Auckland

December 2021

Abstract

Background:

The existing literature suggests that Chronic Obstructive Pulmonary Disease (COPD) and lung cancer may be linked by more than smoking exposure alone. COPD is characterised by reduced expiratory flow rates and quantified as airflow limitation. The studies reported in this thesis explore the relationship between lung cancer and COPD, through several different approaches.

Null Hypothesis:

COPD and lung cancer are not related.

Aim:

To assess the relationship between COPD and lung cancer using different smoking cohorts in order to better understand if any links are evident.

Methods:

Using stratification of smokers according to the presence or absence of COPD, symptomatic airways disease or increasing severity of airflow limitation, to examine relationships and outcomes in lung cancer. This entailed cross-sectional and cohort study designs and covers both epidemiological and genetic approaches.

Results:

Using an epidemiological approach and comparing outcomes in heavy smokers, I have found the following associations-

(1) COPD severity and lung cancer diagnosed prospectively are related in a linear fashion, independent of age and pack years. (Confirmatory)

(2) While COPD is associated with an increased risk of lung cancer, it is also associated with more aggressive cancer, less surgery and more non-lung cancer related deaths in the context of computed tomography (CT) screening. (Novel)

(3) Severe COPD is related to poorer outcomes from lung cancer screening relative to mild or asymptomatic forms of COPD. (Novel)

My case-case ethnicity study found that despite comparable smoking exposure Maori develop lung cancer on average, six years younger than their Caucasian counterpart, and that this is associated with worse lung function (with loss of the expected dose-response relationship), and greater mortality, independent of stage at diagnosis.

Lastly, I confirmed our earlier studies showing that COPD and lung cancer are linked at a molecular level through the nicotine acetylcholine receptor gene variant.

Conclusion:

I conclude that COPD has a strong relationship with lung cancer, independent of smoking that includes genetic linkage, greater risk with increasing airflow limitation albeit with poorer outcomes from screening.

Dedication

To all of those who gave freely of their time and personal information for the benefit of others, in order to participate in the studies represented by the data that was interrogated and interpreted, thank you.

Acknowledgements

I would like to express my special thanks to my supervisor, Dr Robert Young who suggested I undertake this PhD. He has been actively supportive, encouraging throughout the course of my doctoral work and I am grateful for his belief in my ability. This thesis would not have been possible without your guidance. I also thank you for the numerous opportunities to present preliminary findings of this work at conferences and especially for the time given to coaching me in presentation skills.

I wish to thank my parents and siblings who have provided encouragement and support.

I wish to acknowledge the colleagues who have provided helpful critique of my work for this thesis. I cannot name all of these people but they will know who they are.

Collaboration and Sponsorship: This thesis became possible due to the collaboration forged by Dr Young with the American College of Radiology Imaging Network (ACRIN) sites of the National Lung Screening Trial (NLST), in particular Dr Denise Aberle. A Johnson and Johnson (USA), grant was awarded to Dr Young. This gave us access to the ACRIN demographic and bio-specimen data from lung cancer screening participant volunteers, forming the ACRIN-NZ sub-cohort.

Statistical techniques: I wish to thank and owe a great sense of gratitude to Mr Greg Gamble the biostatistician, University of Auckland for his invaluable advice, guidance and input on the statistical techniques used in this thesis, and also for his endless patience and belief in our wider research projects.

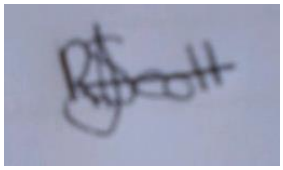
Formatting and printing: For taking my writings and formatting them I wish to thank Sue Knox. I also thank Mr Jim Gilbert, for his advice and expertise in printing the thesis.

I have found that in people's stories if you listen well enough and look hard enough you will hear and see nothing and everything at the same time. Making sense of all the information I have been fortunate to have had access to, has been interesting and enlightening. Arthur M. Schlesinger put it like this*, 'Science and technology revolutionise our lives, but memory, tradition and myth frame our response.'* Arthur M. Schlesinger

Table of Contents

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.



9 December 2021

Signature | Date

Ethics Approval Disclaimer

This thesis involves several post-hoc (secondary) analyses using data from the National Lung Screening Trial. This is through collaboration with The American College of Radiology Imaging Network (ACRIN) sites of the National Lung Screening Trial (NLST). The NLST study protocol and procedures may be accessed from the clinical trials registry and is identified by the unique study number (NCT00047385), (https://clinicaltrials.gov/ct2/home).

The detailed description of the repository samples is available on the ACRIN web site (http://www.acrin.org/TabID/145/Default.aspx). All participants from whom samples were collected signed an informed consent prior to the prospective collection of samples.

All data shared was identified by unique study ID numbers only with no personal identification information. The appropriate ethics approvals were obtained by the NLST steering committee.

Chapter 7 involved a literature search with systematic review and meta-analysis. No ethics approval was required as it used publicly available and accessible manuscripts.

The comparator study, Chapter 8, involved a retrospective clinical audit, and comparison with our previously published work and as such was not subject to ethics approval as described on the Health and Disability Ethics Committee system.

Chapter 1 Health Outcomes of Tobacco Smoking

"History, despite its wrenching pain, cannot be unlived,
but if faced with courage, need not be lived again."
-- Maya Angelou

1.1 The Burden of Smoking

1.1.1 The tobacco plant and nicotine

It is human nature to seek out pleasure. The alkaloid molecule – nicotine – from the tobacco plant, a natural insect repellent, offers a strong 'feel good' effect through rapid absorption into the blood stream via smoke inhaled into the lungs,[1] (Figure 1.1-1.3). This has particularly hazardous consequences.[2-6] Christopher Columbus is credited with introducing the tobacco plant to Europe around 1496 bringing it from the Americas, where smoking tobacco was used as a ceremonial ritual and social interaction, and so began its subsequent widespread cultivation. Despite some enthusiasm for the product not everyone approved of its use. In England, King James the first foresaw the perils issuing a warning in 1604 decrying smoking as being a "*custom loathsome to the eye, hateful to the nose, harmful to the brain and dangerous to the lungs*."[1] Centuries later his comments have proven to be very prophetic given we now know the damaging effects smoking can have for health particularly lung health. How is this possible? Some of the answer lies in our background genetic makeup influencing our responses to lifestyle choices and environmental exposures.[7,8] There is mechanistic plausibility for how smoking may act as a catalyst to set off a gene-environment collusion resulting in the damaging effects seen.[8,9] The nicotine acetylcholine receptor subunit *CHRNA5* locus found on chromosome 15q25, has receptors found in addiction centres in the brain and in the lung epithelium (Figure 1.1-1.3). The brain receptors mediate the effects of smoking exposure on addiction, while those in the lung facilitate smoking induced inflammation giving rise to chronic obstructive pulmonary disease and lung cancer, as well as other diseases.[10-15]

Figure 1.1 The tobacco plant and the molecular structure of nicotine

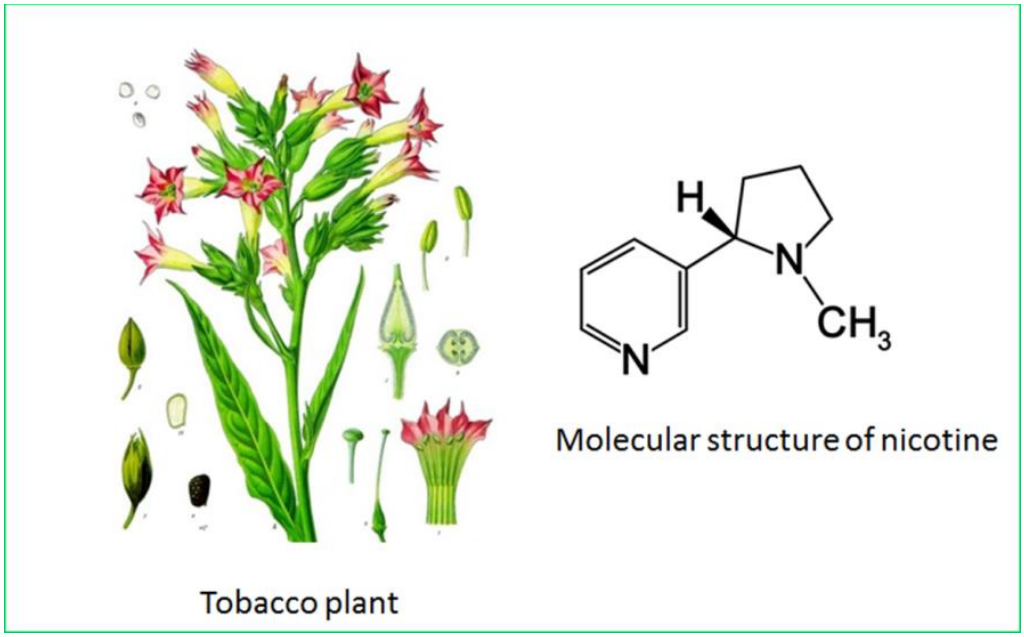
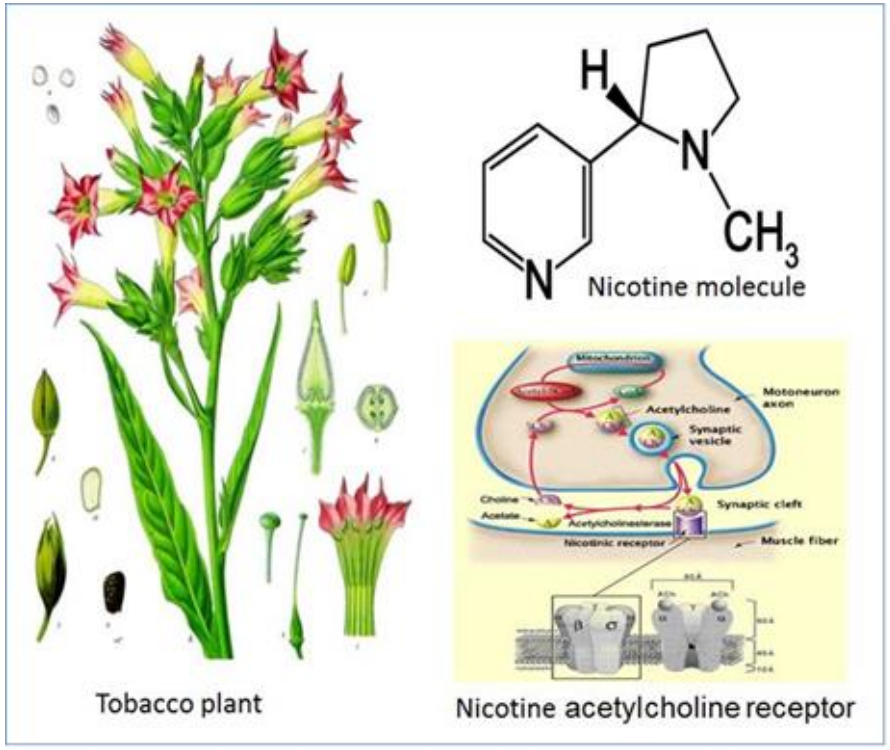


Figure 1.2 The tobacco plant, molecular structure of nicotine and nicotine acetylcholine



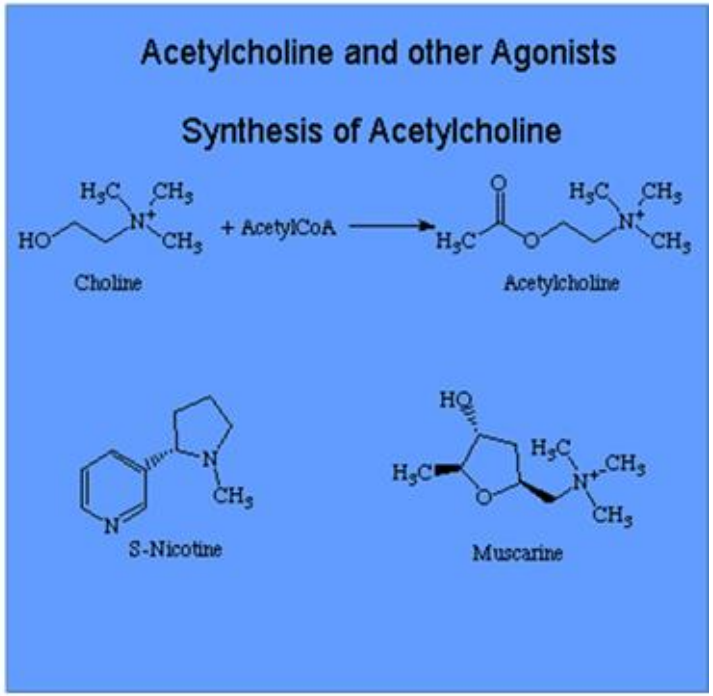
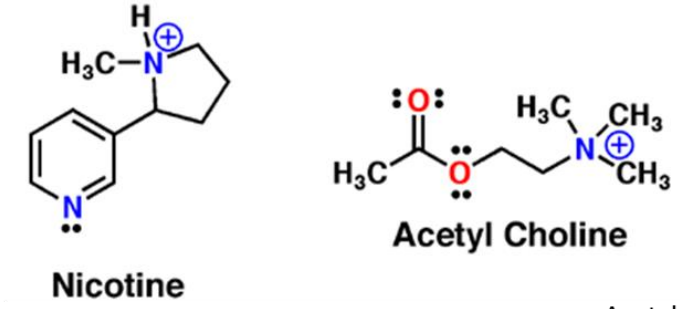
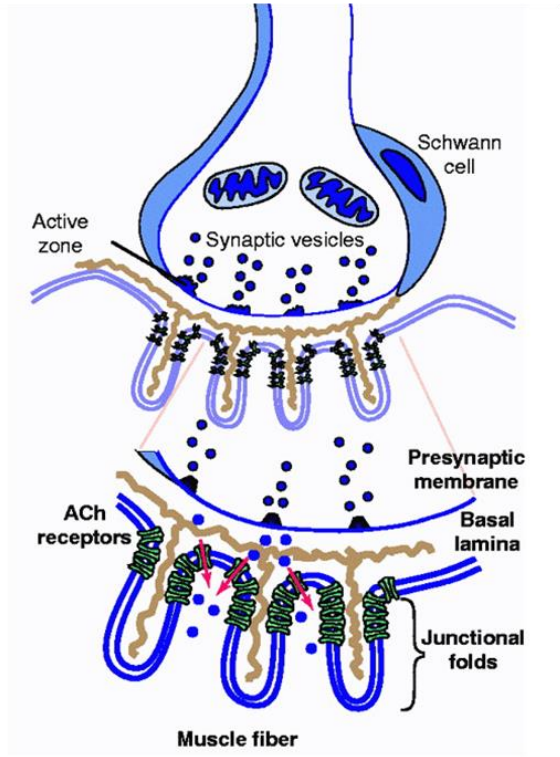


Figure 1.3 Nicotine and nicotine acetylcholine receptors





Acetylcholine is a neurotransmitter. Acetyl choline type synapses are called "cholinergic".

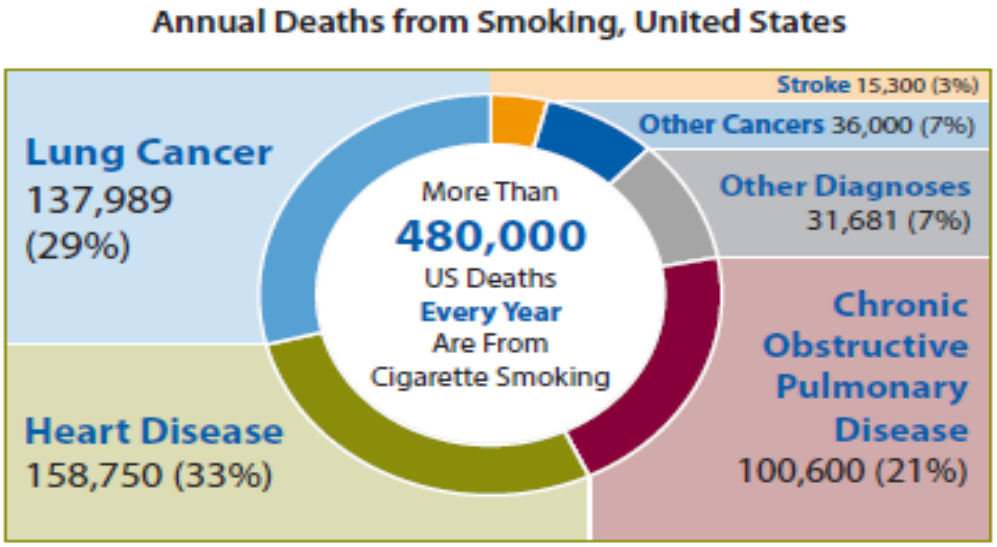
Nicotine has a similar shape and charge distribution to acetylcholine so it binds in acetylcholine receptors and causes the receptors to respond. These are called nicotinic receptors. NICOTINE THUS MODULATES THE ACTIVITY OF THESE KEY ACETYLCHOLINE RECEPTORS. A high dose of nicotine is fatal.

A low dose of nicotine, such as that obtained through smoking a cigarette has excitatory effects, e.g, it increases heart rate, simulates sensory receptors, raises blood pressure, etc.
It is addictive because this increased feeling of "excitement" actually resets the nervous system to operate at normal levels only in the presence of nicotine. Without nicotine, the person notices a difference. It is this resetting of the nervous system at the molecular level that makes nicotine addiction so difficult to overcome while a person quits smoking.

Modified from "The Developing Synapse: Construction and Modulation of Synaptic Structures and Circuits", Susana Cohen-Cory, *Science*, 2002, 298: 770-776.

Tobacco smoking is a global health concern. It contributes significantly to the costs of health care as it is most often associated with chronic morbidity and care dependency leading to many years of lost quality of life.[2-6,15-28] Those who have regularly smoked loose on average, 10 years of expected life meaning that premature mortality from smoking related diseases is a major problem which may have been prevented by avoiding exposure to both direct and second hand smoke, along with other air borne pollutants.[3-6,16-24] Early in the 1950's, the British Doctors Study, a large prospective study looking at smoking and death, identified tobacco smoking as a significant contributor to lung cancer and early mortality.[3] In addition it has been shown that the earlier a smoker stops the greater the benefit.[4] In the United States of America (US) the 2014 Surgeon General's report highlighted that Chronic Obstructive Pulmonary Disease (COPD) and Lung Cancer together made up 50% of all annual deaths from smoking (Fig1.4).[5]

Figure 1.4 Annual US deaths purported to be related to smoking[5]



By 2015 COPD had become the fourth leading cause of death globally and for upper-middle income economies was third behind ischaemic heart disease and stroke, both of which have some relation to smoking behaviour. Smoking remains a leading feature of year's lost and premature death.[16-24] Ischaemic heart disease, cancers of the trachea, bronchus and lung, and COPD ranked as the top causes of death in the US in 2016.[23] Notably COPD had climbed from the fourth leading cause of death in 1990 to be the third by 2016. In the latest US

Burden of Health report published in 2018 tobacco smoking is the number one risk factor accounting for people living with a disability.[23]

In New Zealand (NZ), tobacco use is a leading cause of avoidable disability, disease, and early death just as it is around the world.[22,25] The 'The New Zealand Health Survey' results 2016/17 reported that overall, 15.7% of adults (defined as 15 years old and over) were current smokers and as many as 26% self- identified as being an ex-smoker. This is an improvement on the 2006/7 survey when the current smoking rate was 20.1%.[25] Although Māori adults reported current smoking rates at 37% which is a downward improvement from previous surveys' they continue to smoke at greater rates than the NZ average, as do Pacific Island adults at 24%. After adjusting for age, sex and ethnic differences, neighbourhoods in the lowest socio-economic areas had the highest smoking prevalence with adults residing in these areas being 3 times more likely to be a current smoker than adults living in the least deprived neighbourhoods.[25]

Although lung cancer is the fifth most common type of cancer diagnosed each year in New Zealand it is the leading cause of cancer specific death annually, more than cancer of the prostate, colon, breast combined, with lung cancer having the worst 10-year survival. Lung cancer accounts for the most cancer deaths in men and the second most common cause of cancer death in women. For many cancers, the cancer registration rates for Māori were less than or similar to the cancer registration rates for non-Māori; however, the mortality rates were higher for Māori than for non-Māori. This suggests that Māori with cancer have a higher risk of dying from their cancer than non-Māori.[25]

New Zealand is a signatory to the 2003 World Health Organisation Framework Convention on Tobacco Control (FCTC). Efforts are put into strategies aimed at discouraging initiation of smoking including subsidised nicotine replacement treatment, quitting services available in local communities, and a National telephone based Quitline service. The hope is that smoking rates will continue to decline to negligible by 2025 although it is questionable whether this will be achieved.[25,29]

Since the 2003 WHO-FCTC treaty, growing and harvesting tobacco products in high-income countries has reduced. As a consequence, poorer countries now grow much of the world's tobacco leaf products, (Figures 1.5 and 1.6). For these places their governments view

tobacco farming and cigarette manufacturing as a way to alleviate poverty. Sadly, the health effects of smoking are growing problems and are adding to the diseases of poverty in these regions.[22]

Figure 1.5 Dried tobacco leaf coils, East Africa

Photograph by the author taken 1994 (permission of people featured was obtained).

Figure 1.6 Tobacco cigarette seller, Phenom Pen, Cambodia

Photograph by the author taken 1997, (permission of person featured was obtained).

1.2 Relationship of Lung Cancer and COPD

Lung diseases are very common and manifest most often as an acute reversible event such as a reaction to a stimulus, frequently an allergen, bacteria, or virus.[24] Some people will experience chronic lung health problems that by definition are not completely reversible. These are due mainly to repeated exposure to aero pollutants, for example tobacco smoke exposure, organic and inorganic dusts, and gases, (e.g., asbestos, radon); fumes from fossil cooking fuels and diesel.[22,24,28] How an individual reacts and responds to aero pollutants and irritants is multifactorial and complex encompassing one's biology, pathology and underlying genetic makeup. The human body is a complex system where many individual organs are required to function in synergy for optimal well-being. In this sophisticated system the largest organs have specialist functions designed to reduce harm and this is particularly the case for the airways and lungs which are most often the first line of defence against external aero pollutants and internal environmental stressors e.g., airway and systemic inflammation.[8,24,30,32]

Two complex lung health problems which are particularly disturbing for those diagnosed with them are Chronic Obstructive Pulmonary Disease and lung cancer. These two diseases have a complex relationship. The root of this complexity lies in the connection between environmental exposure (e.g. cigarette smoking), inflammation and genetic susceptibility in vulnerable individuals.[2-15,24 30-32] This connection often becomes apparent only when a person's phenotype according to severity of airflow limitation measured by spirometry lung function is known.[20,33-37] Airway limitation with or without the diagnosis of chronic obstructive lung disease is one of the commonest causes of lost life years, and contributes greatly to disability, morbidity and mortality.[16-21,38-40] Worldwide among all cancer deaths, when stratified by organ origin, lung cancer deaths contribute the most deaths.[41-45] Sex differences in lung cancer incidence and mortality have been reported for the US where lung cancer accounts for approximately one quarter of cancer deaths with decreasing incidence in men offset by increasing incidence for women.[44] With reducing smoking rates, improved early stage lung cancer diagnosis and more treatment options there is an encouraging decline in lung cancer death rates.[42] This mortality decline is somewhat greater for men, than women possibly reflecting that women started smoking later than men and have differences in smoking patterns.[6,42,44]

Lung cancer when diagnosed early enough is often treated with surgery with or without adjuvant radiotherapy/chemotherapy with improving survival rates.[42] However lung cancer diagnosis is often delayed limiting the opportunity for curative treatment.[41,42,46] Nonetheless, improving chemotherapeutics (systemic anti-cancer treatment and immunotherapy), are being developed where 5–10-year survival rates are possible.[24,42]

COPD is often misdiagnosed and under-diagnosed.[47,48] COPD comes on insidiously often going undiagnosed for many years and is characterised by airflow limitation affecting both the small and medium airways where medium airway disease is measured by a portable spirometer or body box lung function testing.[2,4,16,24,26,28] COPD is an umbrella term that incorporates chronic bronchitis, emphysema and small airways disease and the chronic airflow obstruction is not fully reversible with treatment.[2,4,26-28] Chronic tobacco smoke exposure (i.e. cigarettes smoked) is a common contributory risk factor for both lung cancer and COPD. More importantly having COPD increases a person's risk of lung cancer sometime in their future.[2-11,14-15,30,32-37,49] Both lung cancer and COPD are somewhat preventable by

curtailing the uptake of tobacco smoking and improving efforts aimed at smoking cessation particularly before the age of 45 years of age is important.[2-6,24,27-28,50,51]

COPD and lung cancer have traditionally been regarded and treated as two distinct diseases independent of each other. However it is only recently that the degree to which they overlap is coming to be fully realised.[7-11,14,30,33,34,35,51-55] This is despite this possible overlap being identified over three decades ago when it was noticed that the likelihood of being diagnosed with lung cancer increased for smokers as their airflow limitation, measured as forced expiratory volume in one second (FEV_1)/forced vital capacity (FVC), worsened.[3,7,33] Young et al, reported that COPD was an independent risk factor for lung cancer over and above smoking behaviour, age and gender.[34] The origin of this relationship is the overlapping epidemiology with shared risk factors, biology, pathology and genetic fingerprinting, with a number of genetic polymorphisms being present in the two conditions.[7-15,30,32,37,49,52-55] One possible catalyst for both lung cancer and COPD seems to be inflammation driven by inhaled particles and the substances found in smoking cigarettes. Once begun, this inflammatory process appears to continue even in some smokers who manage to stop smoking. This on-going inflammation is associated with the comorbidity common in COPD, and treatment with statin therapy may have utility.[24,30-32,54,55,58-61] There is still much work to do before all the complexities of these overlapping mechanisms can be clarified and understood, but the fact that tobacco smoke exposure (active and passive) is linked to COPD and lung cancer contributing to less lived years, is a fact that nobody should deny.[2-25,28,33,43, 51,54-58,62-67]

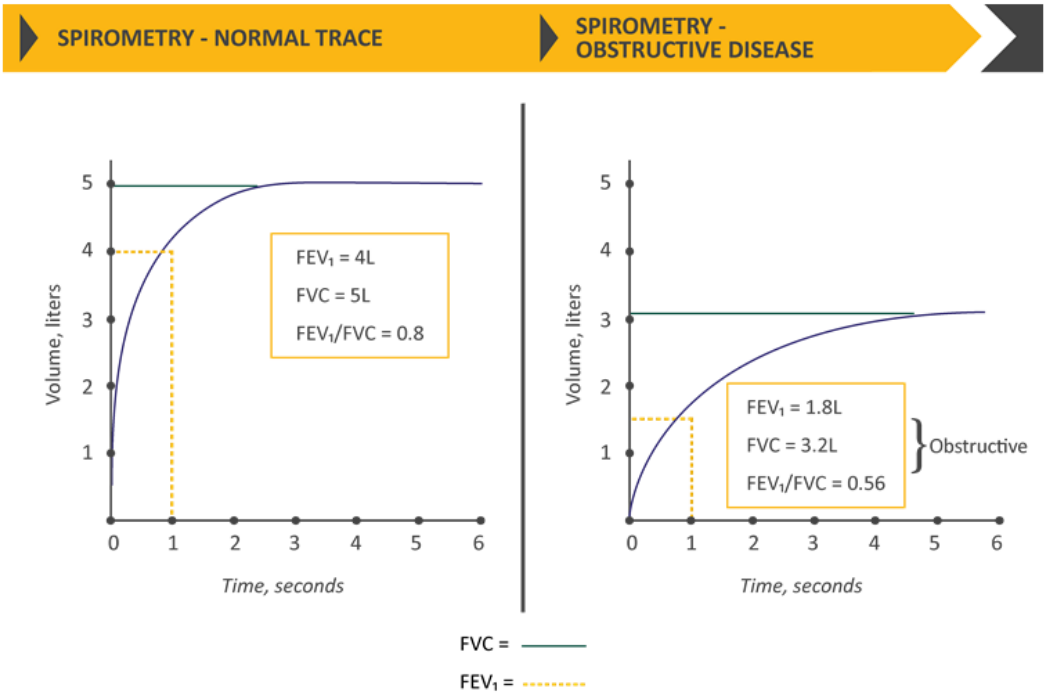
1.3 Chronic Obstructive Pulmonary Disease

Airflow limitation is the cardinal diagnostic characteristic of COPD, a complex heterogeneous disease occurring as a consequence of different pathogenic mechanisms resulting in varied clinical manifestations.[4,17-21,24,62-67] The complex heterogeneous nature of COPD is a consequence of the involvedness of different patho-genic mechanisms resulting in the varied clinical manifestations observed.[4,8,24,26-28,30,35,54-55,58-59,62-67] COPD is a debilitating and complex disease which is yet to be fully understood, but thankfully it is now receiving much more attention from both clinicians and researchers as globally we seek to unravel its mysteries.[25,50,51]

The extracellular matrix (ECM) is the support structure of both the lung parenchyma and airways which controls the cellular response to environmental stressors such as cigarette smoking. Chronic inflammation mediates extracellular matrix destruction resulting in premature apoptosis (cell death) which is a hallmark characteristic of COPD. The destruction of the extracellular matrix elastin fibres is associated with the reduced lung function seen in COPD.[30,40-42,48,49,54,58,59,68,96] Breathlessness on exertion is often the first manifestation noticed by individuals who have airflow limitation which the person often attributes to aging or to the activity/exercise being undertaken. As airflow limitation progresses in an insidious fashion much of the lung parenchyma maybe destroyed before the individual experiences any breathlessness or constraint on their activities. Over time the surreptitious progression to COPD may cause other problems such as cough, sputum production and repeated chest infections (bronchitis and pneumonia), and it is at this stage when most diagnoses are made.[4,24,26-28,62,63,65] Often times COPD is only diagnosed clinically by assessing symptoms, performance status and asking quality of life questions. Because the correlation between the objective measurement of airflow limitation and the subjective perception of symptoms is highly variable it is wise to look at the person's overall health status using validated and reliable assessment tools.[27-28,62-67] For accurate diagnosis of COPD, it is necessary to measure lung function. This process requires the participant to be actively engaged to get it right as effort is required.[4,24,26,28,47-48,62] Portable spirometers are now much more readily accessible in the community setting (primary care), simple to use yet sophisticated enough to obtain good quality acceptable results by adequately trained staff. A Forced Expiratory Volume in one second (FEV_1) to Forced Vital Capacity (FVC) ratio of less than 70% and an FEV_1 less than 80% of predicted is universally accepted as indicative of airflow limitation and COPD (Figure 1.7), (www.GOLD.org, accessed February 2, 2020).[62] The most common parameters measured when performing a spirometry test are:- Vital capacity (VC), Forced vital capacity (FVC), Forced expiratory volume (FEV_1) at timed intervals of 0.5, 1.0 (FEV_1), 2.0, and 3.0 seconds, forced expiratory flow 25–75% (FEF 25–7). The Global Initiative for Chronic Obstructive Lung Disease, recommend spirometry after inhalation of a bronchodilator medication to confirm a diagnosis of COPD. However, in large epidemiological studies the requirement for prior bronchodilator is omitted for practical purposes. Classifications for severity of airflow limitation (FVC/FEV_1 <70%), incorporating FEV_1% predicted are referred to

as GOLD grades -GOLD 1, GOLD 2, GOLD 3 and GOLD 4, (Figure 1.7).[62] Maximal voluntary ventilation (MVV), known as maximum breathing capacity is also included in spirometry.

Figure 1.7 Spirometry result showing Normal and Obstructive patterns and GOLD severity grades[62]



Fev1 as % of predicted	GOLD description	GOLD Grade
FEV1 ≥ 80% (80% +)	Mild	GOLD 1
50% ≤ FEV1 <80% (50-79%)	Moderate	GOLD 2
30% ≤ FEV1 < 50% (30-49%)	Severe	GOLD 3
FEV1< 30% (0-29%)	Very Severe	GOLD 4

Source: Adapted from the Global Initiative for Chronic Obstructive Lung Disease, global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease updated 2020. (www.GOLD.org accessed February 2, 2020)]

The severity of disease at the time of COPD diagnoses is important as it helps guide the management decisions which are most often focused on inhaler-based therapy, pulmonary rehabilitation, and preventing episodic exacerbations. As yet there are no known curative treatments for COPD but it must be stressed that avoiding tobacco smoking is the most efficacious prevention and it is never too late to stop as even those with an already reduced FEV_1 on spirometry benefit from cessation (Figure 1.8).[2-6,24,27-28,50,51] As previously mentioned

people with COPD often have multiple comorbid conditions and are at risk of premature death from all causes most commonly from a cardiovascular disease, other respiratory reasons or lung cancer, indicating that anti-inflammatory medication may be of value in this patient group.[20,24,30-32, ,60-61]

Figure 1.8 Lung function decline attributed to smoking and stopping smoking[4]

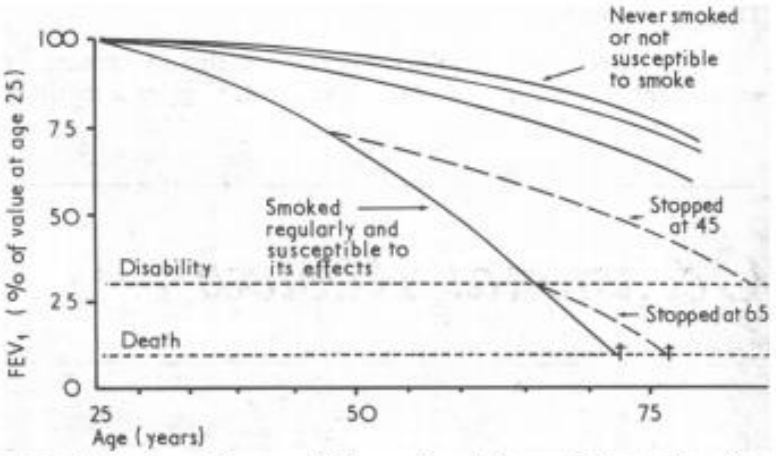


FIG 1—**Risks for various men if they smoke: differences between these lines illustrate effects that smoking, and stopping smoking, can have on FEV_1 of man who is liable to develop chronic obstructive lung disease if he smokes. †=Death, the underlying cause of which is irreversible chronic obstructive lung disease, whether the immediate cause of death is respiratory failure, pneumonia, cor pulmonale, or aggravation of other heart disease by respiratory insufficiency. Although this shows rate of loss of FEV_1 for one particular susceptible smoker, other susceptible smokers will have different rates of loss, thus reaching "disability" at different ages.**

1.4 Lung Cancer

Lung cancer develops in part as a result of deranged apoptosis where the programmed cell death rate is out of control to such a degree that cancerous change occurs.[22] Lung cancer is a complex heterogeneous disease. Many patho-biological pathways underlie lung cancer development. Lung cancer morphology referred to as histology, is divided into small cell lung cancer (SCLC) or non-small cell cancer (NSCLC) which is further sub-grouped into adenocarcinoma, squamous cell carcinoma, unspecified non-small cell, and large cell carcinoma. More than one cell combination has been reported for some tumours.[70-73] The NSCLC histological types together account for 85% of lung cancer diagnosis.[22,42,70-73] Lung cancer resulting from environmental exposures such as tobacco smoke exhibit differing behaviour at both molecular and clinical pathways ranging from slow growing indolent (bronchioalveolar) cancer to the faster growing or aggressive cancers, with different metastatic rates, treatment options and responses.[30,72,74-76] Most people diagnosed with

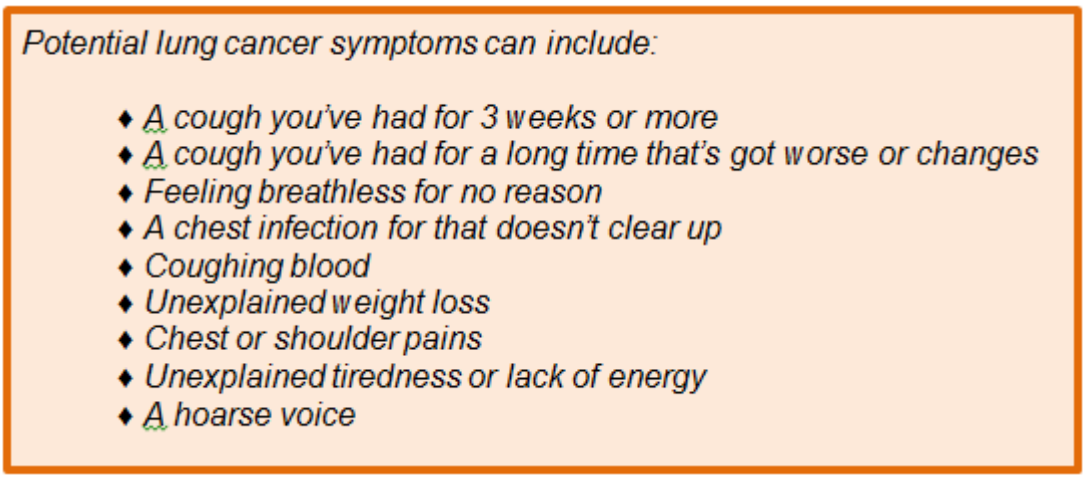
lung cancer have been exposed to air pollutants (e.g. dusts and fumes), with the majority being over 65 years old with a history of smoking tobacco products.[6,45,55] Sadly, even people who have not actively smoked themselves but who have been exposed to other people's smoking fumes also get lung cancer.[71] The vast majority of lung cancer is diagnosed late into the disease process (stages 3-4 disease), by which time it may have already spread from the lungs to neighbouring tissue and other parts of the body.[61] Furthermore concomitant cardio-pulmonary disease is common in lung cancer patients limiting the range of treatment options especially surgery.[74-80] For this reason the possibility of lung cancer when presenting with undifferentiated symptoms such as a persistent dry cough resistant to treatment, a low grade fever, weight loss, haemoptysis on a background of smoke exposure, should be considered and investigated.[34,41,42,45,46,55,77] Appropriate investigations include careful and comprehensive medical history, physical examination, computed tomography (CT) imaging of the chest and pulmonary function testing.[34,55,75,77,79] Critical for improving outcomes in smokers at risk of lung cancer is to be vigilant in identifying warning signs when other smoking-related lung diseases are present such as COPD. COPD increases the risk of lung cancer 2-4-fold and is present in 50-70% of lung cancer cases.[37,42,45,55,74-76,80]

The long-term survival of patients after a diagnosis of lung cancer remains woefully low and is the lowest of many cancers.[42-46,74] This is despite the advances in treatment options including the newer chemotherapy treatments as well as immunotherapy agents based on the biological/genetic mutations of the individual tumours (i.e., tumour specific treatment). This includes antibodies targeting the programmed death pathway (PD-1) which are showing some benefit in NSCLC when used as single or adjunct therapy.[45,74,76] For NSCLC survival time is better than for small cell with five-year survival rates up to 70% for stage I disease with curative intent surgical removal and adjuvant therapy options. Sadly, for small cell lung cancer histology up to 70% have extensive disease at presentation diagnosis with the 5-year survival rate as low as 2%-6%. For those diagnosed with limited stage disease, SCLC 5-year survival can be 12%-15% even with a combination of chemotherapy and chest radiation.[79]

Regardless of histology and treatment options the key to better survival outcomes for all lung cancer is earlier diagnosis when the cancer is at a stage where surgery and other treatment could be most effective.[74-77] To this end the Scottish National Health Service and Healthier Scotland, on-going campaign[81] 'Detect Cancer Early 'designed to raise public

awareness of the signs and symptoms of lung cancer (Figure 1.9), plus initiatives such as the Manchester 'car-park' CT scan mobile lung health assessment clinics,[82] have had some success in facilitating earlier presentation, or diagnoses and better treatment outcomes.

Figure 1.9 Scotland National Lung Cancer Awareness Campaign November 2016[81]



In conjunction with any population based educational or detection programmes, the establishment of rapid referral clinics and fast track treatment protocols are critical. It is important not to relax preventative measures such as reducing tobacco smoking rates and improving indoor and outdoor air quality through pollution control measures. General information that is important for improving public understanding regarding a person's potential risk for lung cancer include the following five points:

1. An individual's vulnerability to lung cancer combines personal genetic and non-genetic (clinical/lifestyle) data, together making up their personal risk;
2. Those who have ever smoked have an overall 10% lifetime risk of lung cancer;
3. The risk of lung cancer increases with age;
4. The risk of lung cancer increases in the presence of airflow limitation (COPD);
5. The risk of lung cancer decreases the longer a smoker has stopped.

Risk assessment models and tools have been developed to help clinicians and individuals calculate the likelihood of developing a chronic disease or cancer sometime in the future.[49,83-86] An assumption behind these tools is that individuals will subjectively perceive their risk to be the same as the level calculated by the assessment tool.[87] Calculated results should be presented in formats that are believable, easy to understand and encourages

positive behaviour changes e.g. smoking cessation, nutrient rich diet and adherence to screening protocols.[45,49,50,88-89]

1.4.1 Advances in screening radiographic technologies and risk prediction models

The advantage of computed tomography (CT) scan over conventional chest radiography (CXR) is that CT enables lung nodules of much smaller sizes to be identified, potentially improving a person's curative prospects.[41,45,82,90-93] It must be remembered that although the sensitivity of low-dose CT (LDCT) versus CXR screening is high, the specificity of LDCT screening is suboptimal due to the many false positive findings. However, work is being done to improve on this. The development and refinement of radiographic technologies underpinning computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI) will enable more precise characterisation of lung abnormalities so that distinguishing benign, indolent, or aggressive nodules should facilitate earlier identification of lung nodules likely to be cancerous and life threatening.[91-93]

CT technology provides an opportunity for mass screening programmes of at-risk populations such as those enrolled into the large North American National Lung Screening Trial (NLST).[90] Consideration needs to be given to the benefits versus harms ratio for individuals being referred for lung LDTC screening.[36,94-100] Due to the sensitivity of CT technology, nodules of benign or indolent type will be picked up along with the more sinister malignant cancer nodules. The intermediate nodules of uncertain significance usually require follow up with further CT scans, biopsy, or surgery to confirm the diagnosis where these investigations themselves may cause unintentional harm to the individual.[97-103] In response to the NLST[90] result which showed a 16-20% reduction in lung cancer specific mortality with CT screening over conventional CXR many societies and specialist groups called for a national screening programme to be approved. In particular it prompted the US Preventive Task Force (USPTF) and other relevant Societies to review lung cancer screening recommendations including when not to screen.[77,94,-10] Most of the final recommendations for yearly LDCT screening[94,104-106,108] are consistent with NLST eligibility criteria.[90] These are people aged between 55 and 75 years old; current or ex-smokers stopped within 15years; with a smoking history equivalent to 30 pack years where a pack year is equal to smoking one packet of 20 cigarettes a day for one year, and no symptoms or signs of lung cancer.

Screening should be stopped or not considered in those who have reached 81 years old, have stopped smoking for more than 15 years, or have a health problem making them an unsuitable surgical candidate if a lung cancer were to be found.[95,100,105] Subsequently lung cancer risk models are being developed and compared to refine recruitment in an attempt to identify better screening criteria.[86,91,95-98,100,103,107-114] Regardless of these recommendations an informed consent and shared decision-making consultation visit with the screening candidate is a mandatory requirement of lung cancer CT screening programmes.[103,110,113,115]

The National Lung Screening Trial, [90] was not a community-based study as would be the case in a nation-wide screening programme. This is important as the outcomes of the NLST may not be consistently reproducible in a community-based screening setting.[82,88,91] In particular for older smokers who have more comorbid disease, the screening process itself may carry unintended risks and the treatment options for those at greatest risk are limited or compromised.[55,79,91,95,96,98,99-103,108,112,114]

The lung cancer risk prediction tools/models that have been developed are based primarily on age and smoking history. While they guide those smokers and ex-smokers considered to be most at risk for lung cancer, they do not discriminate between risk of lung cancer and who may not benefit from screening.[33-34,37,87,95-103,107-108,110,113,116-118] Depending on a person's comorbid (competing risk) status and mortality risk profile, for high risk individuals the outcomes from screening may be attenuated due to death from non-lung cancer causes.[55,100-102,113,116-123] Further, comorbid disease contributes to loss of benefit of screening due to the reduced likelihood of being offered surgery or surviving surgery if a cancer is diagnosed.[75,78,80,93,100-103,113,116-120,123-126] Cost effective analyses using risk-based modelling have shown an unequal benefit across risk levels.[86,90,97,101,122,123] Refinements to these cost effectiveness studies that take into account pre-existing disease are important as high risk smokers also have a high burden of comorbidity such as cardiac disease, COPD or emphysema that will impact screening outcomes.[5-7,15,28,33-38,40,49,55,80,99-103,108,113,116-121,124,126-132] It may be more cost effective to use an outcomes based approach for identifying those who will gain the most benefit from screening as studies suggest smokers of intermediate risk, the middle 60% of risk in risk quintiles 2-4, may gain the most benefit from screening.[102,103,113,117,118,133]

When considering the introduction of a population level lung cancer screening programme there are certain characteristics or elements that are considered to be 'must haves' for the programme to be successful.[110,134] First and foremost is that people coming forward for screening need to be told that screening itself will not prevent the development of cancer, rather the aim is diagnosis of a cancer at a stage when curative treatment may be possible although even this is not a certainty.[41,46,80,95,100,115] Lung cancer and COPD are important intertwined health consequences from smoking which causes reduced quality of life and premature death. There are shared risk factors, overlapping pathways with regards to pathology, and underlying genetic susceptibility for lung cancer, where COPD is a major interdependent risk factor.[34-37,55,56,101-103,113,135] Early diagnosis may alter the outcome trajectory for these conditions. The World Health Organisation (WHO) has specified elements that they consider as being fundamental to any screening initiative.[108] Table 1.1 sets out these principles in relation to lung cancer screening. Case identification using LDCT screening for lung cancer and spirometry for COPD are examples of technologies that might advance these screening approaches. All screening programmes have the potential to preserve health and delay death when the most appropriate people are targeted for screening.[11,94-103,113-114,118] In this way the WHO criteria may need to be adapted in some areas as new knowledge on disease screening outcomes and treatments become available.

Table 1.1 Screening Programme Principles

Criteria	**Comment in relation to lung cancer screening (case-finding) programmes.**
The condition should be an important health problem.	Lung cancer is a major cause of cancer death accounting for 20-25% of all cancer deaths.
The natural history of the disease should be adequately understood	There is incomplete understanding of the natural history of lung cancer
There should be a latent somatic stage or asymptomatic stage for the disease	Lung cancer is asymptomatic in its earliest stages.
There should be a test for the condition	CT screening is appropriate for lung cancer detection.
The test should be acceptable to the population	CT scans of the chest are widely used.
Test used should be sensitive	CT is overly sensitive so 95% of nodules identified are benign.
There should be a treatment for the condition.	Surgical removal is the superior treatment with or without adjuvant chemotherapy/ radiotherapy.
Facilities for diagnosis and treatment should be available	Yes, these are available.
There should be an agreed policy on who to treat	NLST and USPSTF criteria include smokers of a certain age with a specified pack year history.
The total cost of finding a case should be economically balanced in relation to medical expenditure as a whole.	In lung cancer this continues to be an area of debate, work being done on cost of targeted screening and cost-to-benefit analysis.
Case-finding should be a continuous process, not just a 'once and for all' project	Risk profiles should be consistent for case-finding programmes. There are current programmes in the UK, Canada, Australia, Europe, and the USA.

1.4.2 Outcomes in screening for lung cancer

An outcome is the consequence of an action. In health, positive outcomes are the key aim of any activity whether the focus of the activity is education, prevention, surveillance, care, research, or treatment. Achieving maximal beneficial outcomes requires the appropriate matching of risk profiles to available technologies/ treatments specific and sensitive enough to improve both quality of life and life span.[96,97,100,113,117,118] That said, to be successful any health care initiative must be acceptable to all parties so that they engage with it.[49,50,81,82,88-90,115,136-141] Lung cancer and COPD are two important health problems that rely on good

cooperation and trust between patient and provider for meaningful outcomes to be achieved. Screening programmes and enrolment in them are examples of how the afore mentioned can be realised through early case identification and prompt treatment.[47,48,53,55,77,80,85-86,88-91] It is also important to have a supportive management plan that has strategies to support both those screened and the health professions caring for them. Things to consider should include being transparent in giving patients the results of all tests and investigations undertaken and provide information to help them understand the implications of such results. It is important to talk about and understand what a screening participant may wish to prioritise going forward.[103,142] Although not acknowledged, when introducing and maintaining a screening programme the support of clinicians is vital at all stages especially adequate communication regarding referral pathways. On-going assessment and encouragement are important for all involved in screening programmes where psychosocial distress is a very real occurrence in all phases of cancer from the pre-diagnostic and diagnostic stages to management and treatment decisions through to the provision of palliative care.[81-82,90-91,102-106,110,113-114,134,143-144]

1.5 Ethnic Ancestry and Genetic Background

1.5.1 Ethnicity

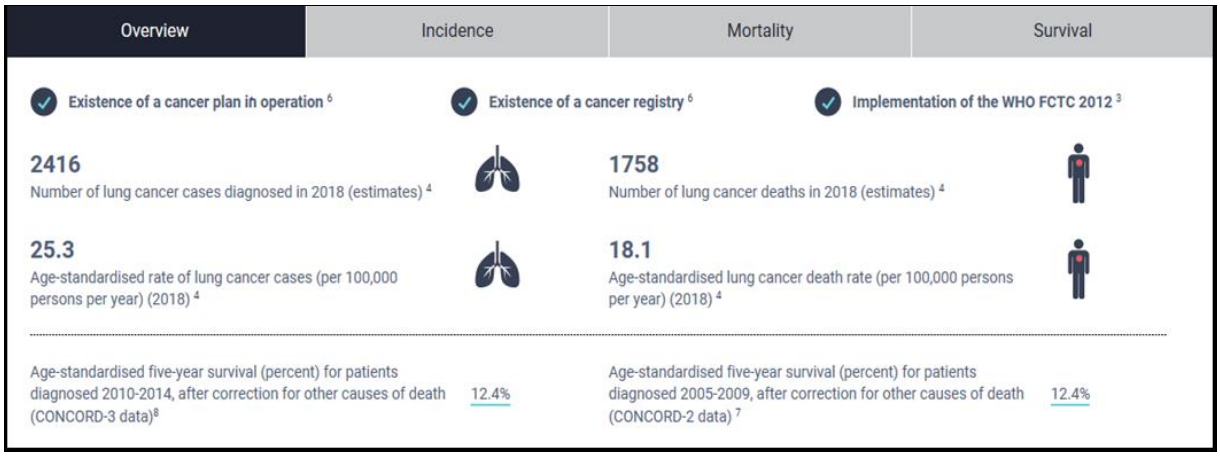
In epidemiological studies where environmental exposure, cultural practices and genetic susceptibility interact to determine health outcomes, it is necessary to also consider ethnicity (ancestral background) in this context.[145-147] Ethnicity and genetic inheritance can inform responses to environmental and lifestyle exposures.[146-151] It has been found that in families with shared smoking histories, susceptibility to COPD is related to an increased risk of lung cancer.[7,150-152] The estimated inheritance of COPD is 50-70% and for lung cancer it is 20-30%,[20,36] this raises the possibility that genetic susceptibility to COPD may be a mediator for susceptibility to lung cancer.[7-12,14-15,20,33-36,49,52-56,93] We have published earlier work suggesting there exist genetic loci that confer risk of both diseases.[10,11,15,35,49,85,153-154] This observation is tested in this thesis.

There have been some important susceptibility differences observed with regards to the effects of smoking in relation to lung function, COPD diagnoses and lung cancer for some ethnic groups.[146,152,155-157] In ever smokers, a study comparing five ethnic groups in the

United States (US) found lung cancer incident rates were highest in African American and those with Polynesian Hawaiian background which remained even after adjusting for age.[146] This study also found a dose-response relationship for smoking and lung cancer risk for other ethnic groups that was not seen in African-Americans and Pacific Islanders. Interestingly, Haiman et al[146] saw that in never smokers, the risk for lung cancer was lowest in African Americans indicating that smoking effects vary with different ancestral/ethnic populations. This has implications in the New Zealand context where the Māori population who share genetic ancestry with Polynesians originating from the Pacific Islands, have lung cancer rates that are 2-4-fold higher than New Zealand Caucasians.[25,157-160]

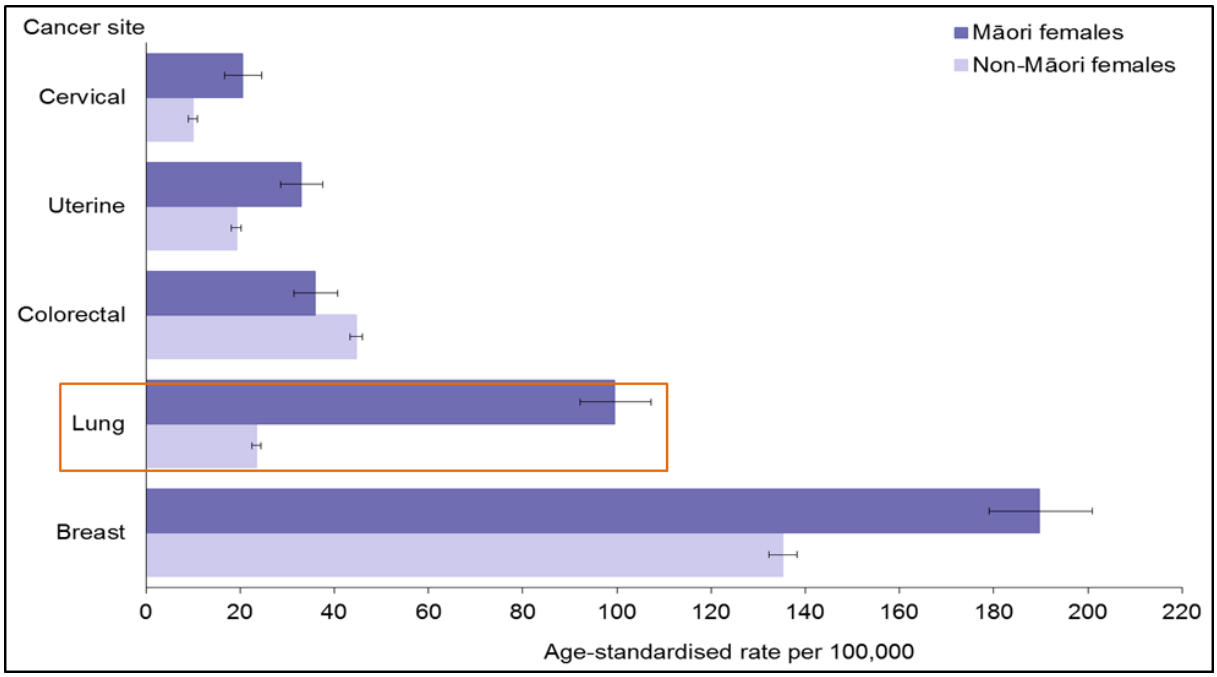
In New Zealand tobacco smoking has been an accepted practice for many decades. Smoking rates differ by area of residency and ethnicity. According to the latest available health survey data the national average for those reporting current smoking is 15.7%, however 37% of Māori adults were current smokers.[25] This is important as lung cancer has the highest annual mortality rate among all cancer deaths in NZ, with the five-year survival rate of 12.4% unchanged between 2005-2014 (Figure 1.10).[25,160] Smoking rates among Māori females is high relative to that reported in the general NZ population. The 2010-12 health statistics revealed that for lung cancer registration and mortality the rates for Māori females were more than 4-times that of non-Māori females and the rates for Māori men were almost 3-times higher than for non-Māori males (Figures 1.11,1.12 and Figures 1.13,1.14).

Figure 1.10 Lung cancer diagnosis and deaths in NZ



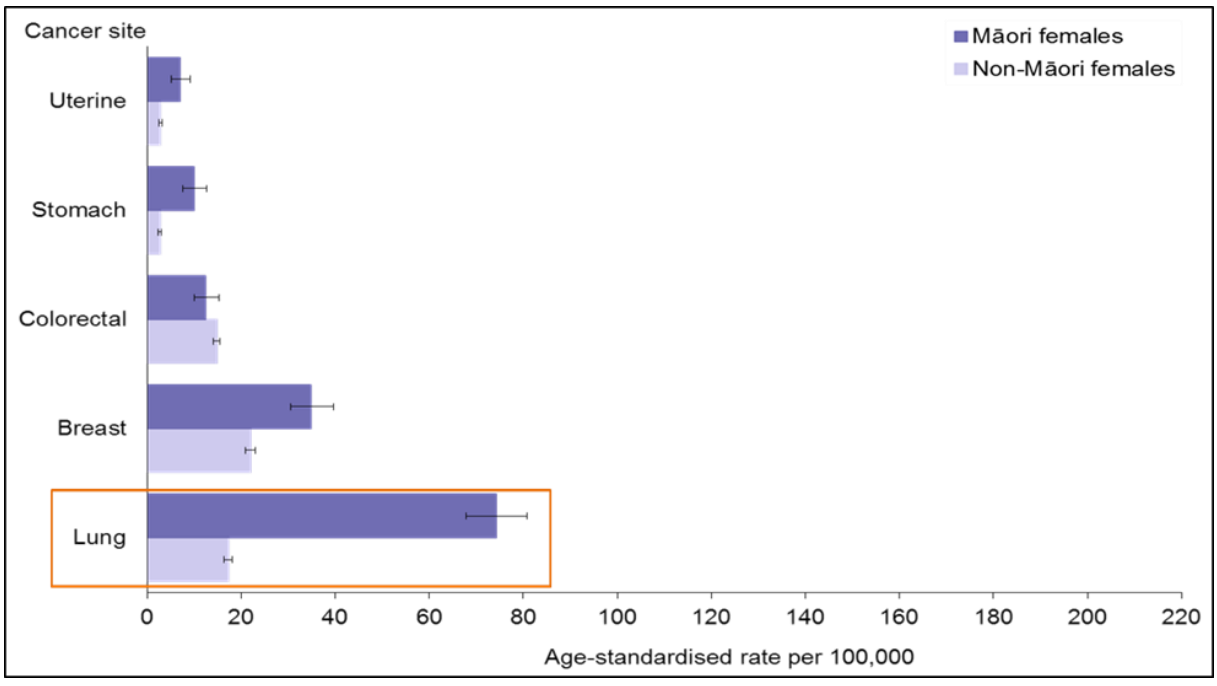
Source: Global Lung Cancer Atlas http://www.lungcancercoalition.org/e-atlas/

Figure 1.11 NZ Females (25+years) cancer registration rates per cancer site, Māori, and non-Māori 2010-12



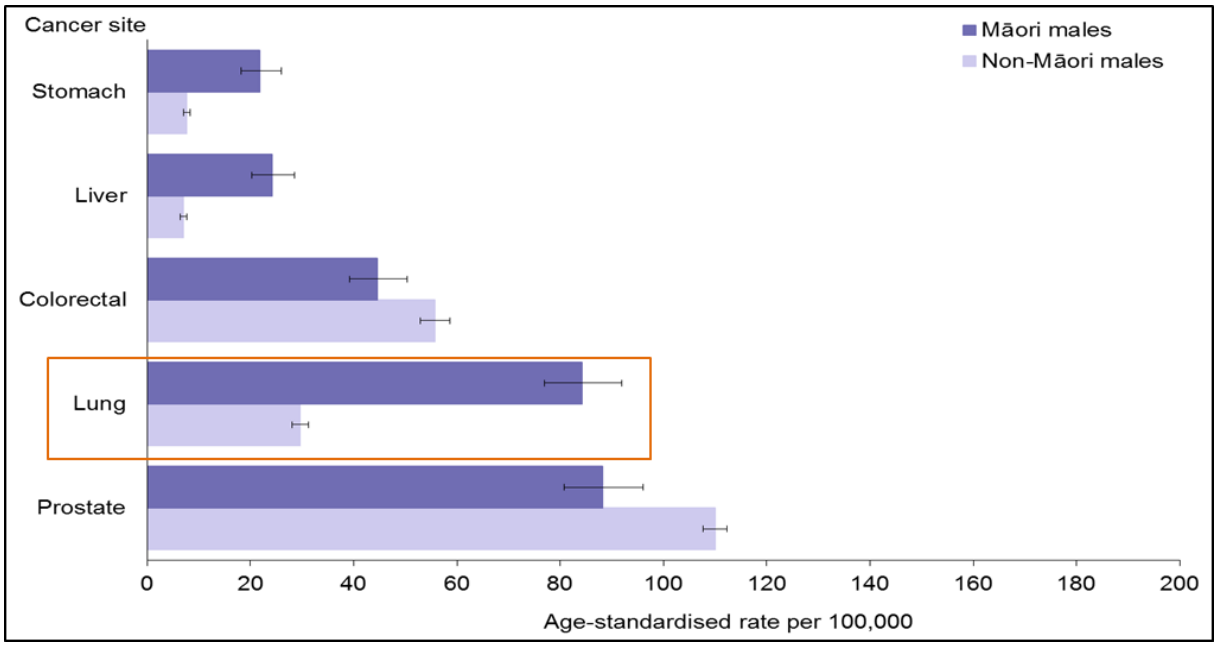
Sources: Mortality Collection Data Set (MORT), and New Zealand Cancer Registry (NZCR) Ministry of Health; New Zealand https://www.health.govt.nz

Figure 1.12 NZ Females (25+years) cancer mortality rates, per cancer site, Māori, and non-Māori 2010-12



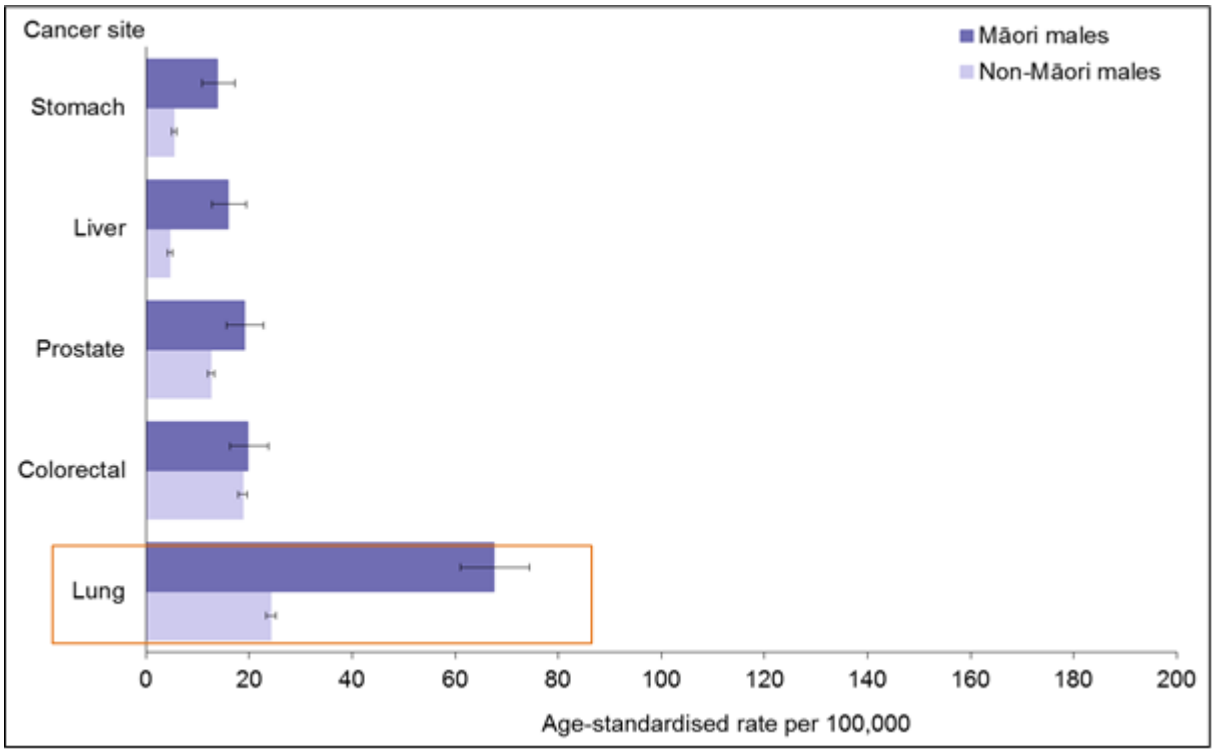
Sources: Mortality Collection Data Set (MORT), and New Zealand Cancer Registry (NZCR) Ministry of Health; New Zealand https://www.health.govt.nz

Figure 1.13 NZ Male (25+years) cancer registration rates, by cancer site, Māori, and non-Māori 2010-12



Sources: Mortality Collection Data Set (MORT), and New Zealand Cancer Registry (NZCR) Ministry of Health; New Zealand https://www.health.govt.nz

Figure 1.14 NZ Male (25+years) cancer mortality rates, per cancer site, Māori, and non-Māori 2010-12



Sources: Mortality Collection Data Set (MORT), and New Zealand Cancer Registry (NZCR) Ministry of Health; New Zealand https://www.health.govt.nz

1.5.2 Genetics

Genome-wide association studies (GWAS) have identified numerous genetic loci related to pulmonary function forced expiratory volume in the first second (FEV_1), forced vital capacity (FVC), and the ratio of the two. Several of these loci have also been linked to the clinical phenotype of COPD, characterised by a low FEV1/FVC ratio (<0.07). Many of these genetic associations have been replicated in more than one study and some have been defined by functional studies to be linked to the development of COPD.[154,155,162-164] GWAS studies have also been undertaken for lung cancer and again several loci have been replicated in large studies.[165-167] However, none of these studies have looked at genetic association in both COPD and lung cancer contemporaneously. This is important because my group has previously reported that confounding and mediating affects between loci may account for associations with lung cancer.[10,11,35-36,153-154,168] Specifically, in the large lung cancer studies, where there has been no capacity to phenotype for co-existing COPD. We have proposed that genetic associations reported to date may be misleading due to confounding effects.[10,11,15,36,153-154] Moreover, these studies are almost exclusively cross-sectional case-control studies where other biases such as survival bias may affect the results. In my thesis I report the results of the first prospective study to explore genetic association with lung cancer in smokers followed for 6 years where all cases of lung cancer were phenotyped for COPD at baseline and prior to the development of lung cancer.

Although genome-wide association studies (GWAS) have identified single nucleotide polymorphism (SNP) associations with lung disease and lung function, in the majority of these discoveries the direction of the association and the SNP functionalities is yet to be determined. In our previous work we have been able to reveal more about the relationship of some SNP variants as they relate to lung cancer and COPD alone or to both diseases.[15,153,154,168] Using a prospective cohort my current work, has validated previous work from our lab. In addition, I have uncovered more about the direction of these SNP effects. In participants from a large screening trial, I have been able to examine associations with risk for COPD and or lung cancer as well as relationships with morbidity, mortality, and outcomes. This new knowledge gives a better understanding of the benefits of case finding within the screening context and has clarified who will have the largest benefit versus those who will have a minimal benefit from screening. This is important work as it should enhance

the experience of individuals during health service interactions especially for 'shared decision' making by personalising risk-benefit to their particular situation.

1.6 The Structure of this Thesis

The overarching hypothesis is that chronic obstructive pulmonary disease is a common comorbidity with lung cancer and is likely to affect lung cancer outcomes.

The key question and aim that this thesis proposes and endeavours to answer is "in what ways does COPD effect lung cancer diagnosis and outcome."

This thesis builds on previous work and earlier publications from our lab (Dr Robert Young, Auckland). Some of my preliminary results have been presented in abstract form at international conferences. The initial premise of my thesis is that the presence or (absence) of airflow limitation has a bearing on which participants developed lung cancer and who will benefit most from a lung cancer screening programme. We have previously reported results from a post-hoc analysis of the National Lung Screening Trial (NLST), showing a greater lung cancer incidence, fewer cancers diagnosed at a late stage and minimal over diagnosed lung cancers in those with baseline airflow limitation compared to those with no airflow limitation.[80,93] This important key finding has implications for screening eligibility in a population who have a high possibility of comorbid disease and death from a competing cause other than lung cancer. The following chapters of this thesis explore, discuss, and refine these associations in more detail.

1.6.1 Brief synopsis

Disclaimer: The data used in the studies within this thesis is based on ever smokers only. Never smokers, are not part of this thesis.

For consistency, in the chapters for each study, the Figures appear within the results section and the Tables appear at the end of the results section. The relevant tables used to derive a figure are referred to in the figure legend.

The essence of this epidemiological thesis is an exploration of how smoking related airflow limitation (COPD) defined by the spirometry lung function test, relates to lung cancer risk and screening outcomes in the setting of a large prospective randomised lung cancer

screening trial comparing the radiographic technologies of computed tomography (CT) scan and chest radiography (CXR).[80,90] Through each chapter I will present and discuss novel discoveries and new insights in a way that shows the broad and comprehensive nature of the relationship between tobacco smoking exposure, airflow limitation, lung cancer risk, genetic susceptibility and lung cancer outcomes in a lung cancer screening population. This thesis is put together in such a way as to show the interconnectedness of these themes.

Chapter 2: Methodology

Chapter 2 explains and describes the methodological approaches and methods used in this thesis. I have been able to incorporate several different study designs appropriate for each of the studies that have gone into building this thesis. Study designs include a large randomised controlled trial comparing CT vs CXR in lung cancer; a systematic review and meta-analysis; a case-case design comparing two ethnically different case series of lung cancer diagnosis in the Auckland region, and a genotyping analysis of over 10,000 lung cancer screening participants. The statistical approaches used to analyse the data for each study were chosen based on the discussions with Dr Young and our biostatisticians and their recommendations. These approaches are outlined in this chapter.

Chapter 3: Airflow Limitation and Lung Cancer Risk

I present the results of a study looking at the role of airflow limitation in lung cancer risk and development. This work has been peer reviewed and published in the *Annals of the American Thoracic Society 2017;* 14(3): 392-402 (DOI: 10.1513/AnnalaATS.201609-741OC).

Rationale: While epidemiological studies consistently show that COPD is associated with an increased risk of lung cancer, debate exists as to whether there is a linear relationship between the severity of airflow limitation and lung cancer risk.

Objective: We examined this in a large prospective study of older heavy smokers from the American College of Radiology Imaging Network sub-cohort of the National Lung Screening Trial (NLST-ACRIN). Airflow limitation was defined by pre-bronchodilator spirometry sub-grouped according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) grade 1-4.

Methods: In the NLST-ACRIN cohort of 18,473 screening participants, 6,436 had airflow limitation (35%) and 12,037 (65%) had no airflow limitation. From these groups, 758 lung cancer cases were prospectively identified. Participants with airflow limitation were stratified according to GOLD group 1 (N=1607), 2 (N=3528), 3 (N=1083) and 4 (N=211). Lung cancer incidence at study end (mean follow-up 6.4 years) was compared between the GOLD groups and those with no airflow limitation (referent group).

Measurements and Main Results Compared to those with no airflow limitation, where lung cancer incidence was 3.78/1000 person years (pyrs), incidence rates increased in a simple linear relationship: GOLD 1 (6.27/1000 pyrs), GOLD 2 (7.86/1000 pyrs), GOLD 3 (10.71/1000 pyrs) and GOLD 4 (13.25/1000 pyrs) - all P≤0.0001 vs referent group.

Conclusion: In a large prospective study of high risk smokers, we report a strong linear relationship between increasing severity of airflow limitation and increasing lung cancer risk.

Chapter 4 Airways Disease and Airflow Limitation Phenotypes

In this chapter I present the results of a Phenotype study where the NLST-ACRIN screening participants have been stratified according to self-reported doctor diagnosed respiratory disease and airflow limitation according to baseline spirometry result (according to GOLD COPD severity grade).

A summary of this work has been peer reviewed and accepted for publication in the journal CHEST, as a Brief Communication.

Rationale: Low-dose chest computed tomography (LDCT) is recommended for lung cancer screening for high risk current and former smokers. However, the importance of pre-existing airways disease and airflow limitation remains poorly understood.

Objectives: To examine the effect of previously diagnosed airways disease and/or airflow limitation on outcomes in the ACRIN subgroup of the National Lung Screening Trial (NLST).

Methods: Assessment of the risk of developing lung cancer, receiving surgery for lung cancer and dying of lung cancer relative to other causes, after sub-phenotyping screening participants according to self-reported doctor diagnosed airways disease, impaired lung function and airflow limitation (COPD).

Measurements and Main Results: Relative to asymptomatic smokers with no airflow limitation (i.e., Referent smokers), those with airflow limitation (pre-existing doctor diagnosed COPD) or undiagnosed COPD (i.e., diagnosed at baseline spirometry), had an increased risk of developing lung cancer, dying of lung cancer or dying of non-lung cancer causes. The absolute benefit from CT screening was greatest in Referent smokers and Undiagnosed COPD with reductions in lung cancer death rate of 37% and 26% (P=0.0008 and P=0.002 respectively), accounting for 96% of all lung cancer deaths averted. In contrast to those with diagnosed COPD, where screening conferred less benefit, in those with Undiagnosed COPD, those screened with CT were more likely to undergo surgery (57% vs 35% respectively, P=0.019), and had fewer lung cancer deaths (Odds of lung cancer death averted=0.34 (95%CI=0.16-0.72), P=0.0048) relative to chest x-ray.

Conclusions: In NLST-ACRIN participants, spirometry helps identify subgroups of eligible smokers for whom lung cancer screening achieves optimal reductions in lung cancer deaths. Specifically, those who are asymptomatic with no airflow limitation and those with undiagnosed COPD benefit most from screening. This outcomes-based approach might help to optimise lung cancer screening efforts by identifying those likely to benefit most.

Chapter 5 Competing Risk of Death

Chapter 5 presents a study looking at lung cancer screening and the effects of competing causes of death in the ACRIN-NLST sub-study. The study shows how competing risk and baseline comorbid disease profiles affect screening outcomes. This work has been peer reviewed and presented as a conference abstract at the American Thoracic Society – *Am J Respir Crit Care Med* 2017;195: A5174

Rationale: Annual computed tomography (CT) is recommended for lung cancer screening in the United States, however the effect of airflow limitation on the competing risks of dying with or without lung cancer remains poorly understood.

Objectives: To examine the relationship between airflow limitation, risk of lung cancer and outcomes from screening in a subgroup of the National Lung Screening Trial (NLST).

Methods: In 10,054 NLST participants, the risk of developing lung cancer, dying of lung cancer or dying of non-lung cancer causes were compared according to airflow limitation or lung cancer risk tertiles ($PLCO_{M2012}$).

Measurements and Main Results: The risk of developing lung cancer (N=395) was linearly related to the severity of airflow limitation, although the risk of dying of lung cancer (N=187) relative to dying of a non-lung cancer death (N=500) diverged. As lung cancer deaths increased comparable increases were seen in cardiovascular deaths and deaths from other causes relative to severity of airflow limitation and risk tertile. Lung cancer risk was associated with worse histological lung cancer subtypes and reduced surgical rates. For those in the highest lung cancer risk tertile, where airflow limitation was most prevalent (48%), the lung cancer-specific deaths averted with CT screening relative to chest x-ray was attenuated (3, 9 and 5 in tertiles 1,2 and 3 respectively, P=0.0436).

Conclusions: In NLST participants, airflow limitation was not only associated with an increased lung cancer risk, but also greater mortality from lung cancer and non-lung cancer causes. This may attenuate the benefit of CT screening for some at-risk individuals.

Chapter 6 $PLCO_{M2012}$ model

Airflow limitation, comorbid disease, and risk of lung cancer ($PLCO_{M2012}$): Effects on lung cancer screening outcomes in the National Lung Screening Trial.

There are numerous models which use clinical and lifestyle information to predict an individual's risk for developing lung cancer over a specified time period and are touted as a prediction risk score for lung cancer. However, these are blunt scores based mainly on age and smoking history with very few including history of COPD and not including objective measures such as spirometry measurement.

In Chapter 6, I discuss the $PLCO_{M2012}$ risk model for lung cancer and show that this model is also a risk model for COPD. I will show why this is important in the context of lung cancer screening programmes. This work has been peer reviewed and presented as a conference abstract at the American Thoracic Society – *Am J Respir Crit Care Med* 2019;199: A4476

Rationale: Annual computed tomography (CT) is now recommended for lung cancer screening in the United States. The assumption that those at greatest risk of lung cancer achieve the greatest benefit from screening has recently been challenged.

Objectives: To examine the relationship between airflow limitation, comorbid disease, and risk of lung cancer ($PLCO_{M2012}$) in regards to outcomes from screening in a subgroup of the National Lung Screening Trial (NLST).

Methods: In 10,054 NLST participants, the presence of airflow limitation and comorbid disease was compared across lung cancer risk quintiles ($PLCO_{M2012}$) and correlated with surgical rates and survival from lung cancer.

Measurements and Main Results: Estimation of risk using the $PLCO_{M2012}$ lung cancer model were comparable between those with lung cancer (N=395) or airflow limitation (N=3255). In a receiver-operator curve-analysis, the area-under-the curve (AUC) for lung cancer risk (AUC=0.67, 95% CI=0.65-0.70) overlapped with the presence of airflow limitation (AUC=0.65, 95% CI=0.64-0.67). Increasing lung cancer risk ($PLCO_{M2012}$) was associated with increasing airflow limitation, increasing respiratory comorbidity, and increasing all-cause mortality, where deaths from non-lung cancer causes diverged with lung cancer deaths (favouring the former). In those randomised to CT, significantly higher surgical rates and lower lung cancer mortality was found in Quintiles 2-4 but not quintile 5. The former accounted for 76% of all lung cancer deaths averted.

Conclusions: Lung cancer screening outcomes were greatest in those of intermediate (20-80%) risk and attenuated in those at greatest risk (top 20%) due to greater comorbid disease, airflow limitation and lower surgical rates.

Chapter 7 Airflow Limitation and Survival after Surgery

Airflow Limitation and Survival after Surgery for Non-Small Cell Lung Cancer: results from a systematic review and lung cancer screening trial (NLST-ACRIN sub-study).

Chapter 7 presents the results of a two-part study where a systematic review looked at the effect having COPD has on survival following lung cancer surgery for early stage I and II disease in unscreened populations. These findings were compared with the results from the

ACRIN-NLST screened population. This work has been peer reviewed and published in the journal *Lung Cancer* 2019; 135: 80-87.

Rationale: Lung cancer remains the single greatest cause of cancer mortality where surgery for early stage non-small cell lung cancer achieves the greatest survival. While there is growing optimism for better outcomes with screening using annual computed tomography, the impact of co-existing airflow limitation on survival remains unknown.

Objective: To compare survival in non-small cell lung cancer patients undergoing surgery stratified according to the presence or absence of pre-surgery airflow limitation.

Methods: We undertook a systematic literature search of non-screen lung cancer that encompassed studies reported between January 1946 and January 2017. Full-text articles were identified following eligibility scoring, with data extracted and analysed using a standardised analytical method (PRISMA). The results of this systematic review in non-screen lung cancers were compared to real-world results from a lung cancer screening cohort (N=10,054), where outcomes following surgery could be compared after stratification according to pre-surgery airflow limitation.

Results: In the systematic review, 7,588 subjects were included from 11 studies; 8 were retrospective, 3 were prospective. There were no signs of heterogeneity or publication bias as assessed by Begg's funnel plot and Egger test (heterogeneity testing $p = 0.504$; $I^2 = 0.0\%$). The overall hazard ratio for 5-year survival comparing those with and without COPD was 0.96 (95% CI =0.91-1.02). In the lung cancer screening sub-study of 10,054 screening participants, we found no difference in 6-year survival in those with and without airflow limitation.

Conclusions: Survival after surgery for non-small cell lung cancer is comparable between those with and without spirometry evidence of airflow limitation. This finding was replicated in lung cancer diagnosed during screening.

Chapter 8 Ethnicity

Are New Zealand Maori More Susceptible to Smoking Related Lung Cancer? - A Comparative Case-Case Study

In this chapter I present the results of a study that used a case-case design in which I have compared our existing NZ Caucasian lung cancer series[85] versus NZ Māori lung cancer cases. Our results show important differences in COPD rates and lung cancer histology between the two groups despite comparable smoking exposure. This work has been peer reviewed and published in the journal *EC Pulmonology and Respiratory Medicine 2019;* 8(1): 1-20.

Rationale: Māori have one of the highest incidences of lung cancer in the world, even after adjustment for age and smoking prevalence. Why this indigenous Polynesian population of New Zealand (NZ) appears so susceptible to lung cancer remains unknown.

Objective: We compared demographic characteristics and risk factors between Māori and NZ Caucasians diagnosed with lung cancer.

Methods: Between Jan 2004-Nov 2014, we identified 472 lung cancer cases in Māori. Data were collected retrospectively and included demographic data, histological type, spirometry, clinical stage, and survival. This was compared to 415 NZ Caucasian lung cancer patients recruited from the same referral centre.

Measurements and Main Results: Relative to Caucasians, despite comparable smoking exposure, we found Māori lung cancer patients were younger at the time of diagnosis (61 vs 67 years old) and had more aggressive histological subtypes, with fewer Adenocarcinomas ($P<0.05$). More importantly at lower smoking exposure (<20 pack years), the expected dose-response relationship observed between smoking dose and airflow limitation in Caucasians was absent in Māori where the prevalence of airflow limitation in the latter was two-fold higher ($P<0.05$). In a Cox-Proportional analyses we found that ethnicity was independently associated with all-cause mortality (Māori HR=1.4, P=0.03), after allowing for advanced clinical stage (Staging HR = 2.6, $P<0.0001$) and histology (Histology HR=1.4, P=0.04).

Conclusions: We found that Māori have more aggressive forms of lung cancer and their greater susceptibility may be mediated in part through a greater susceptibility to COPD

Specifically an increased prevalence of COPD, and a loss of the expected dose-response relationship, at lower smoking exposure. We propose that an ethnicity-by-smoking interaction exists, whereby Māori ancestry contributes directly to worse outcomes from smoking that require greater recognition in future NZ tobacco control policies.

Chapter 9 Genetic Susceptibility to COPD and Lung Cancer

Chr15q25 genetic variant (rs16969968) independently confers risk of lung cancer, COPD and smoking intensity in a prospective study of high risk smokers.

This work is currently in press at the journal Thorax, (thoraxjnl-2020-214839).

Chapter 9 presents the results of a genetic study looking at genetic susceptibility related to smoking exposure, COPD, and lung cancer. Importantly Chapter 9 will show how genetic susceptibility between COPD and lung cancer overlaps. Preliminary results of this study have been peer reviewed and presented as a conference abstract at the American Thoracic Society – *Am J Respir Crit Care Med* 2017; 195: A7052

Rationale: While the cholinergic nicotinic receptor loci (*CHRNA5*) has been consistently linked to lung cancer and smoking addiction, it's co-relationship with COPD is less well understood. To date the interaction between these diseases has not been tested prospectively in a cohort study.

Objective: Using 9,270 Non-Hispanic white subjects from the ACRIN-NLST cohort, a sub-study of the NLST that recruited high risk smokers and followed them for an average of 6.4 years, we examined the association between the *CHRNA5* variant (rs16969968 A/G) in the development of lung cancer relative to its association with COPD and smoking exposure.

Methods: We compared the frequency of the high risk AA genotype according to smoking exposure, lung function and COPD status. We also compared the lung cancer incidence rate according to the *CHRNA5* genotype. Lastly, we used stepwise logistic regression and mediation analysis to examine the role of the AA genotype of the *CHRNA5* variant in smoking exposure, COPD, and lung cancer.

Results: Relative to healthy smoking controls, the AA high risk genotype was associated with lower lung function, greater smoking exposure, the presence of COPD (OR=1.3, P<0.001) and

the development of lung cancer overall (OR=1.5, P<0.01), regardless of COPD status. In the mediation analyses we found this genotype was independently associated with smoking intensity (OR=1.42, P<0.05), COPD (OR=1.25, P<0.05) and getting lung cancer (OR=1.37, P<0.05).

Conclusion: In this large prospective study we found the *CHRNA5* variant to be independently associated with smoking exposure, COPD, and lung cancer (triple whammy effect).

The following thesis follows the thesis with Publication format of the University of Auckland which permits the inclusion of published works from the study period alongside unpublished work in a cohesive style and format. Where chapters represent published work, the citation for the publication is indicated on the front page.

Chapter 2 Methodology

2.1 Introduction

In this chapter the methodological approaches used in this thesis will be described. The organisation of the chapters making up the totality of this thesis is based upon the particular questions asked, and the data interrogation as appropriate for each of the studies undertaken. The specific methodology used is appropriate to answer these questions. For this PhD thesis I have had access to data from the following sources: -

- A large randomised controlled screening trial, the National Lung Screening Trial (NLST) with over 53,000 participants of whom I was able to analyse the results from a subset of 18,000 participants with spirometry and a subset of 10,054 with genetic data (see below);
- A systematic review and meta-analysis study I performed looking at survival outcomes after surgery for early stage non-small cell lung cancer in screened and non-screened patients;
- Two large ethnically different case series of lung cancers diagnosed in the Auckland region in a case-case design; and
- The genotyping results of a cohort of 10,054 from the NLST who underwent genotyping (see above).

These differing study designs allowed me to answer a number of key questions I had regarding the importance of airflow limitation, COPD, and their relationship to lung cancer.

Study designs and approaches used include a post-hoc analysis of data from a cohort study (closed population), a systematic review; a comparative case-case study; and various stratification studies utilising clinical and genetic data from a large population based prospective randomised controlled trial (NLST).

Funding support for sections of this thesis came from the following sources;

- A Johnson and Johnson, United States of America, grant was awarded to Dr Young. This gave us access to the ACRIN demographic and bio-specimen data for lung cancer

screening participant volunteers, and the forming the ACRIN-NZ sub-cohort. This grant provided the funding to support the genotyping for this thesis

- Grants U01-CA-80098 and U01-CA-79778 to the American College of Radiology Imaging Network (ACRIN) for support of the screening study.
- Funding support for the systematic review and meta-analysis came from a summer studentship grant awarded from the University of Auckland
- Funding support for the comparative case-case study came from a summer studentship grant awarded from the University of Auckland.

2.2 Background

The key element of this thesis is the complex interplay between the burden that tobacco smoking places on the development of airflow limitation, lung cancer risk, and deaths from lung cancer and non-lung cancer causes, within the context of a lung cancer screening trial. This includes stratifying study participants by careful phenotyping according to pre-existing self-reported symptoms, and doctor diagnosed comorbidities, relative to confirmed airflow limitation or no airflow limitation according to lung function (spirometry) tested at baseline visit.[2-6,11,80,102,103,113]

Importantly this thesis will look at various characteristics among lung cancer screening participants with respect to lung cancer diagnosis, lung cancer stage, lung cancer histology and lung cancer surgery, as they relate to outcomes for lung cancer specific mortality and all-cause mortality.

Both candidate gene and genome-wide association studies have identified numerous gene (single nucleotide polymorphism) SNP variations to be associated with COPD, lung cancer or both COPD and lung cancer.[10,11,14,35,154,164-168] Because of the complexity of the associations between smoking, COPD, and lung cancer, unravelling and understanding the direction of a genetic influence is complicated. For example, a SNP variant maybe associated with a susceptible, or protective effect on lung cancer where the presence of COPD could have a mediating effect.[11,49,54] This concept will be illustrated and discussed with reference to a study looking at a SNP of the *CHRNA5* nicotinic acetylcholine receptor subunit. Specifically, the AA genotype and its role in smoking behaviour, risk for chronic obstructive pulmonary disease (COPD), lung cancer, and outcomes has been looked at.[15]

The objective thinking process behind this thesis helped inform the various stratifications used to examine the large data set interrogated. Throughout this thesis, the various quantitative type analytical methods used were chosen as the ones most appropriate for the analysis undertaken. In this way I was able to answer the specific questions asked of the available data. Individual studies are represented in a dedicated chapter (Chapters 3-9), see Table 2.1 below.

The concept of lung cancer risk prediction models/tools was introduced and discussed in Chapter 1 (section 1.4.1). One of the most known lung cancer risk prediction models is the $PLCO_{M2012}$. This is a logistic regression model created from the Prostate, Lung, Colorectal and Ovarian screening trial data. In the PLCO study lung cancer group chest x-ray screening was compared to no screening. The PLCO model calculates a person's risk of getting lung cancer within a 6-year period. There are 11 clinical variables taken from questionnaire data including medical history, socio-demographic data, and smoking exposure, that inform the model. For two of the studies in this thesis I have compared the results from our ACRIN-NLST data with results that would have been obtained if using the $PLCO_{M2012}$ model stratified by tertiles or quintiles of risk (Chapters 5-6).

Table 2.1 Study Objectives, Study Type, Study Design, Statistical Methods

Chapter and Title	Objective	Study Type and Study Design	Statistical Methods
Chapter 3 *Published* Reduced Expiratory Flow Rate among Heavy Smokers Increases Lung Cancer Risk: Results from the NLST–ACRIN Cohort (N=18,714).	To examine the role airflow limitation has, in lung cancer risk and diagnosis.	Cohort Study (closed population) Randomised Control Trial (RCT) Population, stratified according to the presence of baseline airflow limitation (FVC/FEV1<0.70, GOLD grade 1-4) and no airflow limitation (FVC/FEV1>0.70 GOLD grade 1-4).	Differences in lung cancer Incident Rates (IRs), stratified by severity of airflow limitation based on GOLD grades 1–4 were compared. Differences in lung cancer IRs were compared using IR per 1,000 person years and Incident Rate Ratio (IRR). Differences in lung cancer histology were compared according to GOLD status using Fisher's exact test and mid-P exact test. Confidence intervals for IRs and IRRs were estimated using the exact method. Significance was defined as a two-tailed P less than 0.05. All statistical analyses were performed using SAS or STATA 4 Statistical software.
Chapter 4 *Accepted for publication as a research letter* Phenotype: Effects of "Airways Disease" and Airflow Limitation on outcomes in high risk smokers screened for lung cancer.	To examine the effect of airflow limitation and airways disease (i.e., phenotype) on outcomes in a subgroup of the National Lung Screening Trial (NLST).	Cohort Study (closed population) RCT - 10,054 NLST-ACRIN participants stratified by their baseline spirometry result according to GOLD-COPD severity grade, and self- reported respiratory disease	Differences in baseline characteristics and screening outcomes following stratification by phenotype were compared using Dunnett's test. Differences in death incidence rates were compared between air flow limitation groups, after adjustment for pack years and age, by general linear modelling assuming a Poisson distribution with a log link function. Contrasts were constructed for pairwise comparisons between the referent smokers' group and each of the increasing airflow limited groups. False discovery protected 'p values' were calculated to maintain an overall 5% significance level. All other statistical analyses were performed using SAS (v9.4 SAS Institute Inc, Cary, NC) or STATA statistical software.

Chapter and Title	Objective	Study Type and Study Design	Statistical Methods
Chapter 5 *Presented as a conference abstract* Competing Risk of death: The effect of airflow limitation and lung cancer risk on lung cancer screening outcomes in the National Lung Screening Trial (NLST).	To examine airflow limitation and lung cancer risk in relation to dying of lung cancer, or dying of a non-lung cancer cause in a subgroup of participants from the NLST	Cohort Study (closed population) RCT - 10,054 NLST-ACRIN participants stratified according to baseline severity of airflow limitation (GOLD grades), and $PLCO_{M2012}$ lung cancer risk tertiles.	The demographic and clinical characteristics of groups defined by either severity of airflow limitation FEV1% predicted and GOLD grade or tertile of lung cancer risk score ($PLCO_{M2012}$), were compared using chi-square test or parametric or non-parametric analysis of variance as appropriate. Differences in lung cancer incidence rates were compared according to cause of death stratified by GOLD status or tertile of lung cancer risk score using general linear modelling assuming a Poisson distribution with a log link function. Contrasts were constructed for pairwise comparisons between the healthy smoker group and each of the increasing airflow limited groups or between all combinations of the lung cancer risk score. False discovery protected p values were calculated to maintain an overall 5% significance level. Cox-Proportional models were used to assess the impact of lung cancer risk ($PLCO_{M2012}$) and airflow limitation on the risk of dying of lung cancer relative to other causes. Clinically important potential confounders (age and pack years) were adjusted for in each model as appropriate. All other statistical analyses were performed using SAS (v9.4 SAS Institute Inc, Cary, NC) or STATA statistical software.
Chapter 6 *Presented as a conference abstract* Airflow limitation, comorbid disease and risk of lung cancer stratified by ($PLCO_{M2012}$): effects on screening outcomes	To examine the relationship between airflow limitation, comorbid disease, and risk of lung cancer ($PLCO_{M2012}$) in regards to outcomes from screening	Cohort Study (closed population) RCT - 10,054 NLST-ACRIN participants. The presence of airflow limitation and comorbid disease was compared across lung cancer risk quintiles ($PLCO_{M2012}$) and correlated with surgical rates and survival from lung cancer.	Differences in lung cancer incidence, stratified by severity of airflow limitation based on FEV1% predicted and GOLD grade were compared. Differences in lung cancer incidence were compared according to cause of death stratified by baseline lung cancer risk based on $PLCO_{M2012}$ quintiles, using Fisher's Exact test and Mid-P exact test. Significance was defined as a two tailed $P<0.05$. All statistical analyses were performed using SAS or STATA statistical software.
Chapter 7 *Published* Airflow Limitation and Survival after Surgery for Non-Small Cell Lung Cancer: results from a systematic	To examine the effect of pre-existing spirometry defined COPD on post-surgical survival in NSCLC.	Study type – 2-part study 1.Systematic review looking at the effect that having pre-existing COPD has on lung surgery for early stage lung cancer 2. Survival analysis following surgery	The data was analysed according to standard methods for a systematic review. For published studies meeting the eligibility criteria, survival according to COPD status was compared using a hazard ratio with 95% confidence interval calculated for each study. The result of this analysis was plotted using a Forest plot with study heterogeneity and publication bias formally assessed using Begg's Test and Egger's Test. Any p-value less than 0.05 was considered statistically significant. Calculations were carried out using the following software package: STATA version 10.1 (Metan, Statacorp, College Station, Texas,

Chapter and Title	Objective	Study Type and Study Design	Statistical Methods
review and lung cancer screening trial (NLST-ACRIN sub-study).		for NSCLC in a subset of participants in the National Lung Screening Trial (NLST), where baseline spirometry was routinely measured	USA). For comparative purposes, in the NLST-ACRIN subgroup the risk of developing lung cancer, receiving surgery for lung cancer and dying of lung cancer relative to other causes was assessed after sub-phenotyping screening participants according to airflow limitation (COPD).
Chapter 8 *Published* Are New Zealand Maori more susceptible to smoking related lung cancer? A comparative case-case study	To compare demographic characteristics and risk factors between New Zealand Māori and New Zealand Caucasians diagnosed with lung cancer.	Case-case comparative study. Both case series were from the same geographical region of Auckland, New Zealand. Maori cases were identified retrospectively using local district health board databases where lung cancer was reported as the 1st,2nd, 3rd diagnosis according to in the ICD 10 code (C34). Maori ancestry assigned by self-reported ethnic identification and Maori surname. NZ Caucasian cases diagnosed with lung cancer were prospectively identified as part of a separate epidemiological study.	Demographic variables were compared by unpaired t-test for normally distributed continuous variables and chi-squared test for discrete variables with Mid-P Exact test on a two-tailed analysis (www.openepi.com access 18/12/2016). Survival was compared using Cox-Proportional Hazards regression model, the assumption of proportionality was tested by including log time dependent variables in the model. SAS statistical package (v 9.4, Cary, NC, USA was used. All tests were two tailed and $P < 0.05$ was considered significant.
Chapter *9* *Paper in press* Chr15q25 genetic variant (rs16969968) independently confers risk of lung cancer, COPD and smoking intensity in a prospective study of high risk smokers.	To examine the association between the *CHRNA5* variant (rs16969968 A/G) in the development of lung cancer relative to its association with COPD and smoking exposure.	Study type Cohort Study (closed population) Genomic analysis for N= 9,270 non-Hispanic white participants from the ACRIN-NLST cohort, using a pre-determined single SNP variant, analysis.	Differences in lung cancer incidence, and screening outcomes stratified by genotype. Differences in lung cancer incidence were compared according to cause of death using Fisher's Exact test and Mid-P exact test. *CHRNA5* variant (rs16969968 A/G) was examined using allelic, genotype and recessive models. Chi-square tests, test for trend or Fisher's exact test were used to compare categorical variable. T-test or Kruskal-Wallis tests were used to compare continuous variables by genotype. Prevalence rates and 95% CIs were calculated per 1,000 person years. Significance was defined as a two tailed $P<0.05$. Cox-Proportional hazard compared survival (HRs and CIs) Multiple logistic regression modelling used with and without adjustments are indicated where appropriate. CAUSALMED procedure of SAS was used for mediation analysis. All statistical analyses were performed using SAS statistical package (v 9.4, Cary, NC, USA).

2.2.1 Statistical analysis

Statistical approaches used in this thesis have varied according to study type. The statistical methods used are appropriate for each study as outlined in Table 2.2, and are described in detail in the individual study chapters. The threshold of statistical significance was set at $p<0.05$ and Hazard ratio 95% CI Statistical analysis was carried out using Stata v.9.4 (StataCorp, Texas, USA).

Statistical Support: Throughout this process I have been supported by an experienced biostatistician, Greg Gamble, (Department of Medicine, University of Auckland).

Table 2.2 Statistical terms used in this thesis

Term	Definition
Lung Cancer Prevalence	The number of lung cancer cases diagnosed as a percentage of the total number of screening participants including both prevalent lung cancers (diagnosed in the first year of screening – T0) and incident cancers (diagnosed during subsequent years T1 to T6). Described as a percentage or per 1000 persons.
Lung Cancer Incidence Rate (IR)	The number of lung cancer cases diagnosed over a defined period as a function of the total number of person years calculated according to the total number of screening participants and years of screening. Described as rate/1000 person years.
Lung Cancer Incidence Rate Ratio (IRR)	The ratio of one lung cancer incidence rate relative to a reference lung cancer incidence rate. Provides a crude estimate of relative risk but in the context of comparing two incidence rates rather than two prevalence rates.
NLST - Study year T0-T6	Study year in which a lung cancer was detected. T0-T2 = Active screening years (i.e., annual CT or Chest X-ray). T0 is baseline scan. T3-T7 = Surveillance/Follow up years
Logistic Regression	This type of analysis looks at associations between categorical (binary) variables. Specifically, it looks at the association of a dependent variable and a set of independent variables without assuming that the independent variables are normally distributed.
Fisher's exact test	A test of independence between nominal or discrete variables to see whether the proportions of one variable are different, depending on the value of the other variable. It is the test of choice when the sample size is small.
mid p–value	Is defined as half the conditional probability of the observed statistic p-value plus the conditional probability of more extreme values. Used in tests of independence in a contingency table.

Term	Definition
P value In this thesis significance was set at p value less than 0.05 (p<0.05), for a two tailed test (i.e., a 5% chance that the result is no different from the null hypothesis.	Is an area in the tail of a distribution curve that tells you the odds of a result happening by chance. It indicates the probability of getting the observed result, if the null hypothesis is true. It involves testing a null hypothesis by comparing the data observed with the predictions of a null hypothesis. If this estimated probability (the *P* value) is below the significance value, then it is unlikely that the null hypothesis is true. Significance was set at p value less than 0.05 (p<0.05), for a two tailed test in my studies.
Two tailed P	Is testing the possibility of a relationship in both directions meaning that 0.025 of the result (i.e., the test statistic) is at each end of the data distribution curve, (i.e., the area under the curve) when significance is set at less than 0.05 (p<0.05).
False discovery protected p values	The false discovery rate (FDR) is the ratio of the number of false positives to the number of total positive test results. The FDR approach adjusts the p-value for a series of tests. A p-value gives you the probability of a false positive on a single test.
Dunnetts' test	Is a posteriori or post hoc test restricted to comparing a single control group against a number of experimental groups. It is used after a significant one-way analysis of variance (ANOVA) has been done. Dunnetts' test determines which differences are significant. It does not test experimental groups against each other.
Poisson distribution	Gives the expected frequency of an outcome. Used to describe the scattering of rare events within a large population over a given time period. It is used to count events as whole numbers occurring independently of each other (i.e., one event does not influence the occurrence of another).
chi-square test	A Chi-square test is for testing independence between categorical data (values) from a single population. It expresses the goodness of fit for how likely it is that an observed distribution is due to chance. It only determines the probability of independence and does not tell anything about the association between the two variables.
Cox-Proportional models	An analysis of survival exploring a number of explanatory effects such as treatment or other influencing variables. Allows for censored data for those still alive at end of observation period.
Parametric test of variance Non-parametric test of variance	Parametric tests of variance usually assume equal variance between two or more samples from two different populations or processes, e.g., mean, mode, median, zero distortion (i.e., the distribution shape is symmetrical). Nonparametric tests do not assume equal variance about the source Distribution(s) and as such are often called "distribution-free.
Likelihood ratio test Akaike Information Criterion Bayesian Information Criterion	These are data metric tests used to assess which models provide the best fit.

Source: Introduction to SAS. UCLA: Statistical Consulting Group.
https://stats.idre.ucla.edu/sas/modules/sas-learning-moduleintroduction-to-the-features-of-sas/ (accessed August 22, 2018

2.3 American College of Radiology Imaging Network (ACRIN) Arm of NLST

Following collaboration discussions with leading investigators (Dr C Berg and Dr D Aberle) from the United States National Lung Screening Trial (NLST),[90] the application process to the ACRIN Data Access Committee applying for access to the ACRIN subject demographic, follow-up data and bio-specimen (spirometry and DNA) was completed. The application underwent a three-way critique by high level review committees, namely - (1) Research Evaluation Panel, (2) Scientific Review Committee, (3) Logistical Review by the ACRIN Data Access Committee. In addition, a parallel review panel from Brown University's Biostatistics Centre also considered the application as this facility has responsibility for overseeing the NLST-ACRIN data management.

The successful completion of the application process confirmed that our proposed study was judged to be in keeping with NLST-ACRIN objectives and the proposed study was robust and had scientific merit. In addition to satisfying these review committees further assessments of the application were made regarding the technical parameters involved including the practicality and feasibility of completing our proposed studies. The clinical and scientific impact resulting from the studies proposed was also considered. After completing the application and review process we have proceeded with a successful collaboration through a Memorandum of Understanding enabling us to produce exciting and novel results from the data made available to us.

The National Lung Screening Trial (NLST), which enrolled 53,454 current or ex-smokers in 33 centres in the United States, is the largest randomised controlled trial (RCT) to assess the utility and benefits of low-dose helical computerised chest tomography (LDCT) compared to chest radiography (CXR) to screen high risk ever smokers for lung cancer. Subjects randomised to the LDCT group are the experimental (intervention) arm, with the control arm being those randomised to CXR. In this study parallel assignment and no blinding to intervention was used. The intervention and control groups included 26,722 and 26,732 participants respectively.

Both the LDCT and CXR groups provided interval health status updates by follow-up at six monthly intervals. The study began enrolling participants in August 2002 with follow up through to December 31st, 2009.

The study involved thirty-three screening centres that were urban tertiary care hospitals having expertise across all aspects of cancer management. Of these, twenty-three were American College of Radiology Imaging Network (ACRIN) sites and ten were National Cancer Institute (NCI)-based lung cancer screening (LCS) sites. A total of 53, 454 volunteers signed consent to participate and were enrolled, although 192 subjects were later found to be ineligible.[169]

Multiple recruitment strategies were used and included advertising on local and national television and radio; physician referral; targeted direct mailing to known smokers and minority groups, community group presentations, and word of mouth, with mass mailing proving to be the more effective strategy.[170] Recruitment in the first year was so successful that 34,688 participants were enrolled which exceeded the expected 25,000 assumed by the original power calculations. The Hu-Zelen model was used for the power calculations which estimated the number of participants needed in each arm at specified time points, based on data from the Mayo Lung Project participants.[171] These participants had to have at least a 30 pack-year smoking history and be aged between 55 and 74 years at entry into the Mayo study. The successful recruitment rate of the NLST, enabled a 20% magnitude of difference in lung cancer death rate, (CT versus CXR) to be detected with 90% power with three screening examinations. Participant retention rate was very high for this trial with 96% follow up until study termination or death whichever came first. Main reasons for non-completion were – withdrawal of consent, lost to follow up (defined as failure to contact participant after repeated attempts), cognitive impairment or medical problems. Where withdrawal of consent occurred, all data for these participants was removed from the trial data base.[90,169-172]

The main eligibility criteria for the NLST study were minimal being only giving signed informed consent; a smoking history at least equivalent to 30 pack years and not stopped, for more than 15 years; aged 55-74 years; able to cope with the positioning required for the CT scanning procedure and lung cancer surgery if appropriate. Additional eligibility criteria included not having a prior diagnosis of lung cancer, and not having had any portion of the lungs removed; no comorbid conditions severe enough to be a risk for mortality prior the study conclusion. Eligible participants were randomised to annual CT or Chest x-ray by random six or eight subject block assignment, stratifying by age, sex, and screening site.

Participants underwent either an annual LDCT or CXR screening for 3 years (i.e., baseline and two subsequent annual examinations), then follow up tracing for a further three years, the average surveillance time was 6.4 years. This timeline reflects that the primary endpoint showing a clear survival benefit in the CT arm had been reached. During the total 7.5 years of the study 2,058 lung cancers were identified and confirmed by histology over the screening and follow up period. Seven Hundred and sixty-eight (768) of these cancers came from within the ACRIN cohort of 18,714 participants across 23 sites.[90,169,171]

For studies such as the NLST it is important that the cohort enrolled is representative of the base population. To check this, the demographic characteristics of individuals in the Tobacco Use Supplement of the US Census Bureau Current Population Surveys meeting the smoking history and age components of NLST eligibility were assessed and found to be similar to the NLST cohort. However, NLST subjects did differ somewhat in terms of higher education level, a lower current smoking status and being slightly younger.[90,169]

Specifically, in this thesis (Chapters 3-6 and Chapter 9), the data I have analysed comes through our collaboration with key personnel from the American College of Radiology Imaging Network (ACRIN) arm of the NLST because these sites had collected at baseline visit both lung function data and blood samples for DNA analysis. Using data collected for 10,054 subjects enrolled across the ACRIN sites of the NLST, I have looked at relationships and associations between baseline demographics, lung function data, self-reported doctor diagnosed comorbid disease status, subsequent lung cancer diagnoses and death. In addition, the blood DNA samples collected at the baseline visit from consenting participants has enabled me to look at results from a genetic panel developed by our group. The genes analysed were chosen based on previously published studies showing associations with either lung cancer, airflow limitation (COPD) occurrence, or both.[11,15,35,153-154,162-168] My thesis has focused on the relationship between COPD and lung cancer in the context of outcomes from lung cancer in the screened and unscreened setting.

To date the NLST is the largest randomised control screening trial for lung cancer where participants in both arms received a screening intervention. Details of the study design protocol, participant characteristics and results have been previously published.[90,169-173] The fact that each arm in this study had an intervention is important as other previous and on-

going lung cancer screening trials used usual care controls. For example, the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO) randomised participants in the lung cancer screening arm to either chest x-ray or continue with their usual medical care (i.e., no intervention until symptoms developed).[171] In smaller European based lung cancer screening studies there is inconsistency in the protocols which will make it difficult to combine them for analysis or a head-to-head comparison. However the European trial approach randomising participants to the screening intervention (i.e., CXR, CT), or usual medical care (no screening), gives some helpful indicators as to the most appropriate screening intervals based on nodule size and growth rate characteristics, for designing future screening protocols and management.[91,174-177] Because of its design, successful recruitment of participants, with high levels of adherence and compliance to follow-up, the NLST was the first and largest lung cancer screening trial to show a positive result, namely that CT screening reduces lung cancer specific mortality by 20% compared to CXR screening.[90] Subsequently, the recently released final results of The Dutch-Belgian lung-cancer screening trial (Nederland-Leuvens Longkanker Screenings Onderzoek (NELSON), has confirmed this mortality benefit even though the study is smaller, predominately male and of younger age than the NLST.[177] In the NLST, for the classification of comorbidities and cause of death, both The International Classification of Diseases Ninth Revision, Clinical Modification (ICD-9-CM) system, and the American Medical Association Current Procedural Terminology code set were used.[178,179]

Apart from the basic demographic, and health status data collected from all NLST participants an important and major strength from the ACRIN study sites is the availability of pulmonary function test results and the bio specimen repository, where collection methods were standardised across sites. Specifics of data collection methods are available from, (https://www.acrin.org and https://clinicaltrials.gov/ct2/search (last accessed 09 March 2018)).

For all participants, background medical and overall information were collected and recorded on designated forms including basic demographic data, lifestyle, socio-economic status, occupational exposures, relevant health status, comorbid diseases, and family history of lung cancer. In addition, a detailed tobacco smoke exposure history was recorded that drilled down on specifics such as cigarettes smoked per day, years smoked, and age

started/stopped, passive smoke exposure, cigarette brands used, and for current smokers any thoughts about stopping. All this clinical information was to be used to facilitate a detailed lung cancer risk profile for each individual.

Lung function tests were completed by trained staff according to the American Thoracic Society recommended guidelines for test reliability and acceptability[180] using the same hand-held spirometry device across all sites, a Spiropro spirometer (eResearchTechnology, GmbH, Germany supplied by SensorMedics Corp. Yorba Linda, CA). The best manoeuvres of acceptable blows were recorded for the subject data-base using the study ACRIN 6654 NLST Pulmonary Function Test Form. NHANES predictive equations appropriate to the US population were used. If the participant was unable to perform spirometry on the day of their enrolment, they were rescheduled a spirometry visit according to a pre-specified time frame. The timing of this rescheduled visit depended on the reason for non-completion and occurred within 3 weeks of the enrolment.[90]

Spirometry values in the NLST were pre-bronchodilator as enrolment was begun in August 2002 through to April 2004. Importantly the first GOLD COPD strategy was published in 2001 and was still being evaluated by the various Respiratory Societies meaning the GOLD suggestion of COPD diagnosis being based on post-bronchodilator values had not been universally accepted at the time the NLST trail began.[28] For study integrity it is important to maintain consistency with the protocol which would have been developed prior to the slow adoption of the GOLD COPD diagnostic recommendations. More importantly the GOLD Board of Directors committee review these recommendations yearly and the criteria for GOLD grades change.[62]

The forced expiratory manoeuvre following the maximal breath inhalation was used. Results for forced expiratory volume in one second (FEV_1), forced vital capacity (FVC) and the FEV_1/FVC ratio were expressed in absolute terms as well as the percentage of a person's predicted value (calculated on height, age, and gender).

Blood was collected from participants who signed an additional informed consent form for specimen collection as per the ACRIN specific protocol and were collected following randomisation. Each site was issued with specimen collection kits from the Colorado Lung SPORE Tissue Bank. The kits contained unique bare codes linked with the NLST participant ID

at each of the ACRIN sites meaning the Colorado Lung SPORE tissue Bank did not have access to participant credentials information. For those not agreeing to sample collection at baseline but who during the time of the study underwent a screening–related tissue biopsy or lung resection were asked to consent to procure some remnant tissue from these procedures.[173]

Because death due to lung cancer and overall mortality were the primary endpoints for the NLST, it was paramount that information on all deaths (no matter what the cause) known to have occurred during the study was obtained. In summary, the three ways study investigators found out about a participants' death were at the time of a study update questionnaire, being informed by a relative, and the National Death Index Plus searches. Study sites made every effort to procure the death certificates and all information from the death certificate was coded and recorded in the trial database. An Endpoint Verification Process (EVP) was in place to review medical records for determination of the cause as being due to lung cancer. Cause of death, on death certificates was based on the US National Center for Health Statistics with the EVP cause was considered the definitive cause and used in the primary end point statistical analyses.[90,171]

2.3.1 National Lung Screening Trial registration and funding details

'The national Lung Screening Trial Clinical trial registration no. NCT00047385'

'The American College of Radiology Imaging Network (ACRIN)–National Lung Screening Trial (NLST) was supported by grants (U01-CA-80098 and U01-CA-79778) under a cooperative agreement with the Cancer Imaging Program, Division of Cancer Treatment and Diagnosis. The Lung Screening Study (LSS) of the NLST was supported by contracts with the Early Detection Research Group and Biometry Research Group, Division of Cancer Prevention, University of Colorado Denver (N01- CN-25514), Georgetown University (N01-CN-25522), the Pacific Health Research and Education Institute (N01-CN-25515), the Henry Ford Health System (N01-CN-25512), the University of Minnesota (N01-CN-25513), Washington University in St. Louis (N01-CN-25516), the University of Pittsburgh (N01-CN-25511), the University of Utah (N01-CN-25524), the Marshfield Clinic Research Foundation (N01-CN-25518), the University of Alabama at Birmingham (N01-CN-75022), Westat (N01-CN-25476), and Information Management Services (N02-CN-63300)'.[90]

2.3.2 Airflow limitation and lung cancer

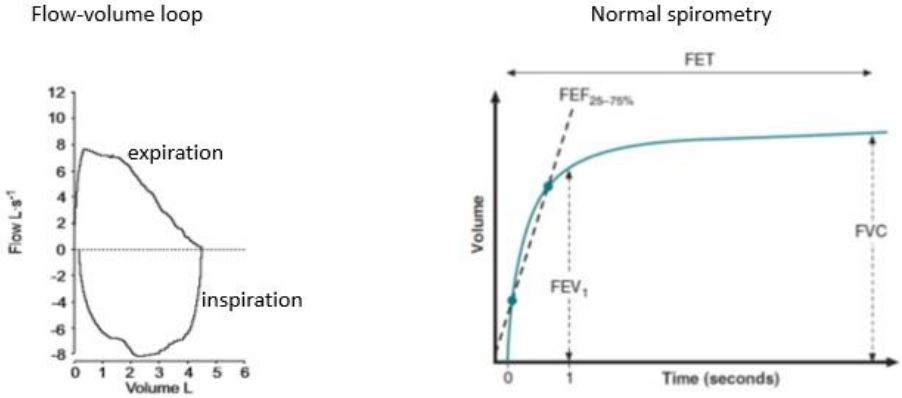
In the ACRIN cohort most participants had lung function measured by spirometry during their enrolment (registration) or screening baseline visit (referred to as T0) when these criteria were met: - had not had a chest infection in the preceding 3 weeks and had not used a short-acting or long-acting bronchodilator inhaler in the preceding 6 hours or 24 hours respectively. Those not meeting these criteria at the enrolment visit did the spirometry testing at a rescheduled baseline visit. All baseline information including spirometry, health, socio-demographic, and quality of life questionnaires were completed 2 to 4 weeks prior to randomisation. The spirometers used at the ACRIN sites were the SpiroPro device, (Figure 2.1). The American Thoracic Society (ATS) spirometry statement for performance standard and quality was followed.[180] Within the ACRIN-NLST cohort of N=18,714 participants, 99% (n=18,473) of volunteers had lung function spirometry as part of their study participation. Of those with acceptable spirometry, airflow limitation was seen in 35% of people based on $FEV_1/FVC < 0.70$.[37]

Figure 2.1 Spirometer

Throughout this thesis, to define airflow limitation I have used the fixed ratio less than 70% of maximum breath out in one second divided by the total maximum breath out, expressed as $FEV_1/FVC < 0.7$, (i.e., the maximal forced expiration in one second (FEV_1) over full expiration (FVC), achieved after taking a maximal inhalation (breath in). To achieve this, it is important that the person must inhale a breath, as deep as they possibly can, then exhale

(blow out) as fast and as hard as they can and keep the exhalation going until there is no force left, felt by the person as if they have run out of air. The measurements achieved in this forced manoeuvre are represented on a volume-time curve and a volume-flow loop. Where the start for the curve is at 'time 0' - 'volume 0' and for the loop 'flow 0 – 'volume 0'. In the spirometry manoeuvre the majority of the air expelled from large airways occurs in the first second. In the volume-time curve the peak of the curve rapidly rises becoming horizontal as the expelled air volume gets lower and lower over time with a plateau occurring at 4-5 seconds. In the volume-flow loop, as air is forcibly expelled the point occurs rapidly before descending straight from top to bottom as the flow decreases, this may be followed by a forced inspiration giving the 'loop' (Figure 2.2).

Figure 2.2 Spirometry Flow-volume and Volume-time



Source: Adapted from the Global Initiative for Chronic Obstructive Lung Disease, global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease updated 2020. (www.GOLD.org accessed February 2, 2020)]

As spirometry is used as a diagnostic tool in respiratory disease the quality of the blow is very important. To this end, criteria for test quality and test reporting have been published as a joint statement by the American Thoracic Society (ATS) and the European Respiratory Society (ERS) Task Force.[180] Generally, three blows is standard, with two of these being within 150mls or 5% of each other. However, some people may have difficulty meeting the quality standard requiring further attempts but it is recommended no more than eight blows are made in one testing block with at least 30 seconds between each blow. The rationale for this is that as it is an effort dependant manoeuvre people get tired so obtaining any better values would be unlikely beyond this point.

The airflow limitation severity classification used for airflow limitation (COPD) is in accordance with the Global Initiative for Chronic Obstructive Lung Disease (GOLD) strategy for diagnosis, management, and prevention (www.GOLD.org accessed February 24, 2019). For simplicity when the FEV_1/FVC is <0.70, the GOLD committee apply specific cut-points to the FEV_1% predicted to allocate a COPD grade. These GOLD grades range from GOLD-1 through GOLD-4 in descending order where GOLD-4 is the most severe, Figure 1.7.

2.4 COPD Status, Surgery for Lung Cancer and Survival: Systematic Review and Meta-Analysis

When non-small cell lung cancer, (NSCLC) is diagnosed at early stage, (stages I and II) disease curative intent surgery is the preferred treatment option. Pre-existing airflow limitation (COPD) increases the risk of lung cancer, may influence surgical decisions, and affect outcomes when surgery is performed. Chapter 7 explores the issue of COPD and the influence that having this co-morbidity has on survival when surgery is performed for early stage non-small cell lung cancer.

A systematic review and meta-analysis were done looking at studies reporting surgical outcomes for non-screen detected lung cancer and these results were compared with outcomes for lung cancer diagnosed in individuals taking part in a randomised CT vs CXR lung cancer screening programme (the NLST-ACRIN cohort). The key inclusion criterion was that pre-operative spirometry defined COPD was available for both the un-screened and screened lung cancer cases.

2.4.1 Systematic review search strategy and inclusion criteria:

Correct reporting of this study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[181,182]

The literature search was performed in Ovid MEDLINE, EMBASE and the Cochrane Database of Systematic Reviews as well as Google Scholar. The date range limit for publication was from 1946 to January 2017. No language restriction applied in the initial searches. For non-English publications, when all attempts to obtain an English translation failed, the study was excluded from further analysis.

Key word search criteria were clustered into five groups [1-5], which allowed for each abstract to be scored based on the number of keyword groupings it contained, Figure 2.3. When an abstract included keywords from at least 3 groups as well as mentioning "*non-small cell lung cancer/lung cancer/adenocarcinoma*", "*pneumonectomy/lobectomy/resection/removal*" and "*mortality/survival*", the full article was read and analysed.

Figure 2.3 Key word search groups

group 1: *"non-small cell lung cancer/lung cancer/ adenocarcinoma",*

group 2: *"chronic obstructive pulmonary disease/COPD/chronic bronchitis/pulmonary emphysema",*

group 3: *"spirometry/bronchospirometry/lung function/airflow limitation",*

group 4: *"pneumonectomy/lobectomy/resection/removal" and*

group 5: *"mortality/survival".*

Study types reviewed included retrospective, prospective or randomised study. Before a study was deemed to be eligible for inclusion into the systematic review, final checklist criteria needed to be met. These criteria were: -

- "*non-small cell lung cancer (stages I-II)*", "*spirometry criteria for COPD*",
- "*results stratified by COPD/lung function*",
- "*preoperative lung function*",
- "*lung surgery*" and
- "*mortality/survival (1-10 years)*".

When the same patient population was encountered in separate publications, only the publication with the most relevant data was included. To ensure no potential study had been missed, the bibliographic references of those studies publications meeting the final eligibility criteria were reviewed. Eligibility for further analysis was checked by two reviewers and a third reviewer resolved any discrepancies. A Microsoft Excel spread-sheet was used to record the reasons for exclusion of studies based on their abstracts.

Only ten eligible studies from the systematic review met the stringent screening process for inclusion. Study heterogeneity and publication bias was formally assessed. Five-year overall survival percentages for cases and controls were then analysed using a forest plot.

Data Extraction: For each Study the characteristics extracted were the lead author's name, year of publication, country in which the study was performed, study design, follow-up period, number of patients (inclusive of the number of cases and controls), how the cases of COPD were defined, the 5-year overall survival rates for cases and controls and the proportion of patients with stage I-II NSCLC. We also included the odds ratio and 95% confidence interval for COPD and its effect on survival post-lung cancer surgery which was calculated on quantitative analysis. The data extraction details are as shown in Table 7.1, (Chapter 7).

Statistical Analysis: The data was analysed according to standard methods for a systematic review (PRISMA).[181] By using PRISMA the 'a priori' plan for this review is declared ensuring its transparency and accessibility. For the ten studies meeting the eligibility criteria, mortality (hazard ratios) according to COPD status was compared using a risk ratio with 95% confidence interval calculated for each study. The results of this analysis were compared using the DerSimonian and Laird Random Effects Model and plotted using a forest plot. This Random Effects Model is considered the easiest to use objective method for testing and estimation of combined characteristics data from a number of individual studies and assigning a weighting with which to analyse and asses a treatment effect. Study heterogeneity and publication bias were formally assessed using Begg's test (Funnel plot) and Egger's test. Any p-value less than 0.05 was considered statistically significant. Calculations were carried out using the software package: STATA version 10.1 (Metan, Statacorp, College Station, Texas, USA).

2.5 Ethnicity and Lung Cancer in the New Zealand Context

New Zealand is a multicultural society where numerous ethnicities reside. In this context health outcomes differ across societal groups where particular vulnerabilities and susceptibility exist due to the effects of socioeconomic circumstance, health beliefs, lifestyle choice, exposure, and genetic background. In New Zealand, health outcomes vary across and between ethnic groups. Over time inter-race marriage means most current people

identifying as Maori also have other ancestral heritage mostly from European heritage. This is important as ancestral heritage informs one's innate immune system and has an influence on one's immune response to pro-inflammatory stimuli such as cigarette smoking. Tobacco smoking is the main exposure linked to lung cancer and chronic obstructive pulmonary disease but the amount and the way that people smoke varies.[31-32,145-152,156] New Zealand Maori have higher rates of both diseases than New Zealand Caucasian.[158-160] To improve understanding of this phenomenon, I undertook a case-to-case epidemiological study, (Chapter 8).[157]

Case-to-case studies: A case-to-case study approach was appropriate allowing direct comparison of several important demographic variables between our two groups, for example smoking history, age at diagnosis, stage at diagnosis, histology, and time from diagnosis to death. Using a case-case approach aided the exploration, description, and explanation of the differences in lung cancer characteristics seen in New Zealand Maori and European within the same source population, (i.e., the greater Auckland metropolitan area, a referral base that encompasses the upper half of the North Island of New Zealand (approximately 40% of the population, 1.5 million). Another advantage to using this design method within a given geographical region is that differences in inter-personal recall about lifestyle behaviours such as smoking and smoking exposure are likely to be less. Further differences in the documented disease characteristics and outcomes are likely to be less, permitting more reliable between group comparisons to be made. This type of study design is better able to ask and answer the 'who', 'what', 'where', 'how, and 'why' questions posed by the differences in lung cancer incidence, type, diagnostic stage and mortality outcomes observed in New Zealand Maori versus New Zealand Caucasian.

Methods: This was a retrospective review of Maori lung cancer case records where cases were identified using local district health board databases with lung cancer reported as the 1st, 2nd, or 3rd diagnosis according to the International Classification of Diseases coding system (ICD 10 C34). This identified lung cases diagnosed between January 1st, 2004 and December 31st, 2014, where the date of diagnosis was taken as the date of the most definitive investigation usually the histology report or CT imaging of the lungs. Ethnicity was assigned according to self-reported ancestry.

NZ Caucasian patients diagnosed with lung cancer were prospectively identified as part of a separate epidemiological study between January 1st 2004 and December 31st 2008.[85] Data collected included; age at diagnosis, gender, ethnicity, smoking history, histological subtype, stage at diagnosis, and months of survival from diagnosis, spirometry and date of diagnostic CT or diagnostic histology.

Lung cancer was staged I-IV, by the TNM Staging was defined according to the IASLC lung cancer staging project 7th edition.[183] Histological subtype from pathology reports were classified as small cell, adenocarcinoma, squamous cell, non-small cell (for cancers with no more precise classification), or other which encompassed the remainder of cancers. Carcinoid tumours, secondary tumours or benign tumours of the lung were excluded from this study. The presence of chronic obstructive pulmonary disease (COPD) was defined according to sighted spirometry reports or lung function results using Global Obstructive Lung Disease (GOLD) criteria. Smoking histories were exclusively based on recorded pack years of smoking exposure.

In two sensitivity analyses, Māori lung cancer cases were restricted to those with Māori surnames to enrich them for Māori ancestry. The first was an age gender matched comparison and in the second, the demographic characteristics were also compared after stratification according to gender, smoking status, and presence of COPD.

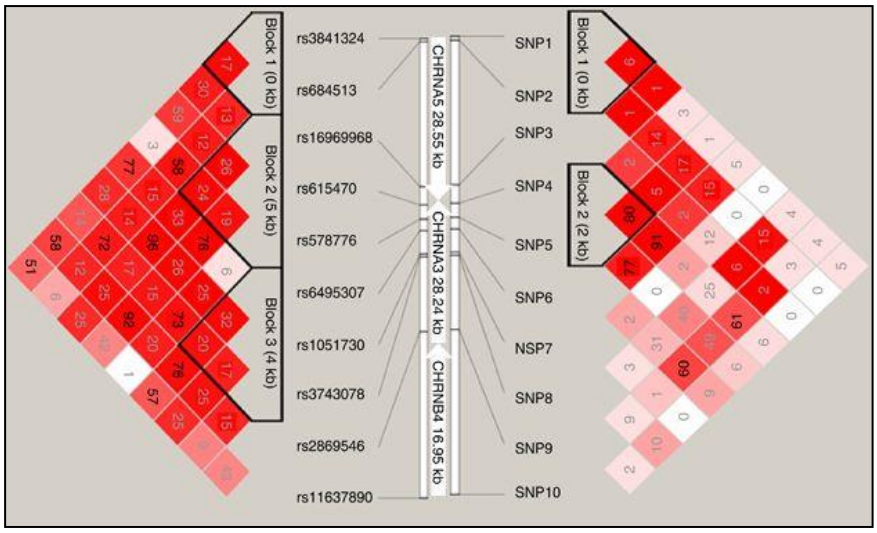
Statistical Analysis: Demographic variables were compared by unpaired T-test and ANOVA for continuous variable and chi-squared test for discrete variables with Mid-P Exact test on a two-tailed analysis. For those diagnosed between 2004 and 2007, survival and factors underlying survival were compared in a Cox-proportional model. Survival was compared using Cox-Proportional Hazards regression model. Statistical tests used SAS statistical package (version 9.4, Cary, NC, USA). Any p-value that was less than 0.05 was taken as statistically significant.

2.6 Genetic Susceptibility to COPD and Lung Cancer

Genetic inheritance and environmental exposure are highly relevant to the study of smoking related diseases especially lung cancer and COPD where background genetic susceptibility can contribute to bad outcomes in those highly susceptible to the effects of tobacco smoke.

In chapter nine of this thesis, I report the results of a study examining the association between the SNP variant rs16969968, of the *CHRNA 5* gene on chromosome 15.25.1q, (Figure 2.4), and the development of lung cancer relative to its association with COPD and smoking. This variant is highly relevant to lung cancer and COPD risk as well as smoking nicotine addiction.[12-15]

Figure 2.4 *CHRNA5* and rs16969968



From our ACRIN-NLST cohort of 10,164 participants, with demographic and clinical data, there were 10,054 unique buffy coat blood samples for which useable DNA was available for analysis. Participants gave a second specific consent for bio-specimens to be collected and stored.[173] Sample anonymity was assured as each specimen was identified only by a study ID barcode. Further, no subject identifying data was released to the laboratory involved, (Figure 2.5 and Figure 2.6).

Figure 2.5 Blood sampling and blood components

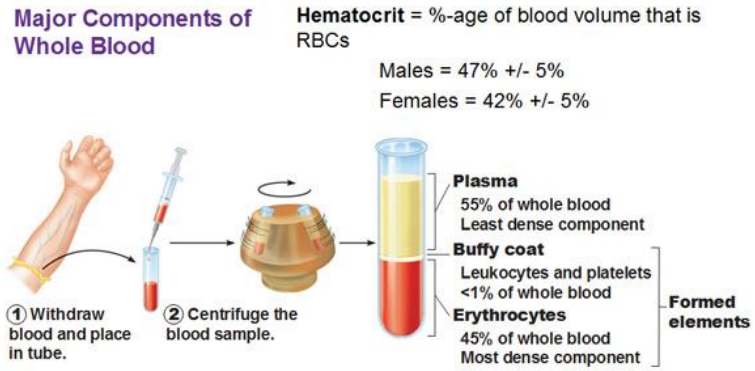
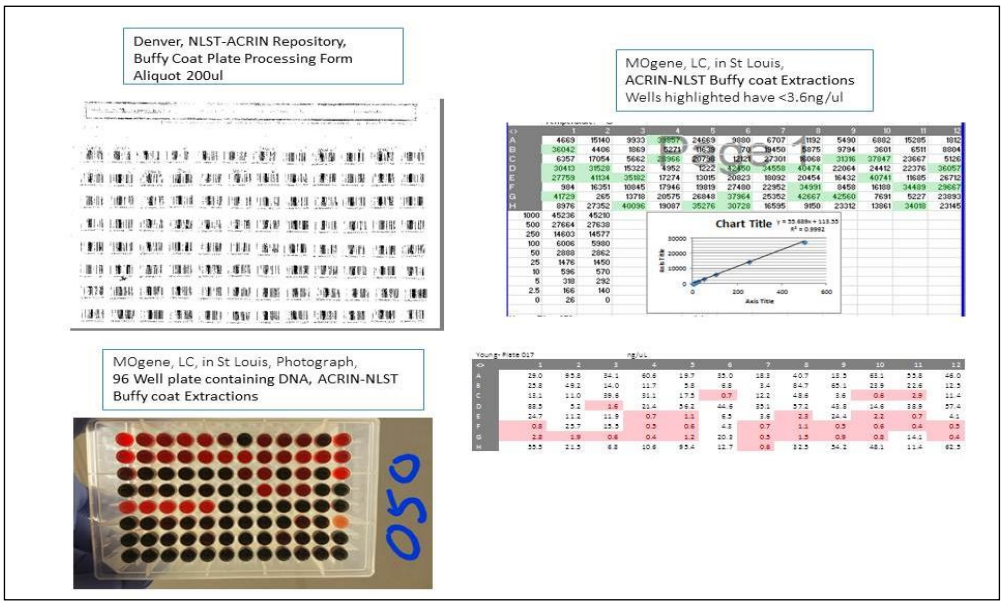


Figure 2.6 ACRIN-NZ Study DNA sample handling

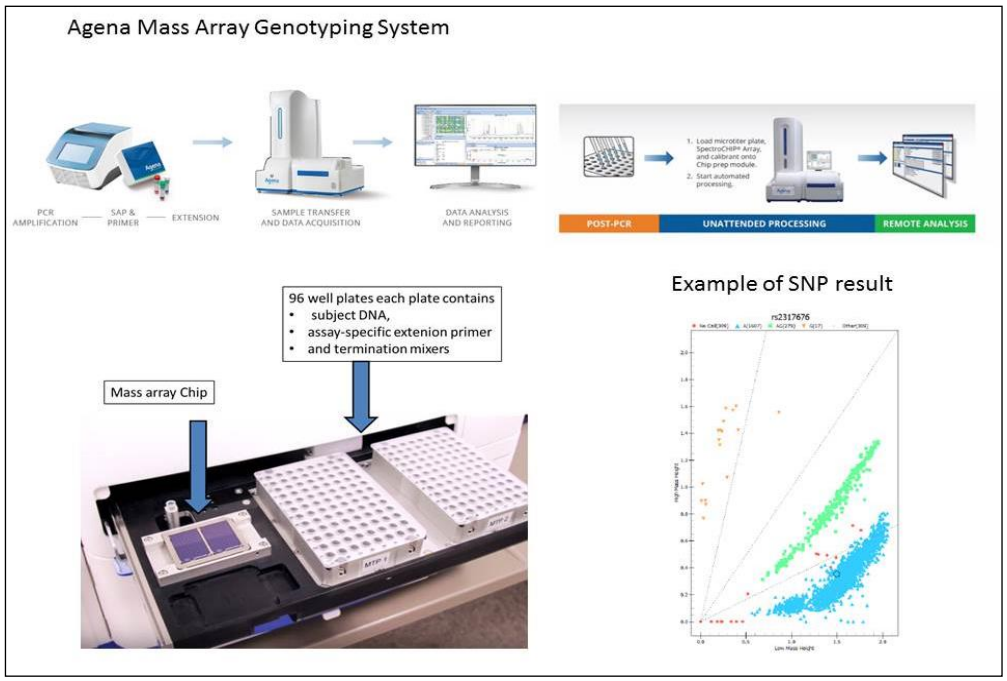


DNA extraction (isolation) and quantitation was performed by nan-o-drop and pico-green for each of the buffy coat sample by MOgene, LC, in St Louis, a US CLIA accredited laboratory. Genomic DNA was extracted by Qiagen column, and underwent a 'bead clean up' procedure to obtain the best quality DNA possible from each sample.

SNP genotyping was done on a Sequenom iPlex platform using iPLEX PRO, at Agena Bioscience, San Diego, a US accredited Sequenom-operated CLIA laboratory blind to the participant outcome and demographic data. The Sequenom sequences were designed in house by Agena with amplification and separation methods (iPLEX™, www.sequenom.com), (Figure 2.7).

Each genotyped assay result is called as either the major allele, heterozygous allele, or minor allele (Figure 2.8). For our genotyping analysis an assay call rate of no less than 95% per gene was required.

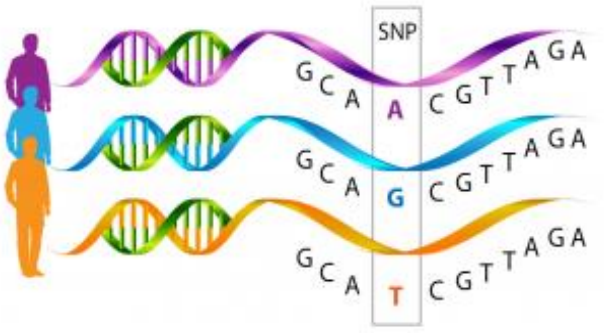
Figure 2.7 ACRIN-NZ Study, DNA Genotyping System (Sequenom Mass Spectrometer)



MassARRAY System 384/96

Dimensions (WxHxD)	18x43x27 in (45x110x68 cm)
Capacity	2 x Microtiter plates (96 or 384 well) 2 x SpectroCHIP 96 Arrays or 384 Arrays
Acquisition Speed	15 minutes for SpectroCHIP 96 Array 60 minutes for SpectroCHIP 384 Array

Figure 2.8 Illustration of single nucleotide polymorphism (SNP) variation in a gene sequence



Source: Source from http://cahartmanfiction.com/understanding-genetics-snps/

Once both the genotyping results data-base and the demographic data-base were cleaned they were combined in order to interrogate the wealth of information through numerous stratification strategies applied to the demographic variables drawn from the large comprehensive questionnaire completed by each participant during their initial enrolment and the subsequent follow-up outcome questionnaire completed over the active study years (average follow up 6.4 years). The enrolment questionnaire was completed prior to randomisation to either the CT or CXR arms. In this way interesting associations and relationships both expected and unexpected between the various components were uncovered. I have made numerous novel discoveries as well extending prior work from our laboratory, (Dr Robert Young, University of Auckland).

Chapter 3 Reduced Expiratory Flow Rate, Among Heavy Smokers Increases Lung Cancer Risk: Results from the NLST-ACRIN Cohort

RJ Hopkins, F Duan, C Chiles, EM Greco, GD Gamble, D Aberle, RP Young. 2017. 'Reduced Expiratory Flow rate among Heavy Smokers Increases lung cancer risk: Results from the National Lung Screening Trial-American College of Radiology Imaging Network Cohort. *Annals ATS* 2017; Volume 14(3): 392-402.'

Annals of the American Thoracic Society, is an official journal of the American Thoracic Society.

This paper attracted an Editorial:

NYT Tanner, NJ Pastis. Chronic Obstructive Pulmonary Disease as a Lung Cancer Risk: Worth Its Weight in "GOLD". *Annals ATS* 2017; Volume 14(3):309-310.

3.1 Introduction

3.1.1 Basic epidemiology

Over the preceding 30 years, both cross-sectional and prospective studies have consistently shown that the presence of chronic obstructive pulmonary disease (COPD), characterised by irreversible airflow limitation and reduced expiratory flow rates, bestows a greater risk of lung cancer than when normal lung function is maintained.[2,16,33-34,52-53,99,184-187] This association is independent of smoking history and is strongest when COPD is defined by measured spirometry with the control group being confirmed to have normal lung function.[2,34,53] In the population-based National Health and Nutrition Examination Survey (NHANES),[53] increasing severity of spirometry-defined COPD was associated with a greater risk of lung cancer. This study showed the risk associated with mild COPD was triple that seen in smokers with normal baseline lung function and 6 fold greater in those with moderate-severe COPD.[54] A similar magnitude of risk (4-6 fold) was found in a cross sectional study of moderate to heavy smokers with lung cancer when compared to a randomly selected group matched for age, gender and smoking histories.[34] When the lung

cancer risk conferred by forced expiratory volume in one second (FEV_1) was examined, an inverse dose-response relationship was found where even a small decrease in per cent predicted FEV_1 (<90%) was associated with an increased risk of lung cancer.[186] This increased risk extends to those with "restrictive" lung disease.[53] However in a recent study, using a clinic based COPD cohort, subjects with Global Initiative for Chronic Obstructive Lung Disease (GOLD) grade 1-2 were found to have a greater risk of lung cancer than those of GOLD grade 4.[187] This unexpected observation argues against a linear relationship between airflow limitation and lung cancer risk, leading to some debate.[188-189] Underlying this debate is the observation that reduced expiratory flow rates have also been associated with increased all-cause, lung cancer and cardiovascular mortality,[16] implicating a differential survival effect.[188] This debate is highly relevant to the current era of CT screening because risk stratifying eligible smokers is considered to be a key element of screening, where the risk of lung cancer must be balanced against the harm-to-benefit ratio of screening relative to dying of other causes.[99]

3.1.2 COPD link with lung cancer – possible mechanism

Why smokers with underlying airflow obstruction are at greater risk of lung cancer than smokers with normal lung function remains unknown.[54,60,190-191] Early hypotheses suggested patients with COPD have a greater accumulation of carcinogens in the airway.[2,33] Other hypotheses link lung cancer with underlying emphysema and lung remodelling.[54,190] The innate immune response has been implicated in both COPD and lung cancer in large prospective studies.[60,191-193] Systemic inflammation, reflecting innate immune hyper-responsiveness, has been linked to both progression of lung remodelling leading to COPD and DNA damage leading to lung cancer.[193] In 2009 it was suggested that the microclimate of remodelling underlying COPD, characterised by excess metalloproteinases and growth factors, might initiate premalignant transformation termed epithelial mesenchymal transition(EMT).[60] Since 2010 when EMT was first identified in patients with COPD,[194] it has been associated with worsening airflow limitation and has subsequently been reported by three further research groups using surgical samples from COPD patients with lung cancer.[195] EMT is considered a precursor to many other epithelial based cancers and shown to be potentially reversible with immune-modulatory treatment.[60]

3.1.3 Relevance to CT screening

Unbiased data on airflow limitation in lung cancer is limited because spirometry is not routinely performed in the work up of unscreened cases of lung cancer. With the recent interest in CT screening for lung cancer, the relationship between airflow limitation and lung cancer can be more robustly examined.[184-186] Using data from the American College of Radiology, Imaging Network (ACRIN) cohort of the National Lung Screening Trial (NLST), a large prospective study in 18,714 subjects, we have re-examined the relationship between GOLD-based airflow limitation and the development of lung cancer. Preliminary results of this study have been reported in abstract form.[37]

Hypothesis: In the context of high risk smokers, worsening airflow limitation increases the risk of lung cancer in a screening cohort.

Aim: To examine the relationship between worsening airflow limitation and lung cancer risk in the participants of the National Lung Screening Trial (NLST).

Objective: To stratify the participants of the NLST at baseline and then compare the rates of lung cancer that developed over the follow-up period.

3.2 Methods

3.2.1 Subjects

The recruitment and study design of the full National Lung Screening Trial (NLST) involving 53,452 screening participants, yielding 2058 histology confirmed lung cancers, has been described elsewhere.[93,171] In the American College of Radiology Imaging Network (ACRIN) cohort of the NLST, participants from 23 centres agreed to undergo baseline pre-bronchodilator spirometry (N=18,714). From this cohort, 768 histology-confirmed lung cancer cases were diagnosed over the study period of 7.5 years[93]

3.2.2 Pulmonary function testing

In the NLST-ACRIN cohort, pre-bronchodilator spirometry was measured at baseline screening (T0) in the majority of participants meeting the following criteria; no chest infection in the preceding 3 weeks and no use of a short-acting bronchodilator inhaler in the preceding 6 hours or long-acting bronchodilator in the preceding 24 hours. Those not

meeting these criteria were rescheduled for spirometry testing at a later visit. The spirometry was measured by trained staff using a Spiropro spirometer (eResearchTechnology, GmbH, Germany). The severity of airflow limitation was defined according to the Global Initiative on Chronic Obstructive Lung Disease (GOLD) criteria grades 1-4 (www.GOLD.org accessed February 24, 2016).

3.2.3 Lung cancer case rates

Lung cancer cases included all those diagnosed during the trial, whether screen or non-screen detected, or prevalent (diagnosed during the first year or T0) or incident lung cancers (diagnosed during subsequent years T1 to T6), (Table 3.1). All lung cancer cases were confirmed on histological sampling according to accepted international classification criteria.[37] We report the old Bronchioloalveolar cancer (BAC) subgroup separately from adenocarcinomas because they have been previously associated with "histology shift" and over-diagnosis.[93]

Lung function results and lung cancer histology results were available for 758 of the 768 lung cancer cases (99% of total). In a sensitivity analysis, where only spirometry meeting strict American Thoracic Society (ATS) criteria (Grade A) were included in the analysis (N=13,530, 72% of the total), the relationships identified in the larger cohort (N=18,714) were re-analysed.

For comparative purposes, the lung cancer incidence rates (IR, per 1000 person years) and incident rate ratios (IRR) in this NLST-ACRIN cohort were compared with those from two other prospective lung cancer studies, a non-screening study[53] and screening study,[185] where airflow limitation was assessed using spirometry at study baseline. In a further analysis we sub-grouped those with no airflow obstruction and compared lung cancer incidence rates in those with restrictive lung disease[53] and GOLD Unclassified (GOLD U),[196] relative to "healthy" smokers.

Table 3.1 Glossary

Term	Definition
Lung Cancer Prevalence	The number of lung cancer cases diagnosed as a percentage of the total number of screening participants including both prevalent lung cancers (diagnosed in the first year of screening – T0) and incident cancers (diagnosed during subsequent years T1 to T6). Described as a percentage.
Lung Cancer Incidence Rate (IR)	The number of lung cancer cases diagnosed over a defined period as a function of the total number of person years calculated according to the total number of screening participants and years of screening. Described as rate/1000 person years.
Lung Cancer Incidence Rate Ratio (IRR)	The ratio of one lung cancer incidence rate relative to a reference lung cancer incidence rate. Provides a crude estimate of relative risk but in the context of comparing two incidence rates rather than two prevalence rates.
Study year T0-T6	Study year in which a lung cancer was detected. T0-T2 = Active screening years (i.e., annual CT or Chest X-ray) T3-T7 = Surveillance/Follow up years

3.2.4 Statistical analysis

Differences in lung cancer incidence rates, stratified by severity of airflow limitation based on GOLD grade 1-4 were compared. Differences in lung cancer incidence rates (IRs) were compared using incidence rate per 1000 person years and incidence rate ratio (IRR) (Table 3.1). Differences in lung cancer histology were compared according to GOLD status using Fisher's Exact test and Mid-P exact test. Confidence intervals for IRs and IRRs were estimated using the exact method. Significance was defined as a two tailed *P* less than 0.05 (P<0.05). All statistical analyses were performed using SAS (v9.4, SAS Institute, Cary, NC) or STATA statistical software (Stata Corp, 2015, Stata Statistical Software; Release 14; College Station, TX).

3.3 Results

3.3.1 Demographics

Table 3.2 shows a comparison of the demographic characteristics of the full NLST trial cohort (N=53,452) and lung cancer cases (N=2,058) compared with the NLST-ACRIN cohort (N=18,714) and lung cancer cases (N= 768). The NLST-ACRIN Cohort participants are very similar to the full NLST study participants.[93] Based on the pre-bronchodilator pulmonary

function testing in the NLST-ACRIN cohort, 64.3% had no airflow limitation, a further 27.4% had airflow limitation approximating GOLD 1-2 COPD, 5.8% had GOLD 3 and 1.1% had GOLD 4 severity COPD (Table 3.2). There was missing lung function data for 239 participants and missing %predicted FEV_1 in 7 participants with airflow limitation.

3.3.2 GOLD grade and lung cancer risk

A comparison of the demographic variables, lung cancer rates and histology according to the presence or absence of airflow limitation is shown in Table 3.3. Regardless of the screening interval, airflow limitation was associated with a 2-fold greater lung cancer incidence rate (IR) than those with no airflow limitation (P<0.0001 for screening and follow-up intervals). Airflow limitation was also associated with significantly less Bronchioloalveolar carcinoma (BAC, now redefined as adenocarcinoma-in-situ or minimally invasive adenocarcinoma) and significantly more non-small cell lung cancer histology (Table 3.3). Airflow limitation was also associated with marginally more squamous cell cancers and marginally less adenocarcinoma. These differences are compared further in Table 3.4, where demographic variables, lung function, and cancer histology are compared across severity of airflow limitation based on GOLD grade.

3.3.3 Lung cancer incidence by GOLD

The lung cancer incidence (per 1000 person years) and incidence rate ratios (referenced against those with no airflow limitation or GOLD 1 airflow limitation), are shown in Table 3.5 and Figure 3.1. In the whole cohort (N=18,714) the incidence in each of the GOLD groups 1-4 were significantly greater than for the referent group with “No airflow limitation” and showed a linear relationship indicating that increasing severity of airflow limitation is directly associated with an increased risk of lung cancer (P<0.001 for trend). This linear relationship extended to include GOLD U (see Table 3.6 and Figure 3.2).

Figure 3.1 Comparison of Incidence rate (IR), Incidence rate ratios (IRR) and 95% CI in lung cancer according to severity of airflow limitation in the NLST-ACRIN cohort (A and B)

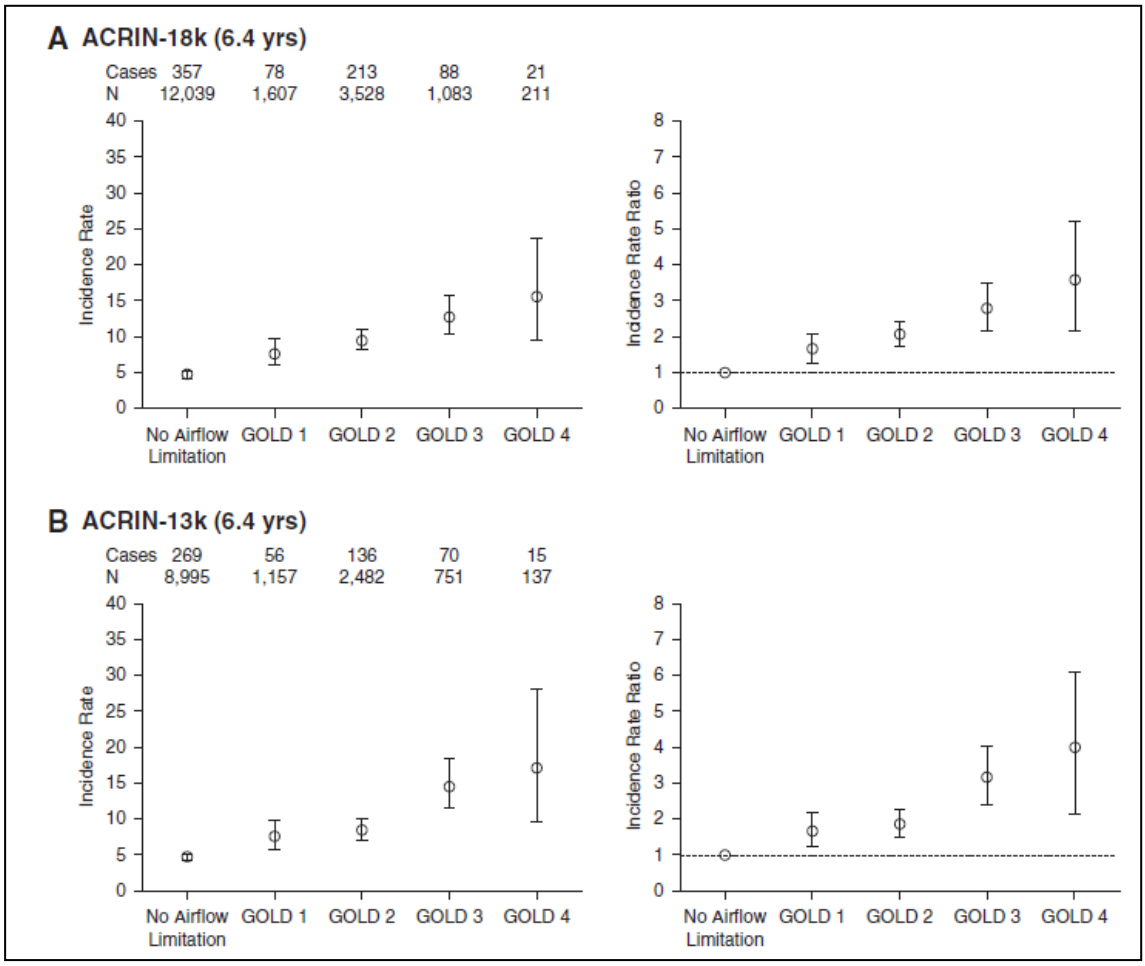
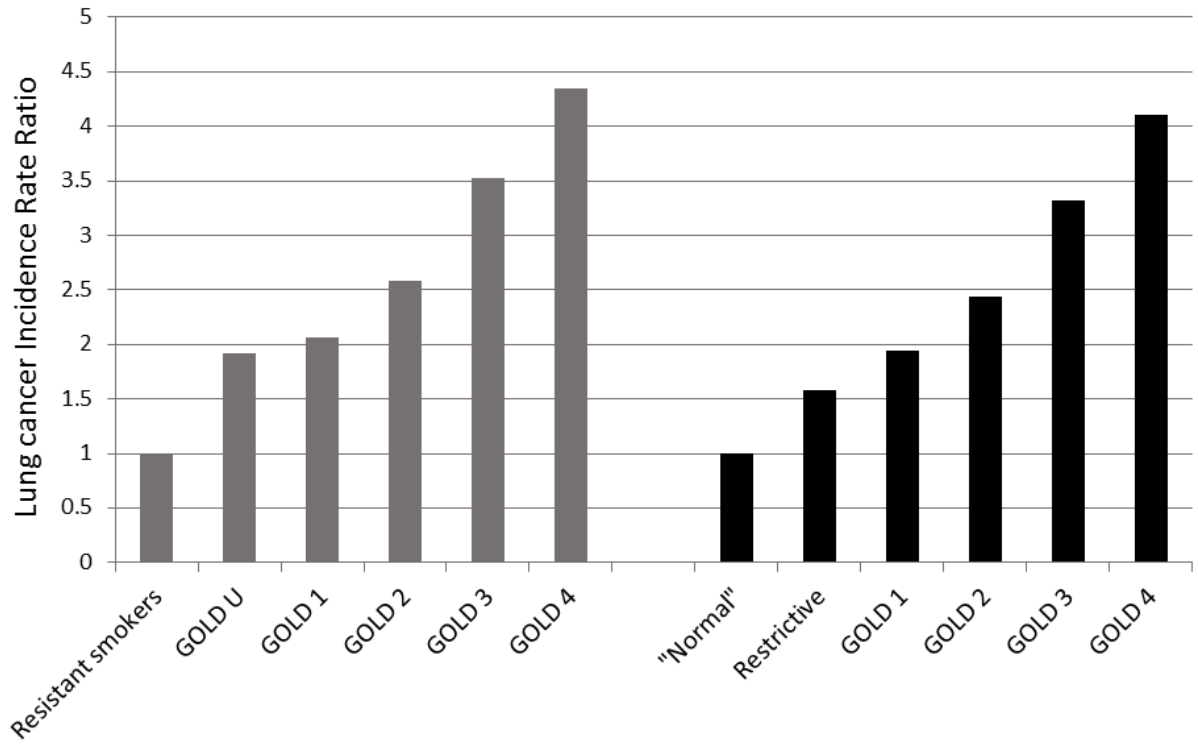


Figure 3.2 Incidence rate ratio of lung cancer according to GOLD severity with sub-phenotyping of those with no airflow limitation into those with "restrictive" spirometry[53] and GOLD U criteria[93]



In Table 3.5 and Figure 3.3, lung cancer incidence stratified by GOLD grade according to baseline spirometry is compared with two other prospective studies,[53,185] one a non-screening study[53] and the other a screening study.[185] Despite representing different risk populations and having different study designs the lung cancer incidence and thus risk, consistently shows a linear relationship with GOLD 3-4 subjects having a 3-6 fold greater risk for lung cancer than those with no airflow limitation and 2-3 fold greater risk for those with GOLD 2 COPD. We found the same relationship across GOLD groups regardless of whether lung cancer was diagnosed in the CT or CXR arms where the later better simulates unscreened lung cancers (Table 3.7, Table 3.8 and Figure 3.4). I also found that %predicted FEV_1 <90% was inversely related to lung cancer incidence (Figure 3.5). In addition, lung cancer rate ratios were increased for those with restrictive spirometry (Figure 3.2).

Figure 3.3 Comparison of Incidence rate (IR), Incidence rate ratios (IRR) and 95% CI in the lung cancer studies according to severity of airflow limitation in the NLST-ACRIN cohort (A and B), PLuSS[185] screening cohort (C) and COPD clinic population[187]

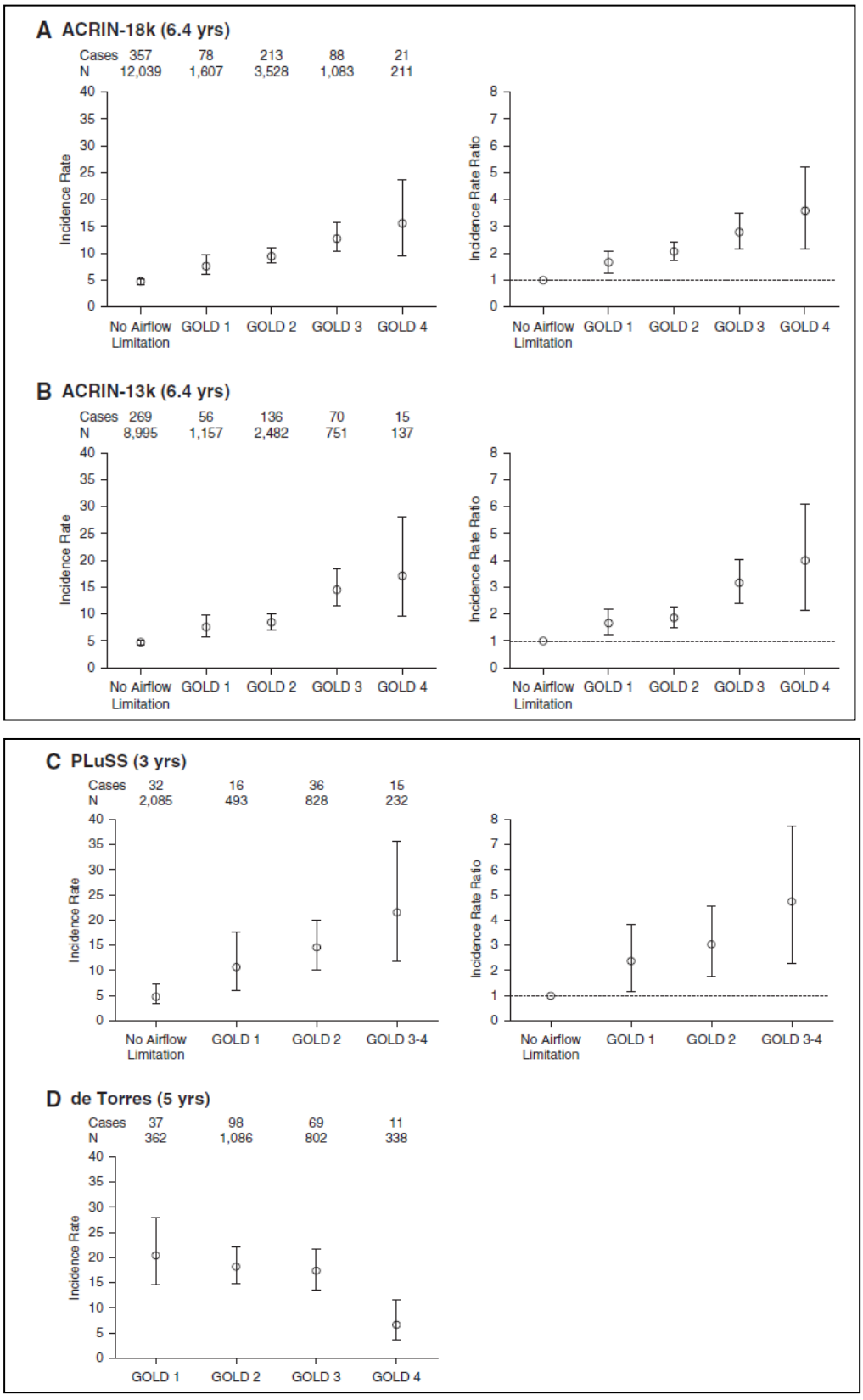


Figure 3.4 Lung cancer incidence rate (per 1000 person years) in the NLST-ACRIN sub-cohort according to pre-bronchodilator spirometry-defined GOLD grade (GOLD 1-4) in the full cohort (18K cohort, grey) and those who meet strict ATS recommendations (13k cohort, black)

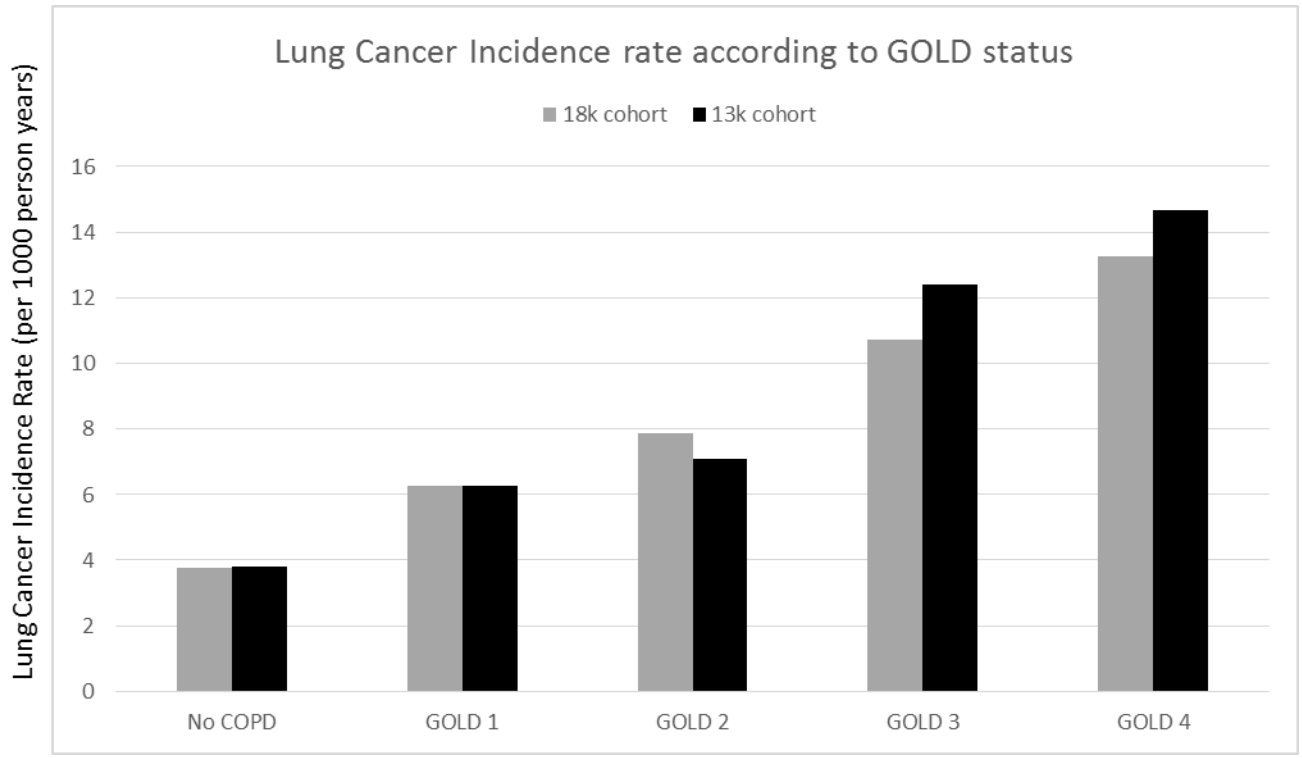
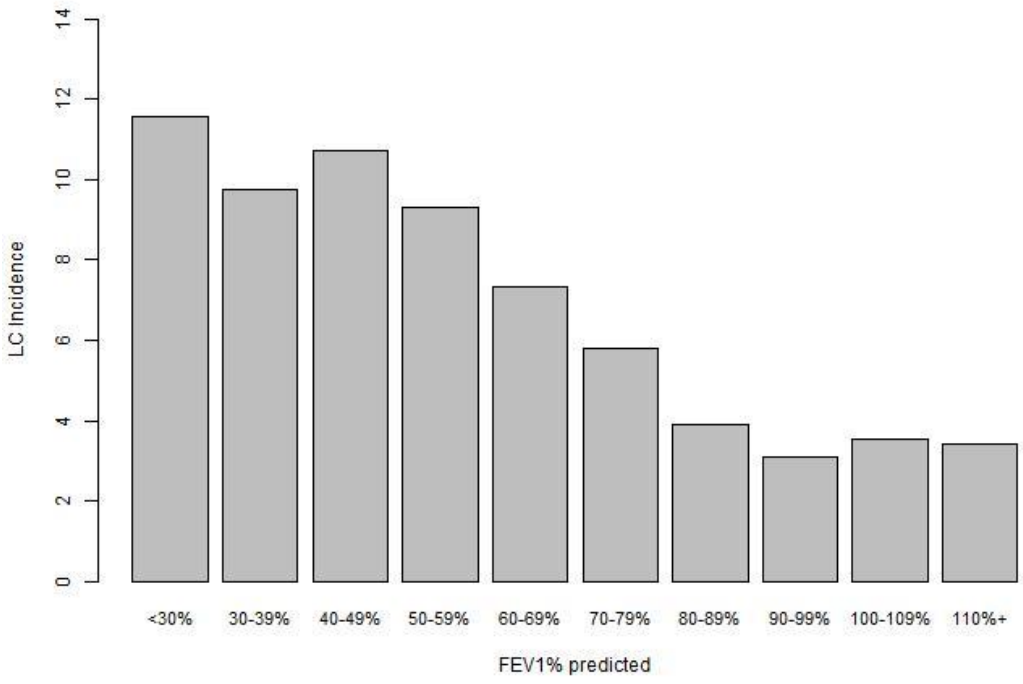
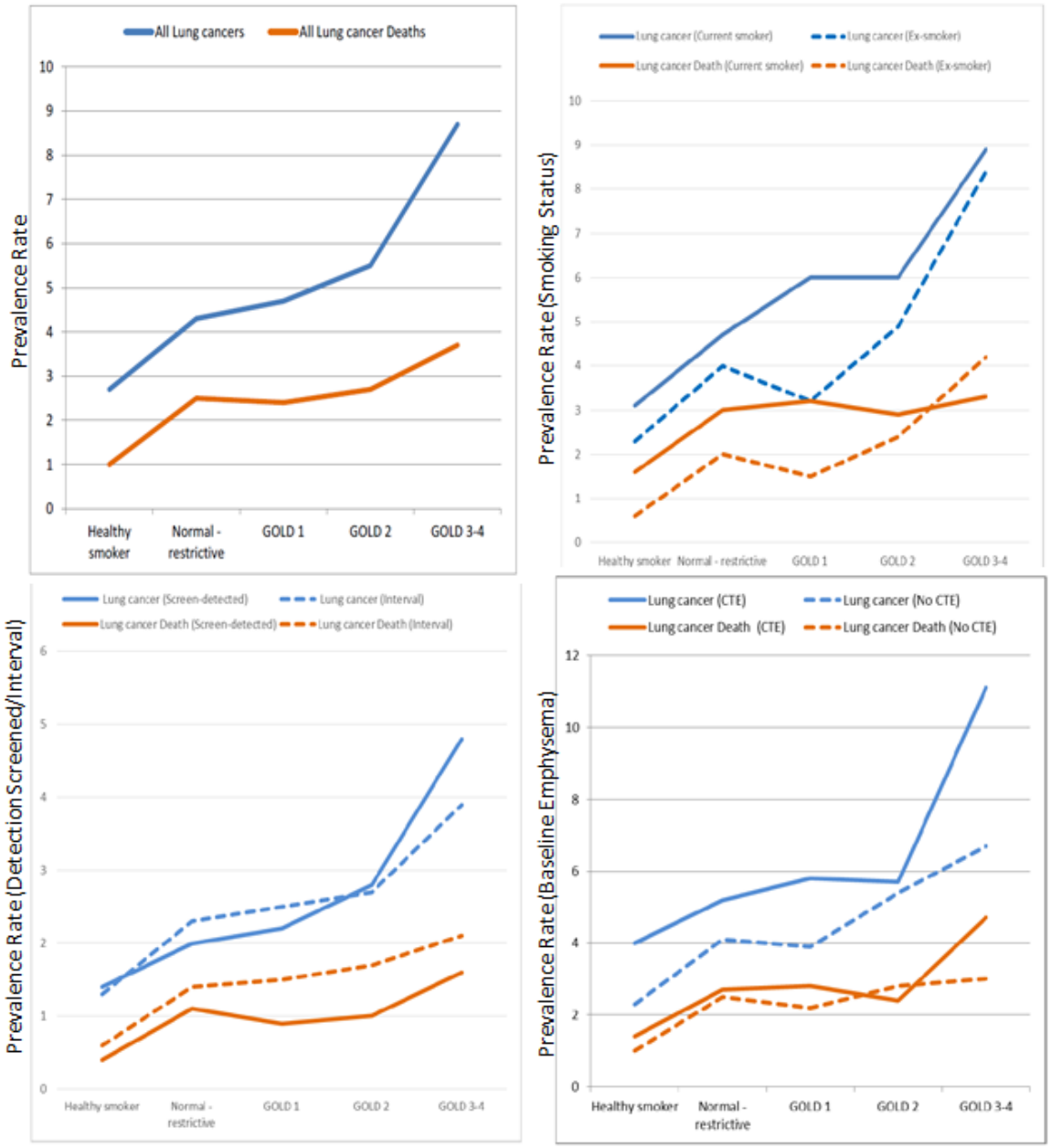


Figure 3.5 Lung cancer incidence (per 1000 person years) according to FEV1%predicted in the NLST-ACRIN sub-study (N=18,473)



Importantly, while lung cancer prevalence was linearly related to the severity of airflow obstruction, after stratification for smoking status, lung cancer detection status and radiological presence of emphysema at baseline, (Figure 3.6), this association was much reduced for lung cancer mortality.

Figure 3.6 Lung cancer prevalence (blue) and mortality (orange) as a percentage stratified by smoking status (Current vs Ex-smoker), lung cancer detection (Screen-detected vs Interval), and baseline emphysema (present (E+) vs absent (E-))



In a sensitivity analysis, when only those whose spirometry met strict ATS criteria were included (73% of the total, N=13,530), I found almost identical lung cancer incidence rates across the GOLD groups (see Table 3.5 and Figure 3.4). The relationship between severity of airflow limitation and risk of lung cancer was again linear (P<0.001 for trend).

Lastly, we assessed the relative contribution of age, pack years and %predicted FEV_1 to lung cancer risk in a multivariate analysis comparable to that reported by Tockman and colleagues.[33] This showed that %predicted FEV_1 was the single most important predictor of lung cancer relative to age or pack years (Table 3.9).

Table 3.2 Comparison of baseline demographics, between the full NLST study subjects, and the NLST-ACRIN cohort, where spirometry defines severity of airflow limitation, using GOLD grading

Screening Trial	NLST-Main Study (N=53,452)		NLST-ACRIN Cohort (N=18,714)	
Subject Demographics	Screening Participants	Lung cancer cases	Total Cohort	Lung cancer cases
Number	53,452	2,058	18,714	768
Mean Age (SD)	61.4 (5.0)	63.7 (5.3)	61.6 (5.0)	63.6 (5.2)
% Male	59%	60%	55%	56%
Mean pack years (SD)	56.0 (23.9)	64.9 (27.1)	55.9 (23.5)	63.9 (27.0)
% Current Smokers	48%	60%	50%	60%
Family History of Lung Cancer	22%	26%	23%	26%
Self-reported COPD#	17%	27%	20%	32%
Mean Body Mass Index (SD)	27.9 (5.0)	26.8 (4.7)	27.8 (5.1)	26.9 (4.9)
Pulmonary Function Tests				
Total§	ND	ND	18,714	768
GOLD 1	ND	ND	1607 (8.6%)	78 (10.2%)
GOLD 2	ND	ND	3528 (18.9%)	213 (27.7%)
GOLD 3-4*	ND	ND	1294 (6.9%)	109 (14.2%)
GOLD Unknown (due to missing height data)**			7 (<1%)	1 (<1%)
Airflow Limitation	ND	ND	6436 (34.4%)	401 (52.2%)
No Airflow Limitation	ND	ND	12,037 (64.3%)	357 (46.5%)
Missing spirometry data	ND	ND	241 (1.3%)	10 (1.3%)

NR=not reported. ND=not done. #Self-reported COPD in the NLST was based on questionnaire responses referring to the past diagnosis of COPD, emphysema, chronic bronchitis, or a combination of these. § Pulmonary function results were available for 99% of screening participants and lung cancer cases. *Stage 4 COPD - 1.1% in total cohort and 2.7% in lung cancer cases. ** Airflow limitation based on $FEV_1/FVC<0.70$ but %predicted FEV_1 not known.
GOLD 1 = $FEV_1/FVC<0.70$, FEV_1%pred≥80%, GOLD 2 = $FEV_1/FVC<0.70$, FEV_1%pred 50%-79%
GOLD 3 = $FEV_1/FVC<0.70$, FEV_1%pred 30%-49%, GOLD 4 = $FEV_1/FVC<0.70$, FEV_1%pred <30%

Table 3.3 Comparison of the demographics, lung cancer prevalence, lung cancer incidence rate and histology according to airflow limitation in the NLST-ACRIN cohort (N=18,473)

Characteristics	No Airflow Limitation	Airflow Limitation	Total	P value
Total	12,037 (65.2%)	6,436 (34.8%)	18,473	-
Mean Age (yrs), SD	61.0 (4.9)	62.5 (5.2)	61.6 (5.1)	< 0.0001
%Male	53%	60%	55%	< 0.0001
Mean Pack years, SD	53.8 (22.3)	59.7 (24.9)	55.8 (23.4)	< 0.0001
%Current smokers	47%	56%	50%	< 0.0001
Family Hx of lung Cancer	23%	24%	23%	0.55
Doctor diagnosed COPD	14%	31%	20%	< 0.0001
Mean BMI, SD	28.5 (5.2)	26.7 (4.8)	27.8 (5.1)	< 0.0001
Lung cancers by screening arm	357 (47%)	401 (53%)	758	<0.0001
CXR	164 (45%)	201 (55%)	365	0.27
CT	193 (49%)	200 (51%)	393	
Lung Cancer Prevalence# (%)	2.97	6.23	4.10	<0.0001
Lung Cancer Incidence rate (per 1,000 person yrs)				
-T0-T6 (Total study)	3.78	8.12	5.27	<0.0001[1]
-T0-T2 (Screening interval)	6.01	12.73	8.33	<0.0001[2]
-T3-T6 (Follow up interval)	2.36	5.14	3.31	<0.0001[3]
CXR	4.28	9.66		<0.0001[4]
CT	4.99	9.82		<0.0001[5]
Histology				
Small cell	51 (14%)	60 (15%)	111 (15%)	0.0035
Squamous cell	73 (20%)	95 (24%)	168 (22%)	
Adenocarcinoma	129 (36%)	127 (32%)	256 (34%)	
BAC‡	40 (11%)	19 (5%)	59 (8%)	
Large Cell	14 (4%)	16 (4%)	30 (4%)	
Non-small Cell	50 (14%)	81 (20%)	131 (17%)	
Other		3 (<1%)	3 (<1%)	
Total	357	401	758	

1 IRR=2.15 (95% CI= 1.86-2.48) **(T6 = combines follow up years 6 and 7)**
2 IRR= 2.12 (95% CI= 1.76-2.55)
3 IRR= 2.18 (95% CI= 1.72-2.76)
4 IRR=2.26 (95% CI= 1.84-2.77)
5 IRR= 1.97 (95% CI= 1.62-2.40)

‡ BAC have been redefined as adenocarcinoma–in-situ and minimally invasive adenocarcinoma and defined separately above from adenocarcinoma as they are representative of lung cancers identified by CT screening and associated with "histology shift" and "over-diagnosis" relative to CXR screening.[93]

Combines prevalent lung cancers (T0) and incident lung cancers (T1 to T6).

Table 3.4 Demographics and lung function at baseline in the full NLST-ACRIN cohort and those with lung cancer, according to severity of airflow limitation by GOLD

Cohorts	Presence of Airflow limitation and its severity					
Demographic Variables	No Airflow Limitation	GOLD 1	GOLD 2	GOLD 3-4	P Trend	P Trend GOLD[†]
Full cohort at baseline (N=18,466)	12,037	1,607	3,528	1,294		
Mean Age (SD)	61.0 (4.9)	62.2 (5.3)	62.5 (5.2)	63.2 (5.2)	< 0.0001	< 0.0001
% Male	53%	64%	58%	58%	< 0.0001	0.0006
Mean pack years (SD)	53.8 (22.3)	56.6 (23.0)	59.8 (24.7)	63.4 (27.2)	< 0.0001	< 0.0001
% Current Smokers	47%	57%	58%	53%	< 0.0001	0.0159
Family History of Lung Cancer	23%	24%	23%	25%	0.63	0.51
Self-reported COPD#	14%	17%	29%	55%	< 0.0001	< 0.0001
Mean Body Mass Index (SD)	28.5 (5.2)	25.9 (4.0)	26.9 (4.8)	26.9 (5.5)	< 0.0001	< 0.0001
Mean Observed FEV1 (SD)	2.7 (0.8)	2.9 (0.7)	2.1 (0.5)	1.2 (0.3)	< 0.0001*	< 0.0001*
Mean FEV1% predicted (SD)	89.2 (16.8)	91.1 (12.2)	65.9 (8.3)	38.8 (8.4)	< 0.0001*	< 0.0001*
Mean FEV1/FVC (SD)	78.7 (54.7)	65.4 (4.9)	61.5 (6.8)	48.3 (10.4)	< 0.0001*	< 0.0001*
Lung Cancer cases (N=757)	357	78	213	109		
Mean Age (SD)	62.8 (5.2)	65.2 (5.4)	64.0 (5.0)	64.3 (5.3)	0.0003	< 0.0001
% Male	52%	69%	58%	57%	0.0324	0.16
Mean pack years (SD)	60.9 (24.6)	64.5 (28.6)	63.4 (26.4)	69.6 (33.4)	0.0288	< 0.0001
% Current Smokers	56%	68%	67%	53%	0.0151	0.0413
Family History of Lung Cancer	27%	19%	25%	28%	0.52	0.42
Self-reported COPD#	22%	14%	39%	58%	< 0.0001	< 0.0001
Mean Body Mass Index (SD)	27.8 (5.1)	25.3 (3.7)	26.1 (4.8)	26.8 (4.7)	< 0.0001	< 0.0001
Mean Observed FEV1 (SD)	2.5 (0.8)	2.9 (0.9)	2.0 (0.5)	1.2 (0.3)	< 0.0001*	< 0.0001*
Mean FEV1% predicted (SD)	85.1 (17.9)	93.8 (19.1)	64.4 (8.2)	38.5 (8.2)	< 0.0001*	< 0.0001*
Mean FEV1/FVC (SD)	76.9 (7.4)	64.9 (5.5)	60.8 (6.9)	49.2 (8.2)	< 0.0001*	< 0.0001*
Small cell (%)	14%	12%	17%	14%	0.70**	0.55**
Squamous cell (%)	20%	18%	23%	28%	0.27	0.26
Adenocarcinoma (%)	36%	36%	32%	28%	0.43	0.56
Bronchioloalveolar (%)[‡]	11%	9%	4%	4%	0.003	0.19
Large cell (%)	4%	6%	5%	1%	0.20	0.10
Non-Small cell (%)	14%	19%	19%	24%	0.08	0.55
Other (%)	0%	0%	1%	1%		

[#]Self-reported (doctor diagnosed) COPD in the NLST was based on questionnaire. [†]Testing difference between GOLD categories only. *Please note that as GOLD was defined by spirometry parameters, differences are highly significant ** p values from the Exact test of the association between each histology (yes/no) and COPD groups. The significance level was adjusted via Bonferroni correction to be 0.05/6=0.008 to control for multiple comparisons. ‡ BAC have been redefined as adenocarcinoma–in-situ and minimally invasive adenocarcinoma.

Table 3.5 Lung cancer Incidence Rate (IR, per 1000 person years) and Incidence Rate Ratio (IRR) according to severity of airflow limitation in screening and non-screening prospective studies[53,185,187]

Study (Follow-up)	Total Cohort	Lung Cancer Cases	Absolute Incidence		Rate versus no airflow limitation		Rate versus GOLD 1 airflow limitation	
	N	N	IR	IR 95% CI	IRR[1]	IRR 95% CI	IRR[2]	IRR 95% CI
ACRIN-18K (6.4 yrs)	18,466	758						
No Airflow limitation	12,037	357	3.78	3.40-4.19	Ref	-	-	-
GOLD 1	1,607	78	6.27	4.96-7.82	1.66†	1.28-2.12	Ref	-
GOLD 2	3,528	213	7.86	6.85-8.99	2.08*	1.75-2.47	1.25	0.96-1.65
GOLD 3	1,083	88	10.71	8.60-13.18	2.83*	2.22-3.89	1.71†	1.25-2.35
GOLD 4	211	21	13.25	8.22-20.19	3.51*	2.14-5.44	2.11‡-	1.24-3.46
Airflow limitation	6,429	401	8.11	7.34-8.94	2.14*	1.85-2.48		-
ACRIN-13K (6.4 yrs)	13, 522	546						
No Airflow limitation	8,995	269	3.81	3.37-4.30	Ref			
GOLD 1	1,157	56	6.25	4.73-8.11	1.64‡	1.21-2.19	Ref	-
GOLD 2	2482	136	7.10	5.96-8.39	1.86*	1.50-2.30	1.14	0.83-1.58
GOLD 3	751	70	12.39	9.67-15.62	3.25*	2.46-4.24	1.98†	1.37-2.87
GOLD 4	137	15	14.68	8.24-24.09	3.85*	2.12-6.46	2.35 ‡	1.23-4.21
Airflow limitation	4,527	277	7.96	7.06-8.96	2.09*	1.76-2.48	-	-
PLuSS (3 yrs)[185]	3,638	99						
No Airflow limitation	2085	32	5.12	3.50-7.22	Ref			
GOLD 1	493	16	10.82	6.18-17.57	2.12	1.16-3.85	Ref	-
GOLD 2	828	36	14.49	10.15-20.06	2.83*	1.76-4.56	1.34	0.74-2.41
GOLD 3-4	232	15	21.55	12.05-35.55	4.21*	2.28-7.78	1.99	0.99-4.03
Airflow Limitation	1,553	67	9.07	7.37-11.04	1.77‡	1.19-2.64	-	-
NHANES (17.9 yrs)[53]	5,402	291						
Normal lung function	2,359	165	0.92	0.66-1.26	Ref			-
Mild COPD	296	14	2.64	1.44-4.43	2.86 ‡	1.55-5.27	Ref	
Mod/Severe	393	38	5.40	3.82-7.41	5.85*	3.74-9.14	2.86†	1.55-5.27
All COPD	689	52	4.22	3.15-5.53	4.57*	3.01-6.91	-	-
Restrictive	300	12	2.24	1.13-3.90	2.42‡	1.27-4.62	-	-
Non-smokers	2054	10	0.27	0.13-0.50	0.29†	0.15-0.59		
de Torres (5 yrs)[187]	2,588	215						
GOLD 1	362	37	20.40	14.60-27.88	-	-	Ref	
GOLD 2	1086	98	18.10	14.73-21.90	-	-	0.88	0.60-1.29
GOLD 3	802	69	17.21	13.49-21.65	-	-	0.84	0.56-1.26
GOLD 4	338	11	6.51	3.42-11.31	-	-	0.31†	0.16-0.62
All COPD (100%)	2,588	215	16.62	14.50-18.95	-	-	-	-

Incidence rate ratio[1] (IRR) compares GOLD groups with No Airflow Limitation (referent)
Incidence rate ratio[2] (IRR) compares GOLD groups with GOLD 1 or Mild COPD (referent)
*P<0.0001, †P<0.001, ‡<0.01, 95% CI= 95% Confidence Intervals

Table 3.6 Incidence Rate Ratio of lung cancer according to GOLD severity, with sub-phenotyping of those with no airflow limitation into those with "restrictive" spirometry[53] and GOLD U criteria[196]

Sub-phenotype	Total Cohort	Lung Cancer Incidence Cases			Rate versus Healthy smokers		Rate versus GOLD U or restrictive	
	N	N	IR	IR 95% CI	IRR[1]	IRR 95% CI	IRR[2]	IRR 95% CI
ACRIN-18K (6.4 yrs)	18,466	757						
Healthy smokers	8845	212	3.04	2.65-3.48	Ref	-	-	-
GOLD U[3]	3192	145	5.85	4.94-6.88	1.92*	1.55-2.39	Ref	-
GOLD 1	1607	78	6.27	4.96-7.82	2.06*	1.57-2.68	1.07	0.80-1.42
GOLD 2	3528	213	7.86	6.85-8.99	2.58*	2.13-3.14	1.34‡	1.08-1.67
GOLD 3	1083	88	10.71	8.60-13.18	3.52*	2.71-4.53	1.83*	1.39-2.40
GOLD 4	211	21	13.25	8.22-20.19	4.35*	2.64-6.83	2.26‡	1.36-3.59
Healthy smokers	8436	214	3.23	2.81-3.69	Ref	-	-	
Restrictive[4]	3601	143	5.10	4.30-6.00	1.58*	1.27-1.96	Ref	-
GOLD 1	1607	78	6.27	4.96-7.82	1.94*	1.48-2.53	1.23	0.92-1.63
GOLD 2	3528	213	7.86	6.85-8.99	2.44*	2.01-2.96	1.54†	1.24-1.92
GOLD 3	1083	88	10.71	8.60-13.18	3.32*	2.56-4.27	2.10*	1.59-2.76
GOLD 4	211	21	13.25	8.22-20.19	4.11*	2.49-6.44	2.60†	1.56-4.13

1 Incidence rate ratio (IRR) (per 1000 person years) referenced against "Healthy smokers"

2. Incidence rate ratio (per 1000 person years) referenced against GOLD U or Restrictive

3. GOLD Unclassified = $FEV_1/FVC \geq 0.70$ and FEV_1%predicted <80%

4. Restrictive = = $FEV_1/FVC \geq 0.70$ and FVC% predicted<80%

*P<0.0001, †P<0.001, ‡<0.01, 95% CI= 95% Confidence Intervals

Table 3.7 Lung cancer prevalence according to Airflow limitation or GOLD status and screening arm (CT vs CXR) in the NLST-ACRIN cohort (N=18,475)

COPD status	CT arm		CXR arm		Total	
	LC/Total	Lung Cancer Prevalence	LC/Total	Lung Cancer Prevalence	LC/Total	Lung Cancer Prevalence
GOLD 1	44/796	5.53%	34/811	4.19%	78/1607	4.85%
GOLD 2	99/1731	5.72%	114/1797	6.34%	213/3528	6.04%
GOLD 3	43/541	7.95%	45/542	8.30%	88/1083	8.13%
GOLD 4	13/110	11.82%	8/101	7.92%	21/211	9.95%
Airflow Limitation	200/3183	6.28%	201/3253	6.18%	401/6436	6.23%
No Airflow limitation	193/6047	3.19%	164/5990	2.74%	357/12037	2.97%
Total	393/9230	4.26%	365/9243	3.95%	758/18473	4.10%

Table 3.8 Unadjusted and adjusted odd ratios from multiple logistic regression for lung cancer associations with CT-based emphysema (curvilinear relationship), and GOLD-based airflow limitation (linear relationship), from the PLuSS, (Table 3 in ref 185).

			Unadjusted		**Adjusted[1]**		**Adjusted[2]**	
	Cases	*Non-cases*	*OR*	*95% CI*	*OR*	*95% CI*	*OR*	*95% CI*
Airflow Obstruction								
None	32	2053	Ref		Ref		Ref	
GOLD 1	16	477	2.15	1.17-3.95	1.66	0.89-3.11	1.13	0.59-2.17
GOLD 2	36	792	2.92	1.80-4.73	2.11	1.27-3.49	1.47	0.87-2.50
GOLD 3-4	15	217	4.43	2.36-8.32	2.86	1.48-5.53	1.87	0.92-3.80
Radiographic Emphysema								
None	24	2068	Ref		Ref			
Trace	22	663	2.86	1.59-5.13	2.58	1.43-4.66	2.48	1.37-4.49
Mild	37	493	6.47	3.83-10.9	5.04	2.94-8.62	4.43	2.53-7.79
Mod-Severe	16	315	4.38	2.30-8.33	3.20	1.65-6.23	2.56	1.26-5.20

1 = adjusted for age, gender, and smoking, 2 = adjusted for age, gender and smoking in addition to CT-emphysema for airflow limitation and GOLD for CT-based emphysema.

Table 3.9 Univariate and multivariate logistic regressions predicting lung cancer according to age, pack years and FEV_1% predicted in the NLST-ACRIN cohort (N=18,473)

Model	**FEV_1%predicted[1]**		**Age[2]**		**Pack years[3]**	
Univariate	*Odds ratio*	*P value*	*Odds ratio*	*P value*	*Odds ratio*	*P value*
FEV1%predicted	2.29	<0.0001				
Age			2.13	<0.0001		
Pack Years					1.82	<0.0001
Multivariate						
FEV_1%pred and age	2.13	<0.0001	1.99	<0.0001		
FEV_1%pred and pack years	2.15	<0.0001			1.71	<0.0001
FEV_1%pred, age and pack years	2.02	<0.0001	1.90	<0.0001	1.62	<0.0001

[1]Odds ratio comparing FEV_1% predicted ≤60% vs. >60%.
[2]Odds ratio comparing age > 59 years vs. ≤ 59 years.
[3]Odds ratio comparing pack years > 49 pack years vs. ≤ 49 pack years.

3.4 Discussion

In an analysis of the NLST-ACRIN cohort, we found a direct linear relationship between increasing severity of airflow limitation, according to GOLD 1-4 spirometry grade, and increasing risk of lung cancer. To our knowledge, this is the largest prospective study to date that has looked at this relationship. In a second analysis (Figure 3.2), we found this linear relationship extended to smokers meeting GOLD U and restrictive spirometry,[53,196] with a magnitude comparable to those with airflow limitation of GOLD 1-2 severity. This study also shows that not only is a reduced expiratory flow rate an important risk variable for developing lung cancer among heavy smokers, it is one of the most important risk variables relative to those of age and pack years (Table 3.2, Table 3.3, Table 3.8).[2,33-34,186,197] This suggests that airflow limitation represents a global barometer of susceptibility to the key smoking related lung diseases including lung cancer.[20]

3.4.1 GOLD grade and linear relation with lung cancer risk

The results reported here, showing that the severity of airflow limitation according to GOLD grade correlates in a linear relationship with lung cancer risk, are consistent with most other studies.[53,185-186] This finding extends to even mild-moderate reductions in airflow limitation (Figure 3.4) according to %predicted FEV_1,[186] a finding confirmed in a meta-analysis.[131] These

findings contrast with those reported by de Torres and colleagues,[87] who found that lung cancer risk is greatest in those with mild to moderate COPD. Strengths of the NLST are its prospective design, large sample size and focus on older asymptomatic smokers, otherwise eligible for CT screening. This contrasts with the participants studied by de Torres, who were symptomatic COPD patients of greater severity. In the de Torres study, the risk of lung cancer was greatest in those who had higher FEV_1 but lower diffusing capacity for carbon monoxide (DLCO) and lower BMI.[187] In their study, variables such as age and pack years were no longer a risk factor for getting lung cancer. These findings are hard to reconcile with other larger prospective studies that report age, pack years and worsening COPD, as the strongest risk variables for lung cancer.[33,53,86,198] One possible explanation for these conflicting results is the very high proportion of COPD patients with GOLD 3-4 severity present in the clinic population reported in the de Torres study (44%) compared to participants in the NLST (20%) and PLuSS (15%) studies. This suggests that the clinic population while being representative of symptomatic severe COPD, and very likely worse emphysema, are not representative of the larger group of relatively unselected smokers at risk of lung cancer. This is important for two reasons. First, while cancer is one of the most important causes of death among COPD patients with mild disease (40-50%), respiratory disease, cardiovascular disease and non-cancer-related causes are more common in moderate-to-severe COPD (where cancer accounts for only 20% of deaths).[200] It has been argued that in a clinic COPD population, the frequency of lung cancer was higher in those with mild COPD secondary to a survivor effect where those with GOLD 3-4 severity are dying before cancer develops.[188,189] Unlike the strong linear association between severity of airflow limitation and risk of lung cancer (Figure 3.2 and Figure 3.3), the association with lung cancer mortality is substantially weakened (Figure 3.6). This contrasts with the strong linear association we found between severity of airflow limitation and non-lung cancer related mortality in this cohort. Second, in PLuSS, the odds ratio for lung cancer shows a linear relationship for GOLD 1-4 (Table 3.5) but a curvilinear (nonlinear) relationship for CT-based emphysema severity (semi-quantitative), where the highest risk for lung cancer was conferred by mild emphysema discussed further below, (Table 3.8).[185] It is therefore possible that the GOLD 1-2 group, identified by de Torres to be at greatest risk of lung cancer, had a greater prevalence of mild emphysema than our screening participants, and more severe emphysema overall. This is consistent with their findings that low DLCO and low BMI also confer increased lung cancer risk.[187] Lastly in the de

Torres study,[187] over 90% of subjects were men and given gender differences in susceptibility, aero-pollutant exposures, COPD prevalence and histological subtypes, their results may consequently be specific to the cohort they studied.

Prevalence of COPD in case-control studies

The results of the current study are notable in showing that the prevalence of airflow limitation in the NLST-ACRIN screening cohort of high risk smokers (35%) was substantially higher in those who went on to develop lung cancer (53%). This means that past lung cancer epidemiological studies, that have almost exclusively failed to account for this difference in prevalence of airflow limitation (or COPD), may spuriously report associations with lung cancer that result from confounding effects with COPD.[36,153,166,201]

3.4.2 COPD and CT-quantified emphysema in lung cancer risk

There remains debate about the respective utility of using airflow limitation or CT-quantified emphysema in assessing the risk of lung cancer.[132,184-185,202-207] Several studies have concluded that the risk of lung cancer was more robust using semi-quantitative assessment of emphysema (visual grading) than with automated emphysema scoring methods.[206,207]

One study found that severe airflow limitation was associated with increased lung cancer risk but not CT-quantified emphysema[202] although others have found the contrary.[184,185] In the Pittsburgh Lung Cancer Screening Study (PLuSS),[185] the investigators concluded that emphysema, based on semi-quantitative scoring (None, Trace, Mild and Moderate-severe) was more strongly correlated with risk of lung cancer than airflow limitation (Table 3.7). However for CT-based emphysema, there is a curvilinear relationship with risk of lung cancer that weakened (but remained significant) after adjustment for airflow limitation based on GOLD.[185] In contrast for airflow limitation based on GOLD, the linear relationship with risk of lung cancer was weakened and lost significance after correction for emphysema severity.[185] These findings together with the observation that the presence of CT-based emphysema in the absence of airflow limitation is associated with lung cancer, lead these investigators to conclude that emphysema is the more relevant “COPD-related” phenotype linked to increased risk of lung cancer.

However, while this may be true, the results of my current study show that airflow limitation has a simple linear relationship with risk of lung cancer that remains clinically useful in the setting of lung cancer risk assessment. This is the case whether the severity of airflow limitation is graded according to GOLD criteria (Figure 3.2 and Figure 3.3) or assessed according to FEV_1%predicted (Figure 3.5). This means that airflow limitation (or FEV_1%predicted) can be included in lung cancer risk models as a continuous variable, across the spectrum of severity, unlike CT-quantified emphysema which cannot because of its curvilinear relationship with lung cancer (Table 3.8). The utility of using airflow limitation as a risk variable also stems from the clinical perspective that it is much simpler and cheaper to perform spirometry than a CT scan.

The most interesting result from the PLuSS was that the smokers with CT-based emphysema alone, GOLD 1-2 airflow limitation or GOLD 3-4 airflow limitation were 3-fold, 4 fold and 6 fold respectively, more likely to develop lung cancer than those with "Normal lungs". This suggests that airflow limitation and emphysema severity can be combined in order to best identify smokers at greatest and lowest risk of lung cancer. In this respect, our group have previously shown that 85% of all lung cancers in the PLuSS study had one or other of these "COPD phenotypes".[99] We estimate 4-fold more people with "normal lungs" will need to be screened to achieve comparable lung cancer diagnostic rates to those with mild-moderate (GOLD 1-2) COPD. The curvilinear relationship between CT-based emphysema and risk of lung cancer, where mild emphysema confers greater risk than trace or moderate-severe emphysema,[185] may explain why studies do not consistently report an association between CT-based emphysema and risk of lung cancer.[132,202-207] This inconsistency may also reflect the differences in scoring emphysema severity using semi-quantitative methods compared to automated scoring systems.[207]

This nonlinear relationship between CT-based emphysema and risk of lung cancer might also underlie why the study of de Torres[187] found greater lung cancer incidence in those with mild airflow limitation (GOLD 1-2), where our group hypothesise that the symptomatic clinic patients have a greater prevalence of mild emphysema than the asymptomatic screening participants. If such a hypothesis were true, then a combined approach to assessing risk of lung cancer might be used where a reduced DLCO[187] may be used instead of a CT scan to assess for the presence and severity of emphysema. Such an approach would also identify

those with smoking-related restrictive lung disease also associated with an increased risk of lung cancer (Table 3.5, Table 3.6).[53] As previously stated, the presence of emphysema was only reported as 'yes' or 'no' in this current study (i.e. dichotomised), while other COPD-related phenotypes such as airway thickness, airway diameter, gas trapping or interstitial changes were not routinely collected and thus not available for analysis in this cohort. That said the impact of airflow limitation on outcomes in CT screening in addition to lung cancer incidence is the subject of further analyses.

This study shows the significant differences in lung cancer incidence and lung cancer histology according to the presence and severity of airflow limitation (Table 3.3 and Table 3.4). The estimated annual lung cancer incidence, regardless of screening interval, was two-fold greater in participants with airflow limitation (GOLD 1-4) compared to those with normal lung function (Table 3.3).

Our results extend those of other studies showing worsening airflow limitation is also associated with more aggressive lung cancer histological subtypes (i.e., less BAC/Adenocarcinomas and more Squamous/Non-Small cell cancers),[93,127-128,208] most notable in those with severe disease (Table 3.4). In a secondary analysis, we also show that smokers meeting GOLD U criteria (also labelled as SGU or Preserved Ratio Impaired Spirometry (Prism)), have an increased risk of lung cancer relative to the healthy smoker (nearly 2-fold) and comparable to those with GOLD 1-2 airflow limitation. In the latter smokers, with only mild reductions in expiratory flow rates, more than 70-80% of them do not know they have airflow limitation (COPD) conferring a greater risk of lung cancer.

A further finding in this study is the difference in histology according to the presence of airflow limitation. In those with airflow limitation at baseline, there were significantly fewer bronchioloalveolar cell cancers (Table 3.3 and Table 3.4). There was also a trend towards fewer adenocarcinomas and more squamous cell carcinomas consistent with other studies. These differences in lung cancer histology, according to the presence of airflow limitation, were also evident with increasing severity of airflow limitation according to GOLD grade, (Table 3.4).

For subjects with more severe airflow limitation (GOLD grade 3-4), we found less BAC and adenocarcinomas, and more squamous cell and non-small cell lung cancers. This may explain

the observation that COPD status is associated with more aggressive lung cancers.[127,208] We and others have shown similar results in studies where lung cancers with a shorter volume doubling time, (faster growth rate), were more prevalent in those with impaired lung function consistent with COPD.[127,128] Moreover in the current study, in those who developed lung cancer but had airflow limitation at baseline, there were no excess cancers between the CT and CXR arms, there was comparable histology and a significant stage shift in favour of early stage cancer over late stage cancer.[93] In those lung cancers found in screening subjects with no airflow limitation at baseline, we found an excess of cancers attributed entirely to early stage cancers of BAC histology. The importance of our observation that airflow limitation is associated with more aggressive lung cancer in the context of CT screening may have a bearing on the attenuated lung cancer-specific mortality found in this group,[211] and remains the subject of further investigation.

3.4.3 COPD and lung cancer screening

With the recent interest in CT screening for lung cancer, there is growing interest in better defining those smokers at greatest risk to maximise the benefit and minimise the harms from screening.[99] Our results confirm that smokers with underlying airflow limitation are at greater risk of lung cancer and this is associated with a 2-4 fold greater lung cancer incidence in screening studies.[99,209] In a simulation study, Lowry and colleagues examined the effect of competing causes of death in CT screening participants and concluded those with COPD may disproportionately benefit from lung cancer screening.[210] However, some argue that screening patients with COPD may not necessarily translate into mortality benefit due to factors such as competing causes of premature death.[211] In a preliminary analysis of the NLST-ACRIN cohort we have found that lung cancer specific mortality reduction was attenuated in those with airflow limitation (15%, $P>0.05$) and enhanced in those with no airflow limitation (28%, $P<0.05$).[80] We suggest that in those with airflow limitation, more aggressive lung cancers (see above) or competing cause of premature death from causes other than lung cancer may indeed be an issue.[80,188] This also suggests that those at greatest risk of lung cancer may not necessarily achieve the greatest benefit from screening.[95,211] This is important because risk models for lung cancer share the same variables associated with COPD,[131] often including a past diagnosis of COPD,[86,198] or the presence of airflow limitation based on spirometry.[199] Personalised risk prediction is a recognised feature of CT screening

for lung cancer and helps assess the relative benefits and harms of screening.[110] Our results in the NLST-ACRIN cohort not only suggests the severity of airflow limitation confers differential effects on lung cancer risk and histology,[93] but that it may also affect outcomes according to lung cancer specific and all-cause mortality.[80] The latter is termed the "competing cause of death" effect and is beyond the scope of this study. However, this effect has relevance to optimising the selection of "healthy" smokers for lung cancer screening (the "sweet spot"),[117] and is the subject of further investigations in this cohort.

3.4.4 Strengths and limitations

There are several limitations to this study. First, while comparable with other CT screening studies,[185,212-214] spirometry in the NLST was performed as a pre-bronchodilator measurement rather than post-bronchodilator as recommended for clinical purposes. This means a proportion of subjects we have classified as having airflow limitation may have had asthma or COPD-asthma overlap rather than COPD (i.e., full, partial, or minimal reversible airflow limitation respectively), although this is likely to be only a small proportion of this elderly cohort of chronic smokers (mean age 64 years and minimum 30 pack years of smoking). A further limitation of the study is that people with a life expectancy of less than 5 years were excluded from the NLST.[37] This might explain the relatively low prevalence of GOLD grade 3 and 4 subjects in this study. A third limitation of this study is that we do not yet have any measure of emphysema severity in the NLST-ACRIN screening participants so that the relationship between emphysema score (severity) and lung cancer incidence could not be examined across the screening participants. Moreover, other CT-based phenotypes such as airway size, gas trapping and interstitial lung disease, were not routinely recorded in the NLST study. There remains much debate as to which of the COPD phenotypes, airflow limitation or CT-based emphysema (CTE), is more closely associated with lung cancer risk.[184,185]

A fourth limitation of our study is that lung cancers identified during CT screening are not identical to those diagnosed in an unscreened cohort.[93,215-216] Results from the full NLST bear this out with an excess of both adenocarcinomas and bronchioloalveolar cancers (now sub-classified under adenocarcinomas) reflecting an "histology shift".[93] Collectively, these results suggest the lung cancers identified during screening include more indolent lung cancers than

occurs without screening.[214] Results from the NELSON trial using volumetric assessment will help answer this question. If the Danish Lung Cancer Screening Trial (DLST) is representative of the histological differences observed between screened and unscreened lung cancer,[213] then 38 cancers of the 45 excess cancers diagnosed in the CT arm (65% excess compared to no screening) were BAC and Adenocarcinoma subtypes. A similar excess was reported by the DANTE trial suggesting that CT screening identifies a subgroup of cancers that would not otherwise come to clinical attention if no screening were done.[215,216] A comparison of the lung cancer histology, after stratification by COPD, from the European trials will help confirm this observation. Lastly, while we have been able to show the GOLD severity relationship with lung cancer risk is maintained after stratification by screening arm (Table 3.7), smoking status, presence of CT-based emphysema, lung cancer detection (Figure 3.6) and gender (unpublished findings). We are unable to comment on ethnicity due to insufficient numbers in minority groups.

3.4.5 Conclusion

To conclude, the presence of airflow limitation identifies smokers at greatest risk for developing lung cancer, with increasing severity being associated with an increased lung cancer incidence rate.[128,209] Airflow limitation is also associated with marginally more aggressive cancers and significantly less (or minimal) over-diagnosis.[93] Another finding of this study is that the risk of lung cancer is directly related to a reduced FEV_1%predicted in a simple linear relationship, where even only minor reductions in expiratory flow rate (<90% in FEV_1%predicted, GOLD U or restrictive subgroups) confer an increased risk. These results do not support the view that smokers with mild-to-moderate COPD are at more risk than those with GOLD groups 3-4. [187-189] The results of this study confirm previous studies showing that the severity of airflow limitation contributes substantially to differentiating smokers at low or high risk of lung cancer.[86,99,185,199,209] We believe that our findings support the routine use of spirometry in asymptomatic adult smokers,[217] especially those otherwise eligible for CT screening where a “sweet spot” may help define those who will achieve the greatest benefit.

Chapters four, five, six, seven and nine, report studies based on the 10,054, NLST-ACRIN sub group of participants who have spirometry and a blood bio specimen available, the ACRIN-NZ sub-cohort.

In this prospective study, the focus of the following subsequent analysis are the relationships between airflow limitation ($FVC/FEV_1 < 70\%$), lung cancer risk, clinical phenotypes; mortality and competing risks; and screening outcomes. Also presented are the results of a genetic analysis of a CHRNA5 variant, showing independent associations with the risk for lung cancer, COPD, and smoking intensity.

Chapter eight is a New Zealand-specific study.

Chapter 4 Effects of "Airways Disease" and Airflow Limitation on Outcomes in High Risk Smokers Screened for Lung Cancer

"*Life consists not in holding good cards but in playing those you hold well*".

– Josh Billings

4.1 Introduction

Based primarily on the results of the National Lung Screening Trial (NLST), eligibility for lung cancer screening in the United States targets high risk current and former smokers with a minimum age of 55 years and a 30 pack year history.[90,218] In the NLST, lung cancer accounted for 24% of all deaths, whereas the majority of deaths in this high risk population was accounted for by cardiovascular disease (25%), non-pulmonary cancers (22%) and respiratory disease (10%).[90] At least one third of NLST participants have spirometric evidence of airflow limitation.[93] Further, as the severity of airflow limitation worsens; there is a linear increase in the risk of lung cancer.[219] Not only is airflow limitation diagnostic of chronic obstructive pulmonary disease (COPD) and associated with an increased risk of developing lung cancer, but it is also a biomarker of increased mortality from cardiovascular disease and reduced life expectancy.[16,19-20,220] Unfortunately, only 30% of NLST participants with spirometric evidence of airflow limitation (COPD) reported a prior diagnosis of "airways disease".[93,219] On further analysis of this sub-cohort of the NLST, we have shown that those with moderate-to-severe airflow limitation and greatest risk of lung cancer have a much greater all-cause mortality than those with normal lung function, particularly from non-lung cancer related causes (competing risk effect).[80,121,221] These findings suggest outcomes from lung cancer screening are not equally distributed across all smokers and that the presence of airflow limitation may be relevant to outcomes in the screening setting.[80] This is important because there are no agreed recommendations to screen asymptomatic high risk smokers for airflow limitation given the paucity of beneficial evidence-based treatments.[217] In smokers with symptoms reflecting airways disease such as chronic bronchitis but normal lung function (GOLD 0),[222] poorer outcomes have been reported compared to those who were asymptomatic with normal lung function.[223] We ask "Do lung cancer screening outcomes in high risk smokers with pre-existing airways disease and normal lung function

(GOLD 0), or airflow limitation with and without a history of airways disease, differ from healthier asymptomatic smokers?"

This study was undertaken to assess the effect of doctor diagnosed (symptomatic) airways disease, with and without airflow limitation, on CT-based screening outcomes.

Hypothesis: In the context of high risk smokers, the presence of comorbid airways disease or airflow limitation worsens the outcomes from CT-based lung cancer screening.

Aim: To examine the relationship of comorbid airways disease or airflow limitation on the outcomes from CT-based lung cancer screening in the NLST**.**

Objective: To stratify the participants of the NLST at baseline according to the presence or absence of comorbid airways disease and/or airflow limitation and then compare the rates of lung cancer that developed over the follow-up period.

4.2 Methods

4.2.1 Subjects

The recruitment and study design of the full National Lung Screening Trial (NLST) involving 53,452 screening participants, yielding 2,058 histology confirmed lung cancers, has been described elsewhere.[90] In the American College of Radiology Imaging Network (ACRIN) cohort of the NLST, participants from 23 centres consented to undergo both baseline pre-bronchodilator spirometry and blood sampling for biomarker analysis (N=10,054). From this cohort, 395 histology-confirmed lung cancer cases were diagnosed over the study period of 7.5 years although another 14 people were diagnosed with lung cancer at post-mortem. Demographic data including doctor diagnosed pre-morbid diseases were collected through an extensive questionnaire (defined in Table 4.1, legend). Although baseline imaging assessed for the presence or absence of emphysema, as methodologies were different and severity was not quantified, this was not further analysed.

4.2.2 Pulmonary function testing

In the NLST-ACRIN cohort, pre-bronchodilator spirometry was measured at baseline screening (T0) in the majority of participants meeting the following criteria; no chest infection in the preceding 3 weeks and no use of a short-acting bronchodilator inhaler in the

preceding 6 hours or long-acting bronchodilator in the preceding 24 hours. Those not meeting these criteria were rescheduled for spirometry testing at a later visit. Of the 10,054 subjects, 174 (2%) did not perform spirometry to a sufficient standard to be included in this current study (N=9,880). The spirometry was measured by trained staff using a Spiropro spirometer (eResearchTechnology, GmbH, Germany). The severity of airflow limitation was defined according to the Global Initiative on Chronic Obstructive Lung Disease (GOLD) criteria grades 1-4 (www.GOLD.org accessed May 24, 2019).

4.2.3 Clinical phenotyping

Each participant of the NLST-ACRIN sub-study with satisfactory spirometry (98%), was assigned to four mutually exclusive phenotype groups based on their personal medical history, reflecting the presence of doctor diagnosed respiratory morbidity, and pre-bronchodilator pulmonary function tests, specifically forced expiratory volume in one second (FEV_1) and forced vital capacity (FVC) (Figure 4.1).

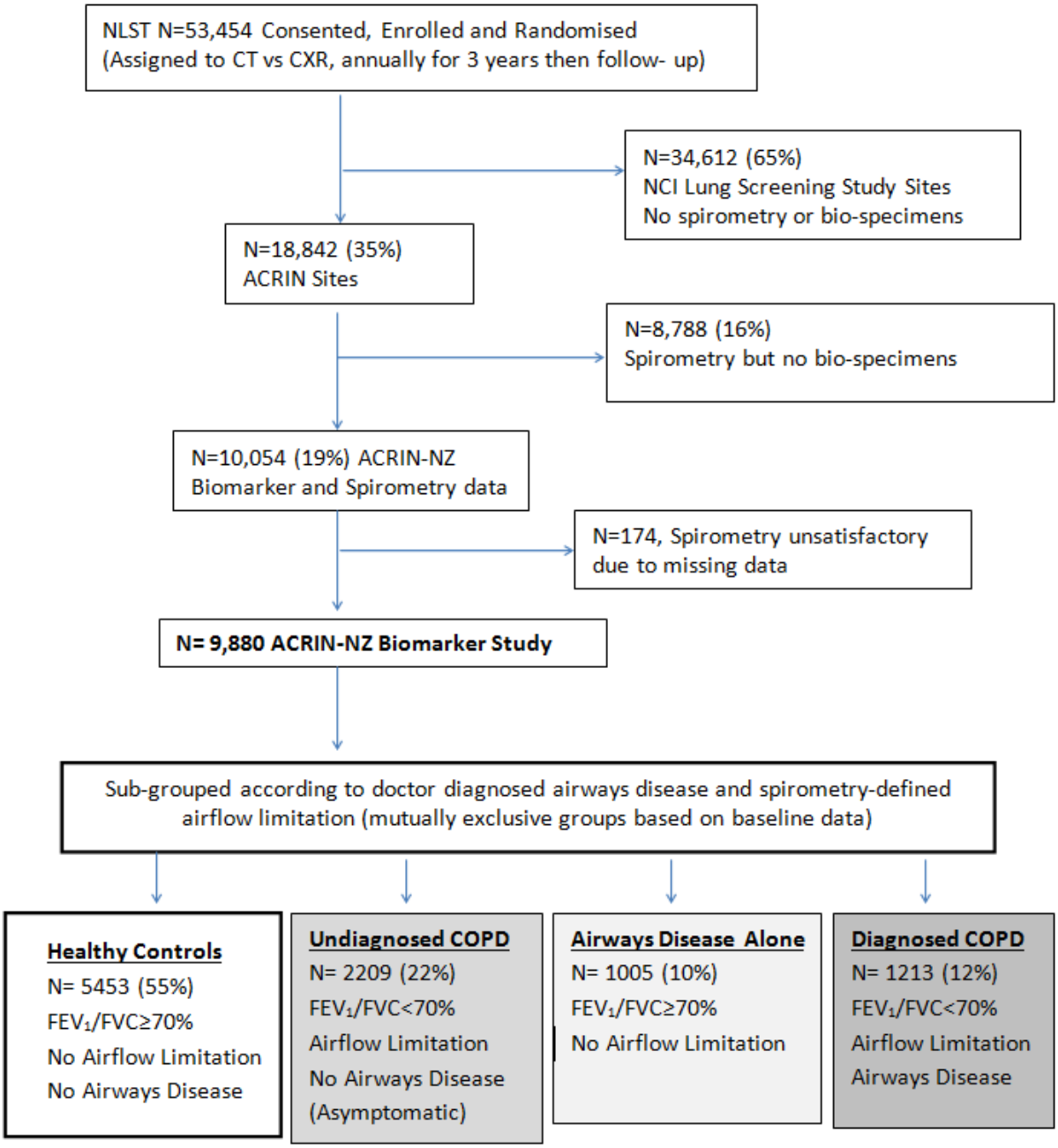
Group 1 (Referent group, (Healthy Controls), included those with no previously diagnosed respiratory morbidity (responded "no" to the question "Has a doctor ever told you that you have any of the following conditions….COPD, emphysema, chronic bronchitis, or adult asthma"), and no airflow limitation $FEV_1/FVC \geq 0.70$;

Group 2 (Airways disease only) included those with previously diagnosed respiratory morbidity (responded "yes" to the question "Has a doctor ever told you that you have any of the following conditions… COPD, emphysema, chronic bronchitis, or adult asthma"), but no airflow limitation (approximating GOLD 0);[222,223]

Group 3 (Undiagnosed COPD) included those with airflow limitation (FEV_1/FVC ratio < 70%) but no known respiratory morbidity (see above, Undiagnosed COPD);

Group 4 (Diagnosed COPD) included those with airflow limitation ($FEV_1/FVC<0.70$) and who had been previously doctor diagnosed with one or a combination of COPD, emphysema, chronic bronchitis, or adult asthma.

Figure 4.1 Subgrouping of NLST-ACRIN subjects according to the presence of pre-bronchodilator lung function and history of respiratory morbidity



4.2.4 Lung cancer and cause-specific mortality

Lung cancer cases included all those diagnosed during the trial (N=395), whether screen or non-screen detected (interval), or prevalent (diagnosed during the first year, i.e., T0) or incident lung cancers (diagnosed during subsequent years T1 to T6).[90] All lung cancer cases were confirmed on histological sampling according to accepted international classification criteria.[90] Lung function results and mortality outcomes were available for 389 of the 395 lung cancer cases (98% of total). The NLST was terminated early when the endpoint of a 20%

reduction in lung cancer specific mortality in the CT arm, relative to the CXR arm, was reached with a mean follow-up of 6.4 years. Cause of death was a primary outcome for the NLST and ascertained through death certification.[90] Cause of death was grouped according to either, lung cancer-related, cardiovascular, respiratory, non-pulmonary cancer and other as per the ICD classification.[178]

4.2.5 Statistical analysis

Differences in baseline characteristics and screening outcomes following stratification by phenotype were compared against the referent smokers' group 1 (Healthy controls), using Dunnett's test for continuous variables and by pairwise Fisher's exact test. For each of the three comparisons the overall significance level for each variable was maintained at 5% by using a false discovery protected p value (Benamini-Hochberg). Differences in death incidence rates were compared between airflow limitation groups, after adjustment for pack years and age, by general linear modelling assuming a Poisson distribution with a log link function. Contrasts were constructed for pairwise comparisons between the referent smokers' group 1, and each of the increasing airflow limited groups. Tests for trend were performed using the Cochran-Armitage test for categorical data and orthogonal contrasts for continuous outcomes. Rate differences were calculated using www.openepi.com (accessed 8/7/2019). Hazard ratios for non-lung cancer death with adjustment for the competing risk of lung cancer death were calculated using the Fine and Gray approach (PHREG procedure SAS) with and without further adjustment for age and sex. The other statistical analyses including ANOVA were performed using SAS (v9.4 SAS Institute Inc, Cary, NC) or STATA statistical software (version 16, College Station, TX: StataCorp LLC).

4.3 Results

4.3.1 Phenotypic sub-groups:

From the cohort of 10,054 NLST participants, 9,880 (98%), had satisfactory spirometry and were classified into 4 phenotypic groups (Figure 4.1) as follows:- Group 1 = 'Healthy controls', asymptomatic smokers (n=5,453 (55%)), Group 2 = doctor diagnosed airways disease but no airflow limitation (airways disease only), (n= 1,005 (10%)), Group 3 = airflow limitation but no doctor diagnosed airways disease described as undiagnosed COPD GOLD 1-

4 grade (n= 2,209 (22%)) and Group 4 = doctor diagnosed, airways disease and airflow limitation consistent with diagnosed COPD GOLD 1-4 grade (n=1,213 (12%)).

4.3.2 Demographic variables

The demographic variables, lung function results and doctor diagnosed, co-morbid diseases are summarised in Table 4.1, with mortality summarised in Table 4.2, and lung cancer and screening outcomes in Table 4.3. Compared to Group 1 (Healthy controls), there were slightly fewer non-Caucasians in Group 4 (diagnosed COPD GOLD 1-4) ($P<0.01$). Compared to Group 1 (Healthy controls), subjects in Groups 3 and 4 (undiagnosed COPD and diagnosed COPD), were slightly older, ($P<0.0001$); a much higher proportion of women in Group 2 ('airways disease' alone) ($P<0.01$); a greater percentage of males in Group 3 (undiagnosed COPD) ($P<0.0001$); more current smokers in Group 3 (undiagnosed COPD) and Group 4 (diagnosed COPD) ($P<0.0001$); greater BMI in Group 2 (Airways disease) and lower BMI in Group 3 (undiagnosed COPD) and Group 4 (diagnosed COPD) ($P<0.0001$). With regards to smoking history, relative to Group 1 (Healthy controls) there was higher pack year exposure in all other groups ($P<0.0001$), likely attributable to greater years smoking (particularly in Groups 3 and 4 with COPD) ($P<0.0001$) and more recent quitting across all groups, ($P<0.01$). While mean cigarettes per day were comparable, there were significant differences between groups ($P<0.01$). There was no difference in family history of lung cancer and minimal difference in education level (Table 4.1).

4.3.3 Pre-morbid disease

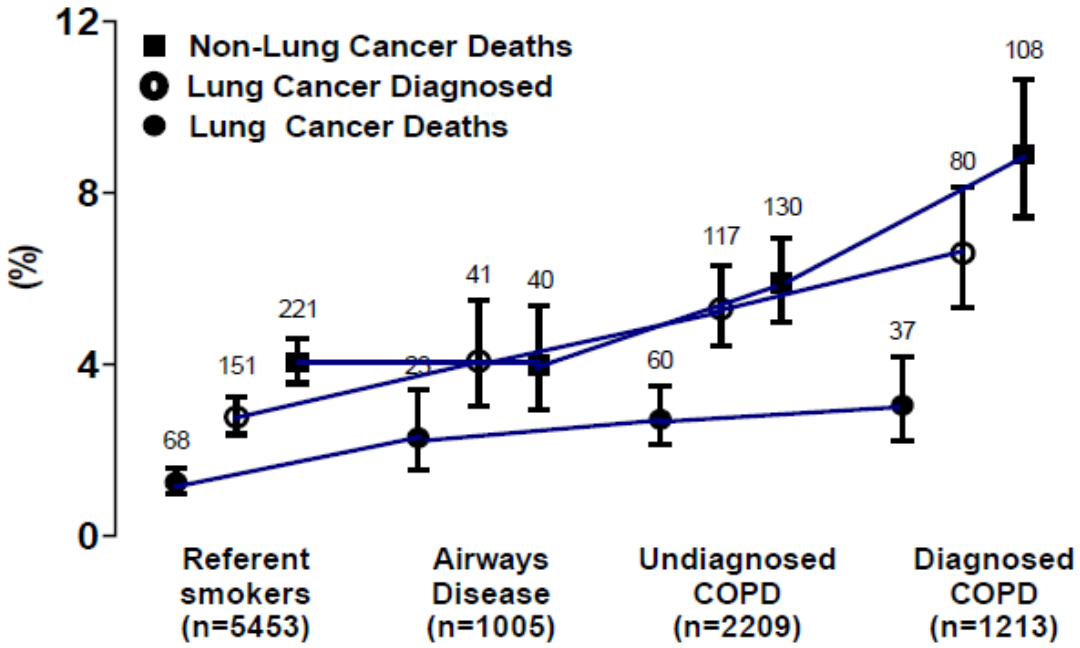
The two variables reflecting pre-morbid diseases were assessed according to reported medical histories (doctor diagnosed), and pre-bronchodilator lung function measured at baseline (Table 4.1). With regards to doctor diagnosed pre-morbid disease compared to those in Group 1 (Healthy controls (Referent smokers)), there was significantly more pneumonia in subjects in Groups 2 (Airways disease) and 4 (diagnosed COPD) ($P<0.0001$); more heart disease and stroke in Groups 2 (Airways disease) and 4 (diagnosed COPD) ($P<0.0001$); and more diabetes and hypertension in Group 2 (Airways disease) $P<0.001$. As expected, the FEV_1/FVC ratio was normal in Group 1 (Referent smokers, 77%), and reduced in Groups 3 and 4 (with airflow limitation, 61% and 57% respectively) ($P<0.0001$). The FEV_1%predicted became increasingly lower across groups 1 (90%), 2 (84%), 3 (71%) and 4

(59%), (P<0.0001). Relative to Group 4 (diagnosed COPD), those in Group 3 (undiagnosed COPD), had more GOLD grade 1 and less GOLD 3-4 COPD (P<0.05). In this current study, doctor diagnosed "airways disease", collectively defined as subjects previously diagnosed with COPD, emphysema, chronic bronchitis, or adult asthma, had a specificity of only 81% and a sensitivity of only 35% in identifying those with airflow limitation.

4.3.4 Screening and health outcomes

The outcomes are summarised in Table 4.2 and Figure 4.2, where deaths are reported as per 100 persons screened (%). We also adjusted for time to death and report mortality per 1000 patient years (Table 4.2). In Figure 4.2, lung cancer incidence in each phenotypic sub-group is compared with lung cancer deaths in comparison with non-lung cancer deaths. While lung cancer incidence increases across the groups 2-4 as lung function worsens (P<0.0001), the increase in lung cancer mortality is somewhat attenuated (P<0.01) (Figure 4. 2).

Figure 4.2 Lung cancer mortality, lung cancer incidence and all-cause mortality (per 100 persons or %) by sub-phenotype (data presented as % + 95% CI)



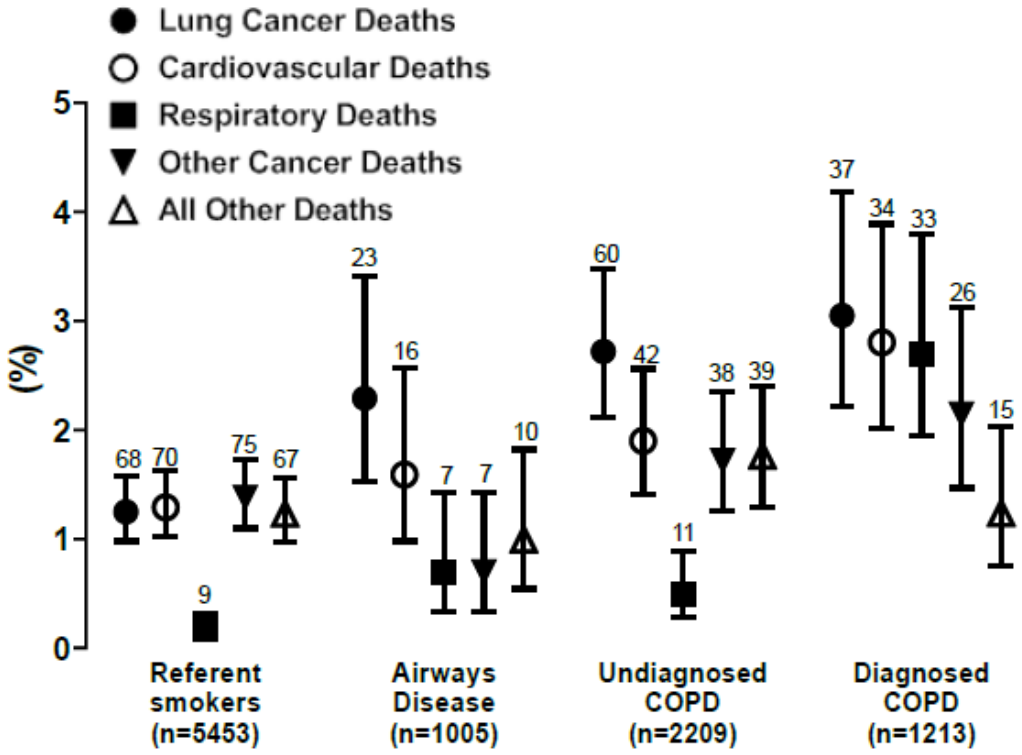
Legend: Error bars are 95% confidence intervals on mortality/100 persons (%) with absolute numbers reported above (see Table 4.2).

In contrast to a rise in non-lung cancer mortality as lung function worsens across Groups 2-4, the lung cancer death rates remain relatively flat (Figure 4.2). This divergence between non-lung cancer mortality and lung cancer mortality, confirmed in a competing cause of death

analysis, (Table 4.2, Figure 4.2), reflects greater deaths from respiratory disease, cardiovascular disease and other cancers (Figure 4.3a, Table 4.2) (P<0.01) observed in Groups 3 and 4. This competing cause of death analysis (Hazard ratio for non-lung cancer death controlling for competing risk of lung cancer death) remained significant in Groups 3 and 4 after adjustment for age and gender. This differential effect on mortality shown in Figure 4.3b indicates smokers in Group 2 (Airways disease) are more likely to die of lung cancer (36%), while those in Group 4 (Diagnosed COPD) are more likely to die of both lung cancer (26%) and cardiovascular disease (23%) (P<0.01). Smokers in Group 4 (Diagnosed COPD) are more likely to die of respiratory disease (23%) despite only marginally worse lung function than those in Group 3 (Undiagnosed COPD). Deaths from other causes in these 2 groups are otherwise similar. Those with diagnosed COPD (Group 4), had the highest rates of cardiovascular and respiratory mortality in contrast to those with undiagnosed COPD (Figure 4.3a), where lung cancer was the leading cause of death.

Figure 4.3 Cause of death by sub-phenotype (%), (see Table 4.2)

(a) Absolute Mortality per 100 persons screened (% ± 95% CI)



(b) Relative Mortality (% from all Deaths by phenotype)

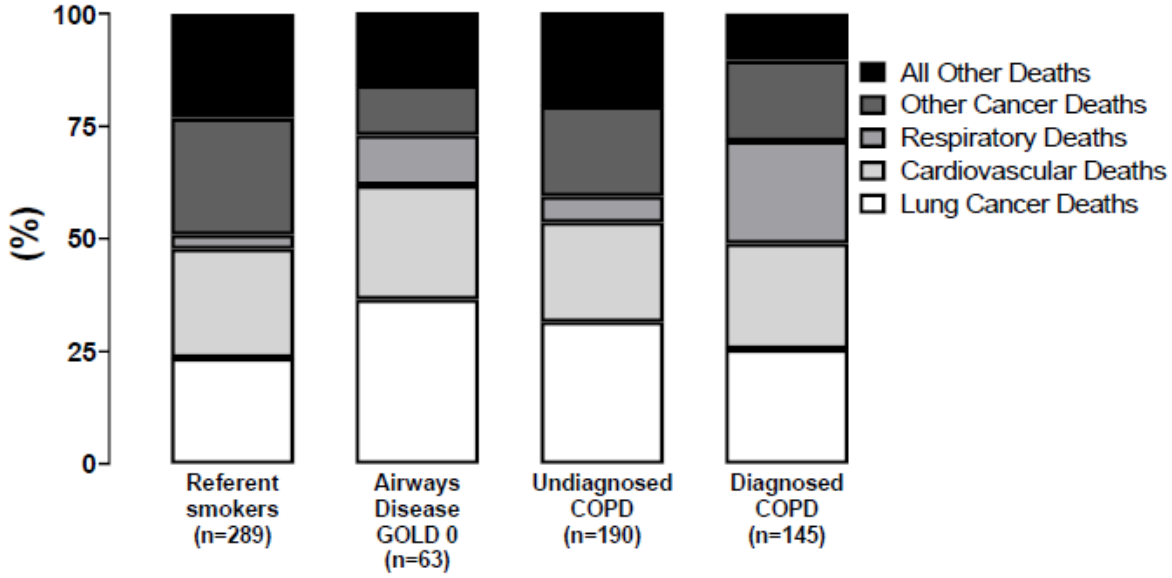
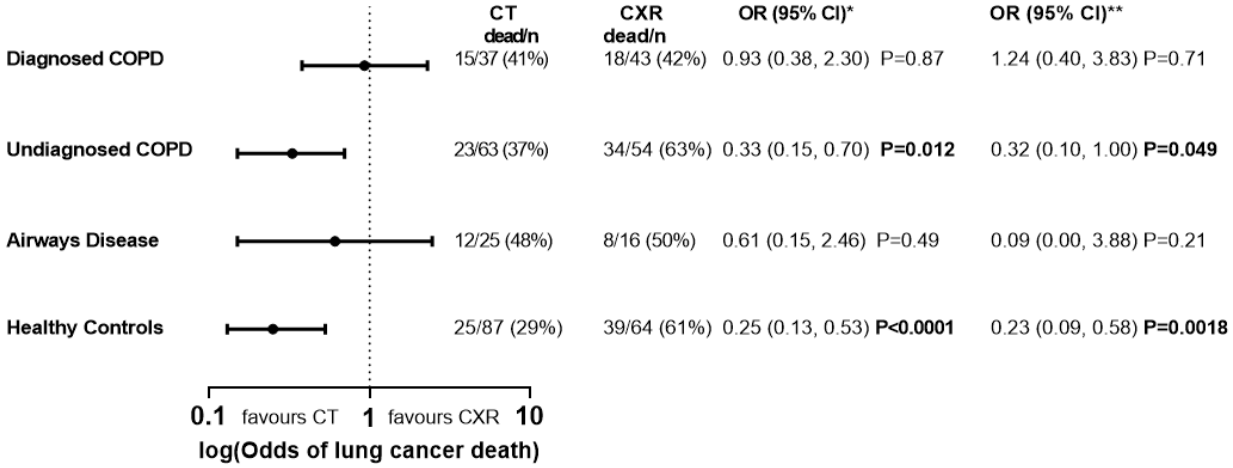


Table 4.3 shows no significant differences between sub-phenotypic groups in respect to histology, stage, and screen detection. Lung cancer surgical rates were lower in all phenotypic groups relative to the Healthy controls, (i.e., Referent smoker, Group 1) (47-49% versus 62% respectively, P<0.01, Table 4.2). However, the surgical rate according to screening arm was significantly different for Groups 1 and 3 only (CT vs CXR, P=0.03 and P=0.02 respectively, Table 4.3). The finding that lung cancer deaths were significantly lower in the CT arm relative to CXR arm of Group 1 (32%, P<0.001) and Group 3 (26%, P=0.002), was as expected. In contrast, in those with "airways disease" alone (Group 2) or diagnosed COPD (Group 4) surgical rate and lung cancer mortality were no different according to screening arm (P>0.05). Groups 1 (Referent) and Group 3 (Undiagnosed COPD) accounted for the vast majority (25/28 or 89%) of lung cancer deaths averted with CT screening observed (Table 4.3). On testing for an interaction between screening arm and phenotype by ANOVA, no significant interaction amongst the demographic variables in Table 4.1, was found. Figure 4.4, provides a summary of the likelihood (Odds) of dying of lung cancer, after adjustment for pack years and age, according to screening arm and clinical phenotype.

Figure 4.4 Odds of lung cancer death‡ in those screening participants randomised to CT versus CXR in whom lung cancer was diagnosed during study follow-up, (see Table 4.3)



* adjusted for pack years and age, ** adjusted for pack years,age,sex, pneumonia, stroke, diabetes, heart disease, any history of cancer and stage of lung cancer

Note: In this Forest plot, Healthy Controls and those with undiagnosed COPD achieved significant reductions in lung cancer deaths following randomisation to CT screening (relative to CXR) even after extensive adjustment (including lung cancer stage).

Table 4.1 Baseline demographic variables for the high risk smokers sub-phenotyped according to the presence of airflow limitation (COPD) and/or respiratory morbidity

Demographic Variables and lung cancer outcomes	No Airflow Limitation		Airflow Limitation	
	Healthy Controls	Airways Disease Only	Un-Diagnosed COPD	Diagnosed COPD*
Phenotype Group	**Group 1**	**Group 2**	**Group 3**	**Group 4**
N=9,880# (% total)	**N=5453(55%)**	**N=1005(10%)**	**N=2209 (22%)**	**N=1213(12%)**
Race and Age				
1 Caucasian	5053 (92.7%)	940 (93.5%)	2062 (93.3%)	1156 (95.3%)*
2 Other	370 (6.8%)	65 (6.5%)	147 (6.7%)	57 (4.7%)
Age mean (SD)	61.1 (4.9)	61.3 (5.0)	62.(5.3)***	63.2 (5.3)***
Gender				
Male	3136 (57.5%)	382 (38%)***	1434 (65%)***	653 (53.8%)*
Smoking History				
-Current Smoker Status	2496 (45.8%)	472 (47%)	1287 (58.3%)***	588 (48.5%)***
Pack Years mean (SD)	52.9 (21.3)	57 (24.1)***	58.8 (24.8)***	62.0 (25.3) ***
Cigarettes/day mean (SD)	27.5 (10.8)	28.9 (11.3)***	28.1 (10.9)*	29.4 (11.3)***
Years Quit mean (SD)	4.0 (5.2)	3.7 (5.0)*	2.8 (4.6)***	3.4 (4.8)***
Smoking duration years mean (SD)	39.2 (7.3)	40.1 (7.5)***	42.1 (7.2)***	42.6 (7.1)***
Other risk variables				
FHx of Lung Cancer -Yes	1285 (23.6%)	249 (25%)	511 (23%)	291 (24%)
Airways Disease (%)*	0	1005 (100%)	0	1213 (100%)
BMI mean (SD)	28.4 (5.0)	29.2 (6.1)***	26.7 (4.7) ***	26.8 (5.1)***
Weight mean kg (SD)	84.1 (17.7)	83.2 (19.1)	80.3 (17.1)***	78.8 (17.9)***
Height mean cm (SD)	171.9 (9.8)	168.4 (10.3)***	173.2 (9.7)***	171.2 (9.8)
Education Level✞ (%)				
-High school or less	1478 (27.1%)	316 (31.4%)***	665 (30.1%)*	418 (34.5%)***
-Post High school training	665 (12.2%)	144 (14.3%)	267 (12.1%)	162 (13.4%)
-Some College	1276 (23.4%)	251 (25%)	477 (21.6%)	265 (21.8%)
- Graduate/Postgraduate	1905 (35.0%)	258 (25.7%)	744 (33.7%)	335 (27.6%)
Lung Function				
FEV1/FVC mean (SD)	77.4 (4.8)	76.8 (4.7)***	61 (8.4) ***	56.5 (10.0)***
FEV1 % predicted mean (SD)	90.3 (15.2)	83.6 (15.9)***	70.7 (18.6)***	58.7 (18.9)***
FVC% predicted mean (SD)	89.1 (14.8)	83.5 (15.6)***	87.8 (20.3)***	78.2 (19.6)***
GOLD Grade				
GOLD 1	-	-	700 (31.7%)	168 (13.8%)
GOLD 2	-	-	1234 (56.0%)	644 (53.1%)
GOLD 3-4	-	-	272 (12.3%)	398 (32.8%)
Doctor diagnosed Comorbidities				
COPD	0	169 (16.8%)	0	468 (38.6%)
Chronic Bronchitis	0	578 (57.5%)	0	514 (42.4%)
Emphysema	0	266 (26.5%)	0	608 (50.1%)
Adult-onset Asthma *	0	328 (32.6%)	0	363 (30%)
Pneumonia	1200 (22.0%)	468 (46.6%)***	550 (24.9%)*	507 (41.8%)***
Heart Disease	660 (12.1%)	168 (16.7%)***	278 (12.6%)	196 (16.2%)**
Hypertension	1931 (35.4%)	407 (40.5%)**	788 (35.7%)	446 (36.8%)
Stroke	141 (2.6%)	42 (4.2%)**	47 (2.1%)	61 (5%)***
Diabetes	539 (9.9%)	138 (13.7%)***	167 (7.6%)*	107 (8.8%)
Any cancer History	189 (3.5%)	55 (5.5%)**	93 (4.2%)**	68 (5.6%)**

#174 (1.7%) Subjects of the 10,054 did not have spirometry and were excluded from the analyses. * Airways Disease =”Has a doctor ever told you that you have any of the following conditions. ”COPD, Emphysema, Chronic Bronchitis, or Adult Asthma (where prior childhood asthma is excluded). Relative to “Healthy Controls” (Group 1) the following achieved statistical differences Dunnett’s continuous variable/FDR categorical test; *P<0.01, **P<0.001 and ***P<0.0001. ✞ Excludes other/unknown

Table 4.2 Mortality outcomes for the high risk smokers sub-phenotyped according to the presence of airflow limitation (COPD) and/or airways disease (respiratory morbidity)

	Healthy Controls	Airways Disease Only	Un-Diagnosed COPD	Diagnosed COPD
Group	Group 1	Group 2	Group 3	Group 4
N=9,880 (98% total)	N=5453 (55%)	N=1005 (10%)	N=2209 (22%)	N=1213 (12%)
Randomised to CT	2757 (50.6%)	526 (52.3%)	1059 (48%)	607 (50%)
Lung Cancer Outcomes				
Lung cancer diagnosis (%N)389	151 (2.8%)	41 (4%)*	117 (5.3%)***	80 (6.6%)***
Lung cancer Surgery-Yes (÷ LC)	93 (1.7%)	20 (49%)	55 (47%)*	38 (48%)*
Mortality†				
Total patient years follow up	33847	6175	13434	7357
Total Deaths (N) - per 100 screened (%)	289 (5.3%)	63 (6.3%)*	190 (8.6%)***	145 (12%)***
Total deaths /1000 pt years (95% CI)	8.5 (8.0, 9.6)	10.2 (7.9, 13.0)	14.1 (12.2, 16.3)**	19.7 (16.7, 23.1)**
Non-lung cancers (LC) Deaths (N) -per 100 screened (%) - % from all non-LC Deaths	221 (4.1%) 44.3%	40 (4.0%) 8.0%	130 (5.9%)** 26.1%	108 (8.9%)*** 21.6%
Non-LC deaths /1000 pt years (95% CI) HR for non- LC death controlling for competing risk of LC death	8.3 (5.7, 7.4) 1	6.5 (4.7, 8.7) 1.0 (0.7, 1.8) P=0.93	9.7 (8.1, 11.5) 1.5 (1.1, 1.8) P=0.0004	14.7 (12.1, 17.7)** 2.2 (1.5, 3.2) P<0.0001
Lung Cancer (LC) Deaths (N) -per 100 screened (%) -% from all Deaths by phenotype	68 (1.2%) 23.5%	23 (2.3%)** 36.5%	60 (2.7%)*** 31.6%	37 (3.1%)*** 25.5%
Lung Cancer deaths /1000 pt years (95% CI)	2.1 (1.6, 2.5)	3.7 (2.4, 5.5)***	4.5 (3.4, 5.7)***	3.1 (2.2, 4.2)*
Odds of LC death	ref	1.5 (1.12, 2.0)*	2.1 (1.7, 2.6)*	3.0 (2.4, 3.8)**
Cardiovascular (CVD) Deaths (N) -per 100 screened (%) - % from all Deaths by phenotype	70 (1.3%) 24.2%	16 (1.6%) 25.4%	42 (2%)** 22.1%	34 (2.8%)*** 23.4%
CVD deaths /1000 pt years (95% CI)	2.1 (1.6, 2.6)	2.6 (1.6, 4.2)	3.1 (2.3, 4.2)**	4.6 (3.3, 6.5)***
Odds of CVD death	ref	1.7 (0.9, 3.0)	2.0 (1.3, 3.1)*	3.0 (1.9, 4.7)**
Respiratory Deaths (N) -per 100 screened (%) - % from all Deaths by phenotype	9 (0.17%) 3.1%	7 (0.7%)* 11.1%	11 (0.5%)* 5.8%	33 (2.7%)*** 22.8%
Respiratory Deaths/1000 pt years (95% CI)	0.27 (0.13, 0.49)	1.13 (0.54, 2.38)**	0.82 (0.45, 1.48)**	4.49 (3.19, 6.31)***
Odds of Respiratory death	ref	7.5 (2.2, 25.5)*	5.3 (1.7, 16.7)**	29.7 (10.5, 84.1)***
Other Cancer Deaths (N) -per 100 screened (%) - % from all Deaths by phenotype	75 (1.4%) 26.0%	7 (0.7%) 11.1%	38 (1.7%) 20.0%	26 (2.1%) 17.9%
Other Cancer Deaths/1000 pt years (95% CI)	2.2 (1.8, 2.8)	1.1 (0.5, 2.4)	2.8 (2.1, 3.9)	3.5 (2.4, 5.2)
Odds of other Cancer death	ref	0.5 (0.2, 1.2)	1.3 (0.9, 2.0)	1.7 (1.0, 2.7)*
Other Deaths (N) -per 100 screened (%) - % from all Deaths by phenotype	67 (1.2%) 23.2%	10 (1.0%) 16.0%	39 (1.8%) 20.0%	15 (1.2%) 10.0%
Other deaths /1000 pt years (95% CI)	2.0 (1.5, 2.5)	1.6 (0.9, 3.0)	2.9 (2.1,.4.0)*	2.0 (1.2, 3.4)
Odds of other death	ref	0.9 (0.5, 1.8)	1.6 (1.1, 2.5)*	1.1 (0.6, 2.1)

Relative to Healthy Controls ("Referent smokers") (Group 1) the following achieved statistical differences (Poisson regression); *P<0.01, **P<0.001 and ***P<0.0001. †False discovery protected P values. pt years=patient years. HR=Hazard ratio, 95% CI= 95% Confidence Interval.

Table 4.3 Lung Cancer outcomes for the high risk smokers sub-phenotyped, according to the presence of airflow limitation (COPD) and/or respiratory morbidity

Demographic Variables	Healthy Controls	Airways Disease Only	Un-diagnosed COPD	Diagnosed COPD
Lung Cancer Diagnosis (LCdx, N=389)				
Lung Cancer Diagnosis N= (% total grp)	151 (2.8%)	41 (4.1%)**	117 (5.3%)***	80 (6.6%)***
Lung Cancer Mortality				
Lung Cancer Death (N=(% total grp incl †)	68 (1.2%)	23 (2.3%)*	60 (2.7%)***	37 (3%)***
Airflow Limitation (N=%LCdx)				
GOLD 1		-	36 (30.8%)	4 (5%)
GOLD 2		-	59 (50.4%)	42 (52.5%)
GOLD 3-4		-	22 (18.8%)	33 (41.3%)
Lung Cancer Histology (N=% total LCdx)				
Small Cell	14 (9.3%)	6 (14.6%)	17 (14.5%)	11 (13.8%)
Squamous Cell	32 (21.2%)	10 (24.4%)	28 (24%)	20 (25%)
Adenocarcinoma	60 (39.7%)	14 (34.1%)	39 (33.3%)	25 (31.3%)
BAC	22 (14.6%)	3 (7.3%)	8 (6.8%)	4 (5%)
Non-small Cell	19 (12.6%)	6 (14.6%)	22 (18.8%)	19 (23.8%)
Large cell/Other	4 (2.6%)	2 (4.9%)	3 (2.7%)	1 (1.3%)
Lung cancer Stage at Dx (N=% total LCdx)				
Stage I-II	78 (51.7%)	19 (46%)	55 (47%)	41 (51%)
Stage III	34 (22.5%)	6 (14.6%)	24 (20.5%)	14 (17.5%)
Stage IV	39 (25.8%)	15 (36.6%)	35 (30%)	21 (26.3%)
Occult carcinoma/ unk	-	1 (2.4%)	3 (2.6%)	4 (5%)
Screen Detection				
Cancer Year (N=% LCdx)				
-T0 (1st yr Screening)	40 (26.5%)	9 (22%)	27 (23.1%)	19 (23.8%)
-T0-T2 (Screening Interval)	95 (62.9%)	19 (46%)	78 (67%)	52 (65%)
-T3-T7 (Follow-up Interval)	56 (37.1%)	22 (54%)*	39 (33%)	28 (35%)
Cancer screen (N=% total LCdx)				
Screen-detected	80 (53.0%)	18 (44%)	64 (55%)	41 (51%)
Missed/Interval	11/4 (10.0%)	0/1 (2%)	13/1 (12%)	8/3 (14%)
Follow –up	56 (37.1%)	22 (54%)	39 (33%)	28 (35%)
Surgery - Yes (N=% total LCdx)	93 (61.6%)	20 (49%)*	55 (47%)*	38 (48%)*
Screening Outcomes (CT vs CXR)	2757/2696	526/479	1059/1150	607/606
Patient years follow up for lung cancer	792.0	206.7	561.4	431.5
LC surgical rate in CT vs CXR arm (%)	60/87 vs 33/64 69.0% vs 51.6%	12/25 vs 8/16 48% vs 50%	36/63 vs 19/54* 57% vs 35%	18/37 vs 20/43 49% vs 47%
Risk Difference in surgical rate (95% CI)	17% (1.8, 33) P=0.03	-2% (-33, 29) P=0.91	22% (4.3, 40) P=0.019	2.1% (-20, 24) P=0.85
LC death rate in CT vs CXR arm (%)	25/87 vs 39/64 28.7% vs 61.0%	12/25 vs 8/16 48% vs 50%	23/63 vs 34/54* 37% vs % 63%	15/37 vs 18/43 41% vs 42%
Risk Difference (95% CI)	-32% (-47,-17) P<0.0001	-2% (-33, 29) P=0.91	-26% (-44, -9) P=0.002	-1.3% (-23, 20) P=0.91
Relative reduction† in lung cancer deaths in CT vs CXR (% reduction)	-51%	+36%	-27%	-17%
Absolute LC deaths averted with CT Relative Risk Difference (95% CI)	-14 ↓36% (23, 52)	NA (+4 LC)	-11 ↓32% (19, 49)	-3 ↓17% (5.8,-39)
LC deaths averted /1000 patient years	17.7 (1.0, 2.9)	NA	19.6 (10.3, 34)	7.0 (1.8, 19)
Odds of LC death‡ in LC cases randomized to CT v CXR	0.25 (0.13, 0.50) P<0.0001	0.61 (0.15, 2.46) P=0.49	0.33 (0.15, 0.70) P=0.0043	0.93 (0.38, 2.30) P=0.15

Relative to Healthy Controls (Group 1) the following achieved statistical differences;* P<0.01, **P<0.001 and ***P<0.0001 ‡ adjusted for smoking pack years and age, NA = Not Applicable as numbers too small and excess lung cancer deaths in the CT arm. †The overall relative reduction in lung cancer deaths comparing CT with CXR was -17% in the full cohort.

4.4 Discussion

The primary aim of this study was to compare mortality and lung cancer screening outcomes, across several respiratory-related phenotypic groups, relative to healthy asymptomatic smokers in the NLST and there are several important findings. First, in contrast to a rise in non-lung cancer mortality as lung function worsens across Groups 1-4, the lung cancer death rates remain relatively flat, reflecting greater deaths from non-lung cancer causes. Second, although mortality for smokers with undiagnosed COPD was nearly as poor as for those with diagnosed COPD, the lung cancer deaths averted with CT screening in the former was two-fold greater (32% vs 17% reduction). Further, relative to healthy asymptomatic smokers, those with self-reported respiratory morbidity without airflow limitation (Airways disease) had a greater risk of developing and dying from lung cancer. Third, screening participants who were 'Healthy controls' (Referent smokers), or who had 'Undiagnosed COPD', gained the greatest benefit from CT screening relative to other phenotypic groups. The former groups were characterised by the absence of doctor diagnosed airways disease, low prevalence of comorbid disease and higher surgical rates in those randomised to CT.

4.4.1 Differences in cause of death

This study found nearly a 2-fold reduction in lung cancer deaths in screened smokers with undiagnosed COPD (Group 4) compared to those with diagnosed COPD (relative reduction 32% vs 17%). Reductions in lung cancer deaths averted per 1000 patient years were 19.6 (Odds of lung cancer death, P=0.0043) and 7.0 (Odds of lung cancer death, P=0.15) in Groups 3 and 4 respectively. This was despite comparable lung cancer risk, histology, stage, detection method and surgical rate. The former group (Group 3) had less airflow limitation, no respiratory morbidity and lower mortality from lung cancer and cardiovascular causes. In contrast, those with diagnosed COPD (Group 4) had greater GOLD 3-4 prevalence and a greater proportion of deaths from respiratory cause relative to lung cancer and cardiovascular deaths. In a preliminary post-hoc analysis of outcomes in the NLST, in contrast to subjects with GOLD 1-2 where screening was effective (40% reduction in lung cancer deaths), those with GOLD 3-4 had minimal benefit from screening.[102] Indeed this study found, within those with COPD (Groups 3 and 4), compared to GOLD 1-2 severity, GOLD 3-4 was associated with a 1.4 fold excess of lung cancer deaths, 1.3 fold excess of

cardiovascular deaths and a 9-fold excess in respiratory deaths (unpublished result). These findings confirm that outcomes from lung cancer screening are not directly proportional to the degree of lung cancer (risk) but are modified by pre-existing co-morbid disease, particularly severe airflow limitation.[37,97,102,113,122,142]

4.4.2 Risk vs benefit of screening

Previously, we and others have shown that the benefits of screening are attenuated in those at greatest risk, where those at the highest risk die prematurely from their comorbid disease.[102,122,142] This is very important as the clinical labels used to define comorbid disease reveals that doctor diagnosed "airways disease" or COPD, collectively defined as subjects previously diagnosed with COPD, emphysema, chronic bronchitis, or adult asthma, in this study had a specificity of only 81% and a sensitivity of only 35% in identifying those with airflow limitation. Put another way, while as much as 65% of those with airflow limitation were undiagnosed, this group had two-fold better outcomes from screening than those with diagnosed COPD. This was associated with a better lung function, lower GOLD 3-4 prevalence and fewer deaths from respiratory disease, cardiovascular disease, and lung cancer. By examining spirometry-defined phenotypes and assessing respiratory morbidity, clinicians are able to better categorize screen-eligible participants into different but clinically meaningful groups. These sub-phenotypes appear to have a bearing on outcomes of screening for lung cancer notably surgical rate (Table 4.3) and lung cancer mortality benefit (Figure 4.4) and may better inform the shared-decision making process. This is not advocating that screen-eligible smokers in Groups 2 or 4 (symptomatic groups) are not screened, rather this result is suggesting patient factors may be relevant in attenuating the outcomes of screening and that these may be important to consider where there are cost or resource constraints, or doubt about the individualised benefits of screening.[102,103,113]

There is growing interest in the significance of smokers with symptoms of airways disease, primarily a chronic productive cough or exertional breathlessness, but not spirometric evidence of airflow limitation.[222-223,228] The classification of GOLD 0 was coined for this group of people but then abandoned in part because GOLD 0 subjects were subsequently shown to be at no increased future risk of developing COPD.[222] However, this group is at an increased risk of dying compared to their asymptomatic counterparts.[223,228] The current study shows

that Group 2 (Airways disease, approximating GOLD 0 status) is female predominant, with high rates of chronic bronchitis and pneumonia, and at a particularly high risk of developing and dying of lung cancer. This is consistent with the findings of the Swedish study using a similar definition.[223] Poor outcomes from screening in this group, where symptom load is high, may be due in part to more stage 4 cancers diagnosed in the follow-up interval.

4.4.3 Strengths and limitations

There are several strengths and weaknesses underlying this study. First the study was limited to heavy smokers aged 55-74 years old with a minimum 30 pack year history who were volunteers rather than randomly selected from the community such that the generalisability of these findings are limited to this group. A second limitation of the study is that for reasons of powering, mortality outcomes have been grouped across organ systems according to organ-specific ICD codes rather than providing specific cause of death outcomes. The latter approach is the subject of a larger follow up study involving over 18,000 subjects from the NLST.[93] A third limitation of this current analysis of screening outcomes (Table 4.3) is small sample sizes, although differences in surgical rates and lung cancer mortality according to screening arm for Groups 1 and 3 are statistically robust. Consistent with any mortality study, the reliability of the cause-specific data relies heavily on death certification. While this may be questioned, in the NLST cause of death was the primary endpoint and carefully adjudicated on this basis.[90]

This study has several strengths. The first of these is that this study confirms that the routine use of office spirometry helps identify those most at risk for dying of lung cancer or dying of non-lung cancer causes.[16,19,20,217,219] In addition, the findings indicate office spirometry helps distinguish who benefits most from screening among asymptomatic smokers eligible for lung cancer screening. This is important because, unlike spirometry, which is not recommended in asymptomatic smokers,[217,230,231] lung cancer screening of asymptomatic smokers is a clinically proven[90,177] and widely recommended intervention, supported by the United States Preventive Services Task Force.[105] Moreover, by identifying these spirometric-based phenotypes, primary care providers may be prompted to assess and treat several modifiable risk factors (smoking, blood pressure, metabolic syndrome, cholesterol and diabetes) which could improve a current or former smokers' overall health.

4.4.4 Conclusion

In conclusion, the presence of airflow limitation affects cause of death and outcomes during lung cancer screening. However, the findings of this study suggest that those at greatest risk of lung cancer are not necessarily those who gain the most benefit from CT-based screening and that there exist sub-groups of high risk smokers who appear to gain more from screening than others. While much of the work since the publication of NLST has focused on individualising risk, little consideration has been given to the effects of comorbid disease and reduced life expectancy on outcomes from screening. In a recently described simulation model of lung cancer screening, it was reported that life expectancy was shortest in those at greatest lung cancer risk and that this might be due to increased mortality from comorbid disease attenuating the benefits of screening.[133] The presence of severe COPD (GOLD 3-4) has been shown to reduce life expectancy by as much as 10 years in ever smokers, particularly those with reduced exercise tolerance.[19] This study suggests that routine use of simple spirometry in the screening setting might allow for a better assessment of likely outcomes from lung cancer screening,[80,102] a better assessment of life expectancy,[16,19,20] and a more informed decision-making consultation about the benefits of screening.

Chapter 5 The Effect of Airflow Limitation and Lung Cancer Risk on Lung Cancer Screening Outcomes: All Cause and Lung Cancer Specific Mortality

5.1 Introduction

Following the findings of the National Lung Screening Trial (NLST), annual computed tomography (CT) screening for lung cancer is recommended in the United States.[90,218] This recommendation was based on a 20% relative risk reduction in lung cancer specific mortality in the CT arm compared to those randomised to annual chest-x-ray (CXR) screening.[90] However, reduction in all-cause mortality was only 7% suggesting the benefit of reducing lung cancer mortality does not translate into a comparable reduction in overall death.[90] Recognising that current screening eligibility criteria include current and or former smokers with very wide ranging risks of lung cancer,[96] attention has turned to risk-based approaches to screening selection.[232] This approach assumes that increasing risk is linearly related to increasing benefit from screening although some have questioned this assumption.[112,122,142]

The primary aim of screening for lung cancer is to affect a stage shift whereby more lung cancers can be identified at a sufficiently early stage to facilitate long-term survival by surgical intervention. In this context "overtreatment" (or inappropriate screening) in the setting of lung cancer screening can occur when the harms of work-up or treatment outweigh the benefits.[233,234] This is important in CT screening for lung cancer because lung cancer screening involves older heavy smokers for whom overall background morbidity and mortality is high due to co-existing smoking related diseases, primarily cardiovascular disease and respiratory disease.[20,102] We have previously shown that airflow limitation, that characterises smokers with chronic obstructive pulmonary disease (COPD), is a marker for premature death from all causes[20] and affects about 35% of NLST participants.[93] We have also shown in both unscreened and screened lung cancer cases, that worsening airflow limitation increases the risk of developing lung cancer in a linear relationship.[219] In a secondary data analysis of the American College of Radiology Imaging Network (ACRIN) sub-group of the NLST, we observed that although airflow limitation at baseline was associated with a two-fold greater risk of lung cancer, it was also associated with an almost halving of the lung-cancer mortality reduction with CT, relative to NLST participants with normal lung

function.[80] We propose that the increased risk of developing lung cancer associated with worsening airflow limitation might also be associated with a greater risk of dying from lung cancer (more aggressive disease and less surgery) and a greater risk of dying from a disease other than lung cancer. This raises two questions, "As airflow limitation worsens, is there a differential effect on lung cancer-specific mortality relative to other causes of death?" and "As the risk of lung cancer increases, could a differential effect from airflow limitation on mortality and screening outcomes attenuate the benefits of CT screening?

In an analysis of the ACRIN sub-cohort of the NLST participants, where baseline spirometry was available and the risk of lung cancer could be estimated, we undertook this study to examine the relationship between pre-morbid disease, cause-specific deaths, and outcomes from screening.

Hypothesis: In the context of high risk smokers, worsening airflow limitation is associated with greater premorbid disease and a differential effect on lung cancer versus non-lung cancer death.

Aim: To examine the relationship of worsening airflow limitation on doctor diagnosed, comorbid disease and mortality from lung cancer and non-lung cancer causes.

Objective: To stratify the participants of the National Lung Screening Trial at baseline according to the presence of airflow limitation and then compare the rates of premorbid disease and deaths from lung cancer or other causes.

5.2 Methods

5.2.1 Subjects

This is a secondary data analysis of the National Lung Screening Trial (NLST). The recruitment and study design of this trial involving 53,452 screening participants, yielding 2058 histology confirmed lung cancers, has been described elsewhere.[90,173] In the ACRIN sub-cohort of the NLST, participants from 23 centres agreed to undergo baseline pre-bronchodilator spirometry and blood sampling for biomarker analysis (N=10,054). From this cohort, 395 histology-confirmed lung cancer cases were diagnosed over the study period of 7.5 years although another 14 people were diagnosed with lung cancer at post-mortem.[120] Demographic data, including doctor diagnosed (self-reported) history of pre-morbid disease,

were collected through an extensive questionnaire, and baseline imaging was assessed for the presence or absence of emphysema but not quantified due to the different radiology screening methods. My preliminary results of this analysis have been published in abstract form.[120]

5.2.2 Pulmonary function testing

In the NLST-ACRIN cohort, pre-bronchodilator spirometry was measured at baseline screening (T0) in the majority of participants meeting previously published criteria.[93] The spirometry was measured by trained staff using a Spiropro spirometer (eResearchTechnology, GmbH, Germany) (see Chapter 2, section 2.3.2). The severity of airflow limitation was defined according to the Global Initiative on Chronic Obstructive Lung Disease (GOLD) criteria grades 1-4 (www.GOLD.org accessed March 2, 2018). Those with no airflow limitation were further sub-grouped into those with normal lungs ("Healthy Smokers") and GOLD Undefined (GOLD U), where the latter is defined as $FEV_1/FVC \geq 0.70$ and FEV_1%predicted <80% (http://www.copdgene.org/study-design accessed November 2, 2018).

5.2.3 Lung cancer and cause-specific mortality

Lung cancer cases included all those diagnosed during the trial (N=395), whether screen or non-screen detected (interval), or prevalent (diagnosed at T0 baseline screen), or during the first year) or incident lung cancers (diagnosed during subsequent years T1 to T6).[90] All lung cancer cases were confirmed on histological sampling according to accepted international classification criteria. Lung function results and mortality outcomes were available for 388 of the 395 lung cancer cases (98% of total). The NLST was terminated early when the endpoint of a 20% reduction in lung cancer specific mortality in the CT arm, relative to the CXR arm, was reached with a mean follow-up of 6.4 years.[90] Cause of death was a primary outcome for the NLST, and ascertained through review of clinical records and death certification.[90] Cause of death was grouped according to the International Classification of Diseases Ninth Revision (ICD)[178] codes into lung cancer, other cancers, respiratory, cardiovascular, external and other (remaining) deaths, where external deaths included those from accidents, self-harm and surgical procedure.

5.2.4 Lung Cancer risk, airflow limitation and CT outcomes

In this current study the ACRIN-NLST participants were stratified according to the Prostate, Lung, Colorectal, Ovarian Cancer trial $(PLCO)_{M2012}$ Model for lung cancer risk,[86,123] to assign tertile of risk and to assess whether there was any effect on lung cancer-specific mortality reduction conferred by CT screening relative to CXR screening. Cox-proportional Hazard Ratios were compared for risk variables found to associate with cause-specific mortality.

5.2.5 Statistical analysis

The demographic and clinical characteristics of groups defined by either severity of airflow limitation, FEV_1% predicted and GOLD grade, or tertile of lung cancer risk score ($PLCO_{M2012}$), were compared using chi-square test or parametric or non-parametric analysis of variance, as appropriate. Differences in lung cancer incidence rates were compared according to cause of death stratified by GOLD status or tertile of lung cancer risk score using general linear modelling assuming a Poisson distribution with a log link function. Contrasts were constructed for pairwise comparisons between the healthy smoker group and each of the increasing airflow limited groups or between all combinations of the lung cancer risk score. False discovery protected p values were calculated to maintain an overall 5% significance level. Clinically important potential confounders (i.e., age and pack years) were included in each model as appropriate. All other statistical analyses were performed using SAS (v9.4 SAS Institute Inc, Cary, NC) or STATA statistical software (version 16, College Station, TX: Stata Corp LLC).

5.3 Results

Differences in demographic variables, according to a healthy, restrictive, or obstructive airways disease phenotype, based on baseline pre-bronchodilator spirometry (N=10,054), are outlined in Table 5.1. Those with obstructive airways disease were sub-grouped according to GOLD grade criteria. Of the 10,054, 186 were removed from further analysis due to incomplete information, leaving 9,868 subjects.

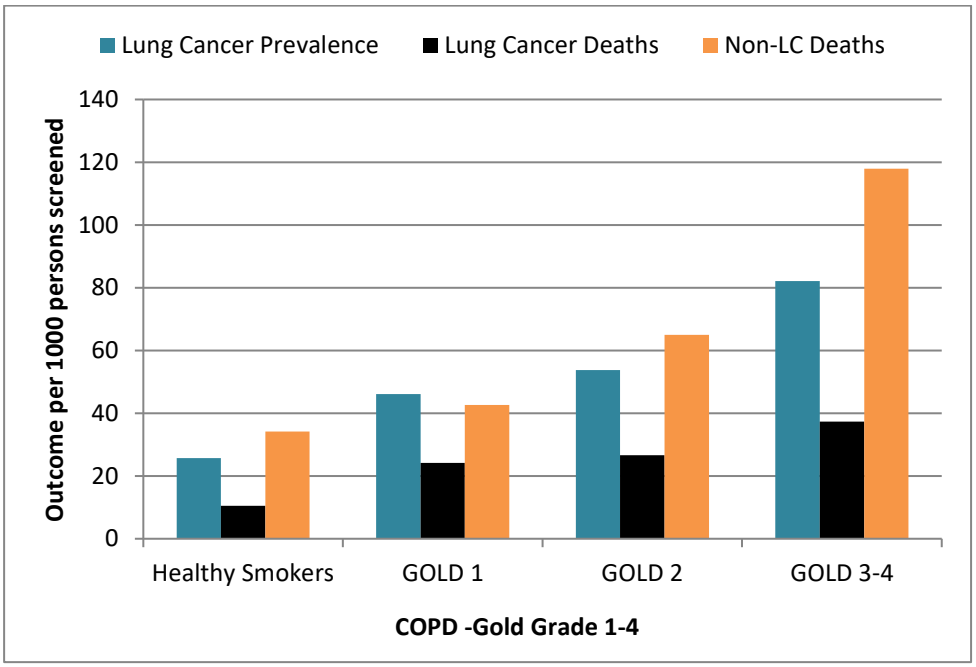
Compared to “healthy smokers”, those with a “restrictive pattern” (N=1,594, 16%) were significantly more likely to be female, have a higher BMI, and significantly lower college education and FEV_1% predicted. Additionally, while they had greater smoking exposure and

cardio-respiratory morbidity and mortality, these trends did not reach statistical significance. This heterogeneous group were excluded from further analysis.

Compared to “healthy smokers”, and across increasing GOLD grade severity, there were significant differences among most demographic variables known to underlie the risk of lung cancer, (except a family history of lung cancer). Specifically, those with spirometry evidence of COPD were older, more likely to be male, had greater smoking exposure, more often self-reported COPD, had lower BMI and lower education status. Across GOLD grade groups, lung function was significantly lower while self-reported pre-morbid respiratory diseases were significantly greater (Table 5.1). As the severity of GOLD grade COPD increased, there was a significant increase in lung cancer prevalence, lung cancer mortality, and deaths, from other causes, mostly due to cardiovascular and respiratory causes (Table 5.1, Table 5.2, and Figures 5.1and 5.2).

Figure 5.1 Lung Cancer deaths and Non-Lung Cancer deaths, relative to lung cancer prevalence, per 1000 persons after stratification in the NLST-ACRIN sub-cohort

(a) Outcomes per 1000 persons stratified according to COPD (GOLD grade 1-4)



(b) Outcomes per 1000 persons stratified according to $PLCO_{M2012}$ Lung Cancer risk model by tertile

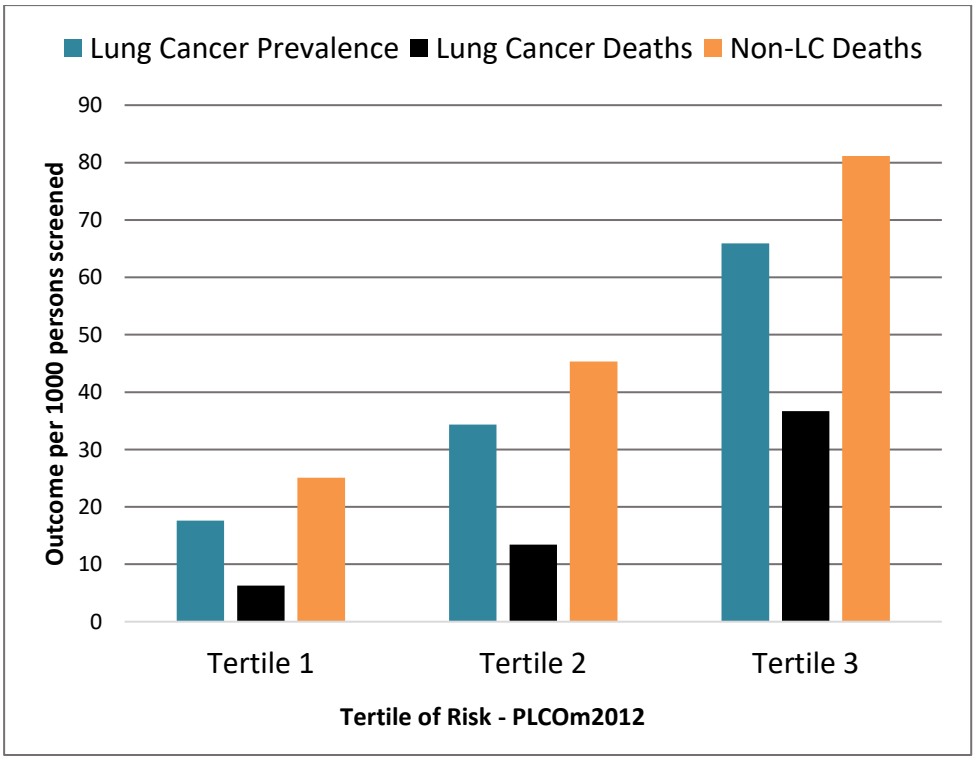
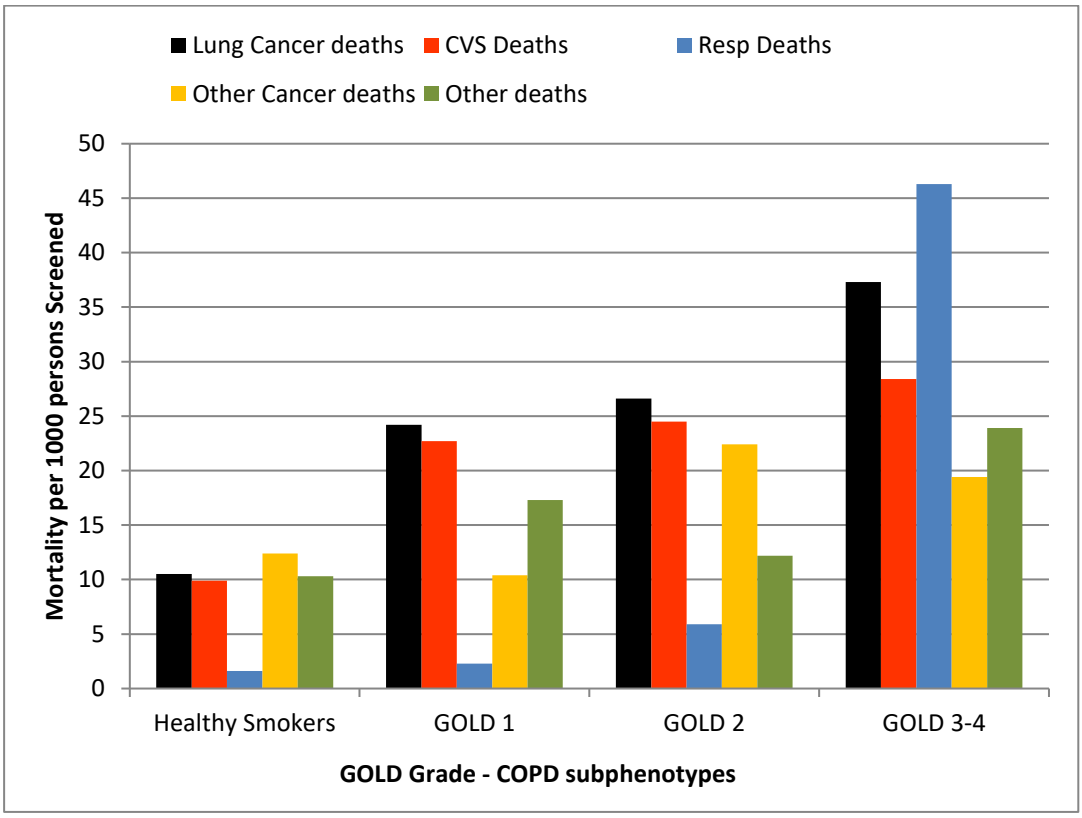
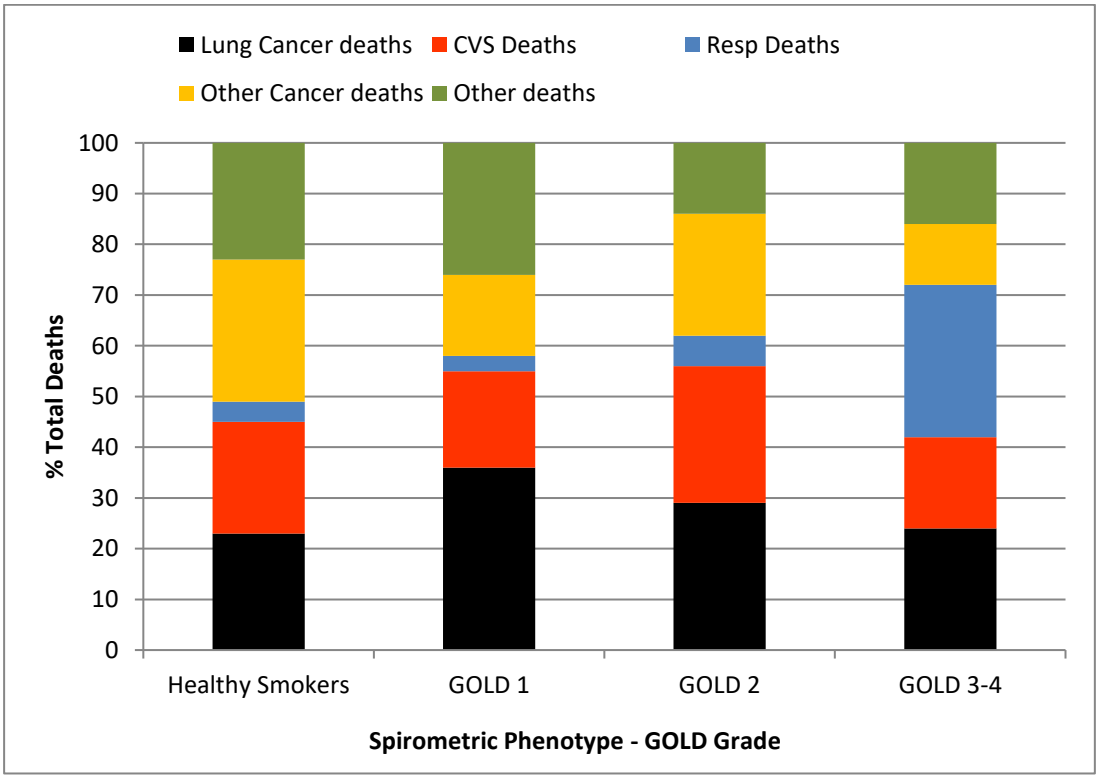


Figure 5.2 Disease-specific deaths (per 1000 persons) stratified by GOLD grade (COPD sub-phenotypes) in the NLST-ACRIN sub-cohort

(a) Mortality per1000 persons for Lung Cancer deaths relative to other disease specific deaths



(b) Percentage of all deaths within each "Spirometric Phenotype" attributed to specific causes



When comparing the lung cancer deaths with non-lung cancer deaths (per 1000 persons screened), stratified according to COPD GOLD grade 1-4 or the $PLCO_{M2012}$ lung cancer risk by tertile, there was divergence in mortality favouring non-lung cancer deaths as the COPD severity increased (Figures 5.1 and 5.2a). A similar divergence effect was found with increasing lung cancer risk according to the risk tertile (Figure 5.1b). When the prevalence of cause specific deaths (per 1000 persons screened) were compared according to GOLD grade (Figure 5.2a), deaths from lung cancer, cardiovascular disease and other cancers increased according to GOLD grade, whereas for GOLD 3-4, respiratory deaths became the most prevalent cause of death (Table 5.1, Figure 5.2b). Notably the death rate for lung cancer, as a proportion of all deaths, reduced as GOLD grade increased (Table 5.1, Table 5.2, Table 5.3, and Figure 5.2b) and for those in the GOLD 3-4 group, respiratory deaths (30%) appeared to replace deaths from lung cancer (24%), cardiovascular disease (18%) and other cancers (13%), (i.e., "substitution effect"), (Figure 5.2b). A similar increasing trend in deaths from lung cancer, cardiovascular disease and from other cancers was seen across $PLCO_{M2012}$ risk tertiles (Table 5.4, Figure 5.3, and Figure 5.4).

Figure 5.3 Disease-specific deaths (per 1000 persons) stratified by $PLCO_{M2012}$ model by tertile in the NLST-ACRIN sub-cohort

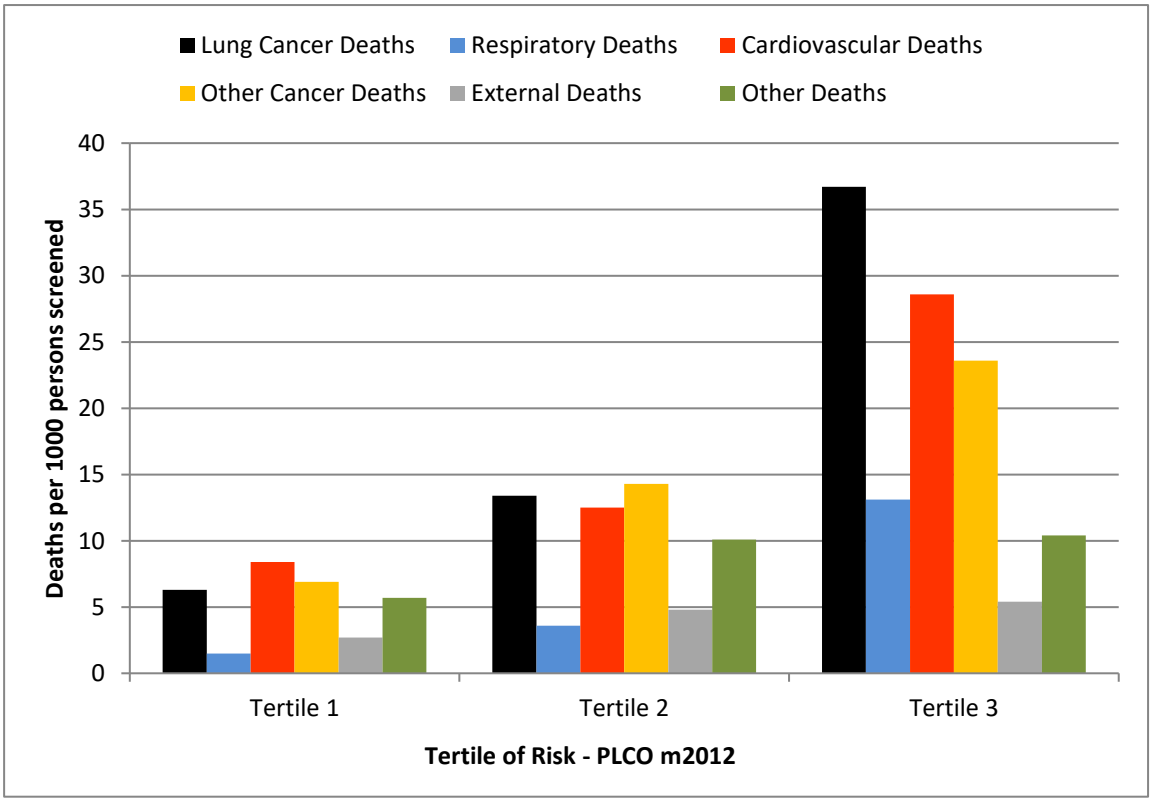
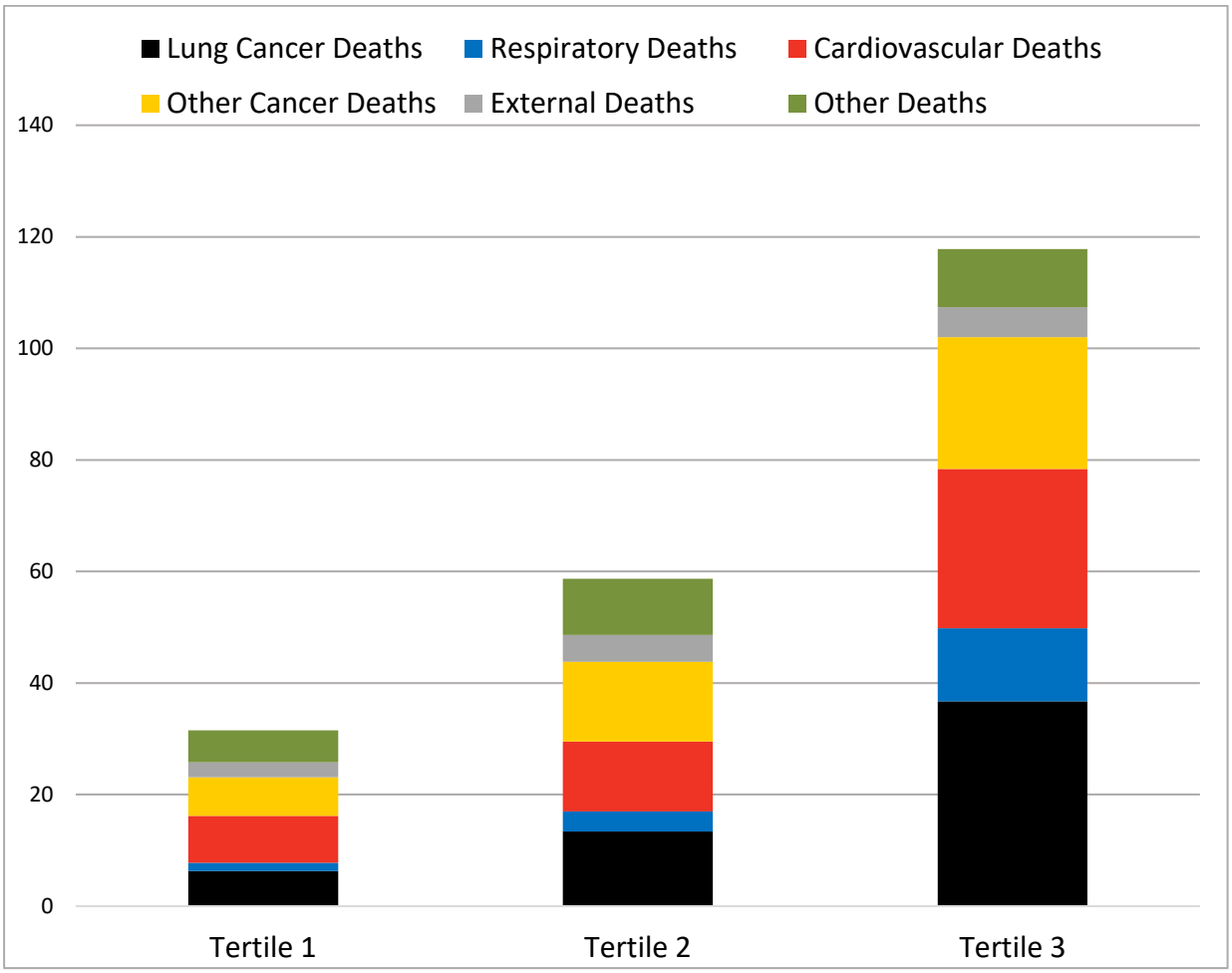


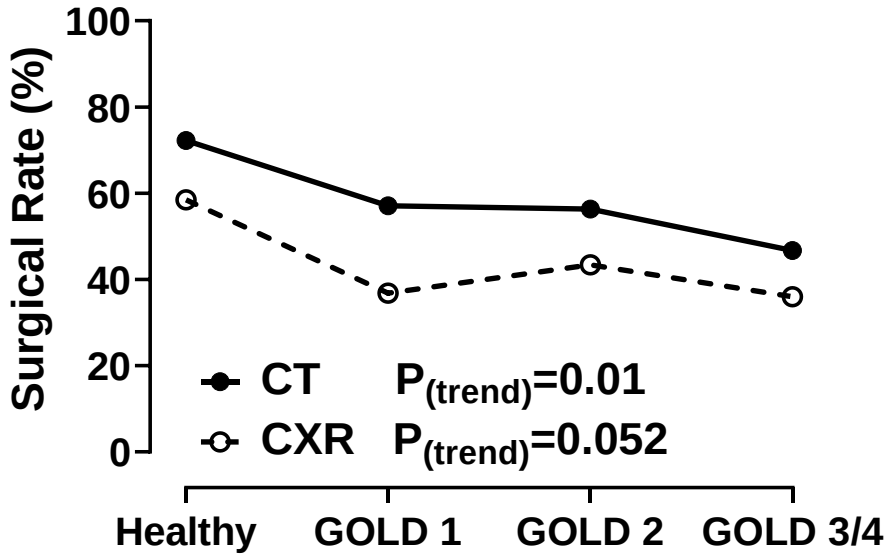
Figure 5.4 Cumulated deaths per 1000 persons screened according to risk of lung cancer ($PLCOM_{2012}$ Model)



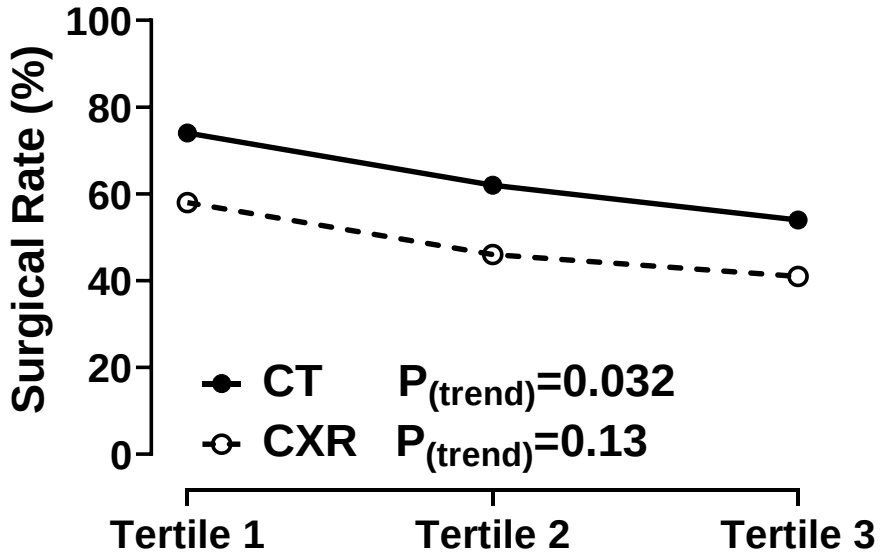
In an adjusted and unadjusted analysis of death rates from specific causes according to risk of lung cancer, estimated using the $PLCO_{M2012}$ Lung Cancer risk calculator,[86,123] there is a trend with increasing death rates for lung cancer, other cancers, and cardiovascular disease across each risk tertile (Table 5.4, Figure 5.3). The prevalence of spirometry-based COPD was 21%, 33% and 48% in tertiles 1-3 respectively (Table 5.4). As the lung cancer deaths/1000 persons screened increases with each tertile of risk, comparable increases in absolute terms are seen for deaths from cardiovascular disease and other cancers (Figure 5.3). In tertile 3, there is an increase in lung cancer, cardiovascular and respiratory deaths relative to other causes of death (i.e., substitution).[102] With increasing risk of lung cancer, there is an increase in the prevalence of COPD ($P<0.0001$) (Table 5.4), more aggressive lung cancer subtypes ($P=0.05$) and less surgery ($P<0.05$) (Table 5.5 and Figure 5.5).

Figure 5.5 Surgical rate as a percentage (%) of all lung cancers according to (a) GOLD grade and (b) lung cancer risk ($PLCO_{M2012}$ model) by screening arm

(a) COPD severity - GOLD grade



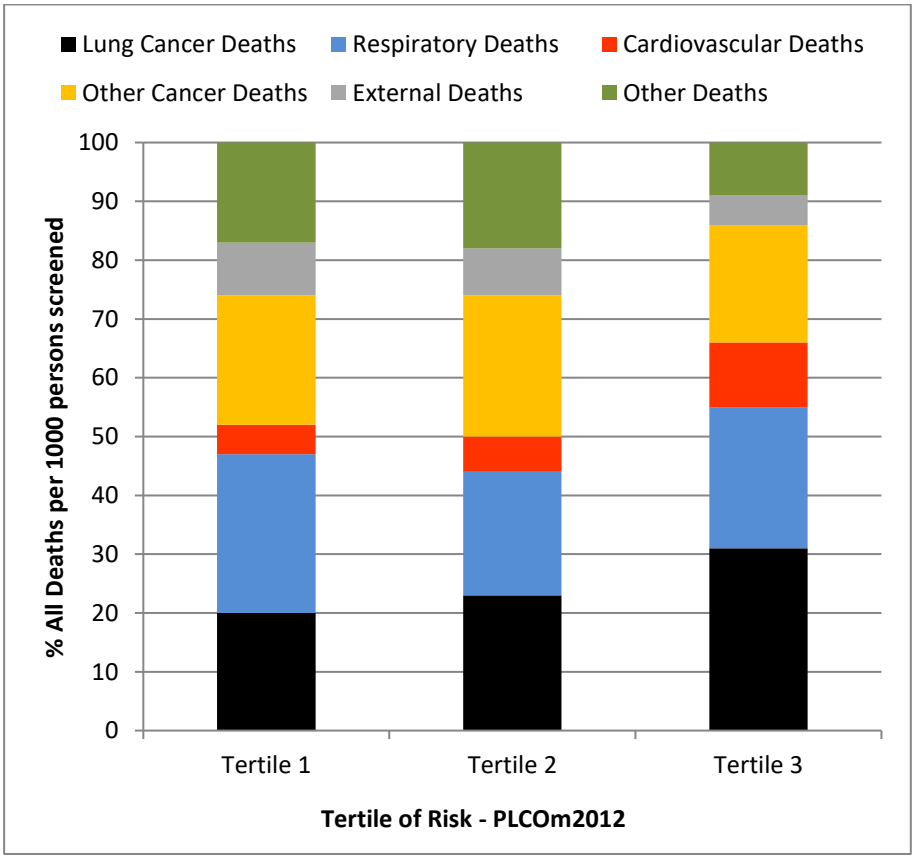
(b) Lung cancer risk - $PLCO_{M2012}$ model



In a Cox-proportional hazards multivariate survival analysis, COPD (GOLD 1-4) was associated with death from lung cancer (HR=2.02), cardiovascular disease (HR=1.77) and respiratory disease (HR=5.1) (Table 5.6). The presence of COPD conferred a similar risk of cardiovascular death as diabetes does (Table 5.6). These observations indicate that after adjusting for other causes of death, COPD confers an increased risk of death from lung cancer and non-lung

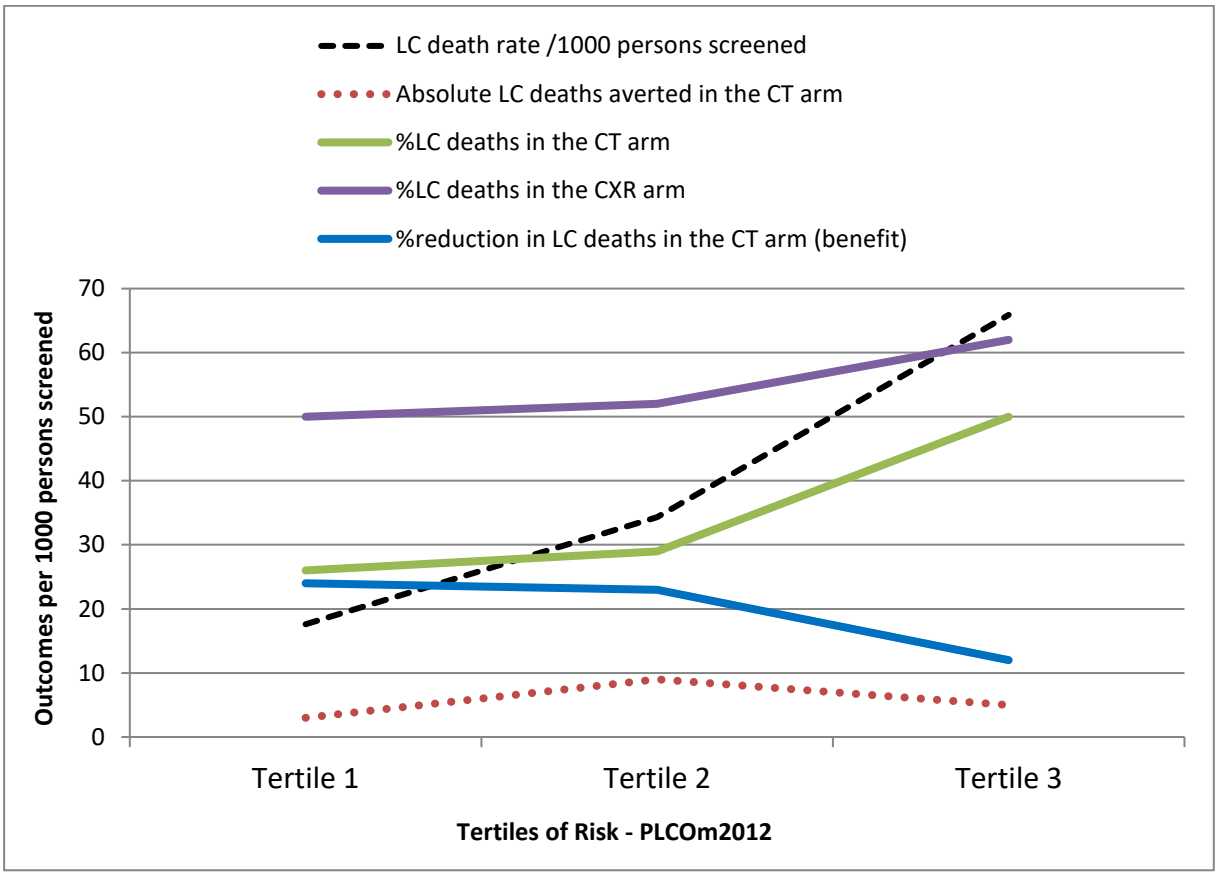
cancer causes attributed largely to cardiovascular causes. When comparing the outcomes of CT versus CXR screening in these NLST-ACRIN participants, stratified by tertile of lung cancer risk according to the $PLCO_{M2012}$ model,[80,219] those in the highest risk tertile had significantly greater pre-morbid disease (Tables 5.4 and 5.5), greater non-lung cancer related mortality (Figures 5., 5.3 5.4,5.6) and lowest lung cancer specific mortality reduction (12%) from CT screening (Table 5.5 and Figure 5.7). Most importantly, the difference in lung cancer specific mortality conferred by CT screening was attenuated (i.e., converged with CXR mortality) in the highest risk tertile (Figure 5.7, Table 5.5). While the relative reduction in lung cancer specific mortality was similar in tertiles 1 and 2 (24% and 23% respectively), the absolute number of lung cancer deaths averted per 1000 persons screened in tertile 2 (N=9), was 3-fold greater than in tertile 1 (n=3,) and nearly 2-fold greater than tertile 3 (N=5) (P=0.044, Table 5.5). In Figure 5.5, the surgical rate for lung cancer is compared according to COPD grade and lung cancer risk ($PLCO_{M2012}$ tertiles). Worsening GOLD grade, and increasing lung cancer risk to a lesser degree, are associated with lower surgical rate (P<0.05).

Figure 5.6 Disease-specific deaths (per 1000 persons) stratified by $PLCOM_{2012}$ model by tertile in the NLST-ACRIN sub-cohort



Percentage of all deaths within each risk tertile attributed to specific causes.

Figure 5.7 Lung cancer mortality and outcomes from screening stratified according to risk of lung cancer ($PLCO_{M2012}$ Model) by tertile



Legend: The percentage lung cancer deaths according to screening arm (CXR=Purple line and CT= Green line) converges as the risk of lung cancer increases (Black dotted line) according to tertile of risk. The net difference between lung cancer deaths in the CT arm and CXR arm (Blue line) decreases as the risk increases. The superiority of CT screening in reducing lung cancer deaths is attenuated in the highest risk tertile in both relative and absolute terms.

Table 5.1 Demographic and Mortality data for NLST-ACRIN participants stratified according to baseline airflow limitation and its severity

	Healthy Smokers	Restrictive (GOLD –U)	COPD GOLD Grade 1	COPD GOLD Grade 2	COPD GOLD Grade 3/4
N=9868†	4858	1594	868	1878	670
Patient years	30288.77	9693.97	5370.41	11440.21	3942.16
Demographics					
Age (mean)	61.00	61.79	62.52‡	62.77‡	63.23‡
Gender (% Male)	55.4%	51.8%	66.6%‡	59.7%‡	57.0%‡
Smoking status (Current)	43.8%	52.4%	53.8%‡	55.6%‡	53.7%‡
Pack Years	52	58	57‡	60‡	64‡
Cigarettes/day (mean)	27	29	28	28‡	30‡
Years Quit (mean years)	4.23	3.22	3.40‡	2.87‡	2.96‡
Smoking duration (mean years)	38.75	40.97	41.23‡	42.47‡	43.19‡
Family history of Lung Cancer	23.9%	23.4%	24.3%	23.0%	23.6%
Self-reported history of COPD	11.4%	21.2%	16.6%‡	30.2%‡	56.3%‡
BMI (mean)	28.18	29.53	25.95‡	27.06‡	26.82‡
Education Level					
-High School (HS) or less	1275 (26%)	519 (33%)	239 (28%)	594 (32%)	249 (37%)
-Post HS training, Some College	1725 (35%)	608 (38%)	288 (33%)	642 (34%)	218 (32%)
-College graduate	863 (18%)	229 (14%)	170 (20%)	305 (16%)	95 (14%)
-Postgraduate/Professional	867 (18%)	201 (13%)	153 (17%)	288 (15%)	67 (10%)
-Other/unknown	128 (3%)	37 (2%)	18 (2%)	49 (2%)	21 (3%)
Lung Function					
FEV_1 % predicted (mean)	95 .51	70.30	90.46‡	45.26‡	38.60‡
FEV_1/FVC (mean)	77.74	75.95	65.02‡	60.92‡	47.72‡
Pre morbid (self-reported) (N %)					
COPD	556 (11%)	338 (21%)	144 (17%)‡	567 (30%)‡	377 (56%)‡
Asthma	193 (4%)	123 (8%)	33 (4%)	205 (11%)‡	124 (19%)‡
Pneumonia	1167 (24%)	500 (31%)	230 (27%)	563 (30%)‡	264 (39%)‡
Heart Disease	541 (11%)	286 (18%)	95 (11%)	270 (14%)‡	109 (16%)‡
Hypertension	1636 (34%)	701 (44%)	254 (29%)	726 (39%)‡	254 (38%)‡
Stroke	104 (2%)	79 (5%)	14 (2%)	64 (3%)‡	30 (5%)‡
Diabetes	438 (9%)	239 (15%)	38 (4%)	170 (9%)	66 (10%)
Any cancer History	181 (3.7%)	63 (4.0%)	35 (4.0%)	93 (5.0%)	33 (4.9%)
Outcome and Mortality (/100 persons)					
Lung cancer Diagnosis	125 (2.6%)	67 (4.2%)	40 (4.6%)‡	101 (5.4%)‡	55 (8.2%)‡
Total deaths	217 (4.5%)	136 (8.5%)	58 (6.7%)‡	172 (9.2%)‡	104 (15.5%)‡
- Lung Cancer (LC) Deaths	51 (1.0%)	40 (2.5%)	21 (2.4%)‡	50 (2.7%)‡	25(3.7%)
- Cardiovascular (CVD) Deaths	48 (1.0%)	38 (2.4%)	11 (1.3%)	46 (2.4%)‡	19 (2.8%)
- Respiratory Deaths	8 (0.2%)	8 (0.5%)	2 (0.2%)	11 (0.6%)‡	31 (4.6%)‡
- Other cancer Deaths	60 (1.2%)	22 (1.4%)	9 (1.0%)	42 (2.2%)	13 (1.9%)
- Other Deaths	50 (1.0%)	28 (1.8%)	15 (1.5%)	23 (1.2%)	16 (2.4%)

†Unknown COPD status N=174; ‡ P<0.0001 compared to healthy smokers. #P values are adjusted for row-wise false discovery rate and are compared to healthy smokers, statistical significance remained unchanged for all pairwise comparisons after adjustment for age and pack-years.

Table 5.2 Event or death rate per 1000 patient years for NLST-ACRIN participants stratified according to baseline airflow limitation and its severity

	Healthy smokers	Restrictive (GOLD –U)	COPD GOLD Grade 1	COPD GOLD Grade 2	COPD GOLD Grade 3/4
N=9868†	4858	1594	868	1878	670
Event or Death rate/1000 patient (pt) years (95% Confidence Limits, CI)					
Patient years	30288.77	9693.97	5370.41	11440.21	3942.16
Lung Cancer Diagnosis/1000 pt years#	4.1	6.9	7.4	8.8	14.0
(95% CI)	(3.5, 4.9)	(5.4, 8.7)	(5.5, 10.2)	(7.3, 10.7)	(10.7, 18.2)
P value		0.007	0.012	<0.0001	<0.0001
Total Deaths/1000 pt years#	7.2	14.0	10.8	15.0	26.4
(95% CI)	(6.3, 8.2)	(11.8, 16.5)	(8.3, 14.0)	(12.9, 17.5)	(21.8, 32.0)
P value		<0.0001	0.0055	<0.0001	<0.00014
- Lung Cancer (LC) Deaths /1000 pt years#	1.7	4.1	3.9	4.4	6.3
(95% CI)	(1.3, 2.2)	(3.0, 5.6)	(2.5, 6.0)	(3.3, 5.8)	(4.3, 9.4)
P value		<0.0001	0.0012	<0.0001	<0.0001
-Cardiovascular (CVD) Deaths /1000 pt years#	1.6	3.9	2.0	4.0	4.8
(95% CI)	(1.2, 2.1)	(2.9, 5.4)	(1.1, 3.7)	(3.0, 5.4)	(3.1, 7.6)
P value		<0.0001	0.44	<0.0001	<0.0001
-Respiratory Deaths /1000 pt years#	0.3	0.8	0.4	1.0	7.9
(95% CI)	(0.1, 0.5)	(0.4, 1.6)	(0.1, 1.5)	(0.5, 1.7)	(5.5, 11.2)
P value		0.02	0.66	0.005	<0.0001
-Other Cancer Deaths /1000 pt years#	2.0	2.3	1.7	3.7	3.3
(95% CI)	(1.5, 2.5)	(1.5, 3.4)	(0.9, 3.2)	(2.7, 5.0)	(1.9, 5.7)
P value		0.59	0.64	0.002	0.10
-Other Deaths /1000 pt years#	1.7	2.9	2.8	2.0	4.1
(95% CI)	(1.2, 2.2)	(2.0, 4.1)	(1.7, 4.6)	(1.3, 3.0)	(2.5, 6.6)
P value		0.018	0.07	0.43	0.002

†Unknown COPD status N=174; ‡ P<0.0001 compared to healthy smokers. #P values are adjusted for row-wise false discovery rate and are compared to healthy smokers, statistical significance remained unchanged for all pairwise comparisons after adjustment for age and pack-years.

Table 5.3 All-cause, lung cancer specific mortality data and screening outcomes for the ACRIN-NLST participants stratified according to GOLD sub-phenotypes

	Healthy Smokers	Restrictive (GOLD U)	COPD GOLD Grade 1	COPD GOLD Grade 2	COPD GOLD Grade 3-4
Airflow Limitation status	No Airflow Limitation†		Airflow Limitation‡		
N=9896 (%Total)	4858 (49%)	1594 (16%)	868 (9%)	1878 (19%)	670 (7%)
Patient years	30288.77	9693.97	5370.41	11440.21	3942.16
Death rate/1000 patient (pt) years (95% Confidence Limits, CI)					
Total deaths /1000 pt years# (95% CI) P value	7.2 (6.3, 8.2)	14.0 (11.8, 16.5) <0.0001	10.8 (8.3, 14.0) 0.0055	15.0 (12.9, 17.5) <0.0001	26.4 (21.8, 32.0) <0.00014
Lung Cancer (LC) Deaths /1000 pt years# (95% CI) P value	1.7 (1.3, 2.2)	4.1 (3.0, 5.6) <0.0001	3.9 (2.5, 6.0) 0.0012	4.4 (3.3, 5.8) <0.0001	6.3 (4.3, 9.4) <0.0001
Non-LC Deaths /1000 pt years# (95% CI) P value	5.48 (4.7, 6.4)	10.0 (8.1, 12.1) P<0.0001	6.9 (5.0, 9.5) P=0.21	10.7 (8.9, 12.7) P<0.0001	20.0 (16.1, 25.0) P<0.0001
Screening Outcomes					
Lung cancer	125	67	40	101	55
Patient years	675.6	323.1	184.0	519.5	282.9
Lung cancer surgical rate in CT vs CXR arm (%) Risk difference (95% CI) P value	52/72 vs 31/53 72% vs 59% 13.7 (-3.1, 30.6) 0.12	20/40 vs 10/27 50% vs 37% 13.0 (-11.0, 36.9) 0.31	12/21 vs 7/19 57% vs 37% 22.4 (-12.0, 56.9) 0.22	27/48 vs 23/53 56% vs 43% 12.9 (-6.5, 31.2) 0.21	14/30 vs 9/25 47% vs 36% 10.7 (-15.3, 36.6) 0.44
Lung cancer death rate in CT vs CXR arm (%) Risk Difference (95% CI) P value	16/72 vs 30/53 22% vs 57% -35.4% (-50.8, -17.9) <0.0001	21/40 vs 17/27 53%% vs 63% -10.5% (-34.4, 13.4) 0.41	9/21 vs 11/19 43% vs 58% -15.0% (-45.7, 15.6) 0.37	16/48 vs 31/53 33% vs 58% -25.2% (-44.0, -6.4) 0.013	12/30 vs 10/25 40% vs 40% 0% (-26, 26) 0.99
Lung cancer deaths averted in the CT arm (%) (95% CI) Lung cancer deaths averted /1000 pt years (95% CI) P value	14 11% (6.5, 17.7) 20.7 (12.7, 35.0)	+4 60 % (2.3, 14.4) 12.4 (4.6, 33.0) 0.36	2 5% (1.4, 16.5) 10.9 (2.7, 43.9) 0.39	15 14.9% (9.2, 23.6) 28.9 (17.4, 47.9) 0.37	+2 3.6% (1.0, 12.3) 7.1 (1.8, 28.3) 0.15

†Unknown COPD status N=174; ‡ COPD Unknown severity N=6.

#P values are adjusted for row-wise false discovery rate and are compared to healthy smokers, statistical significance remained unchanged for all pairwise comparisons after adjustment for age and pack-years.

Table 5.4 Demographic variables and outcomes for the whole cohort according to risk of lung cancer ($PLCO_{M2012}$ Model) by tertiles

Demographic Variables	Lung Cancer Risk ($PLCO_{M2012}$ Model)			
Total	**Tertile 1 N=3351**	**Tertile 2 N=3351**	**Tertile 3 N=3352**	**P value**
Demographics				
- Age (mean, yrs)	59.0 yrs	61.0 yrs	65.0 yrs	**<0.0001**
- Gender (% males)	57%	56%	57%	0.58
- Cigs/day (mean)	27/day	28/day	29/day	**<0.0001[a]**
- Years smoked (mean)	35 yrs	40 yrs	46 yrs	**<0.0001**
- FHx Lung cancer	9.3%	22.4%	39.2%	**<0.0001**
- Self-reported COPD	7.4%	18.3%	34.8%	**<0.0001**
- BMI (mean)	29.4	27.8	26.4	**<0.0001**
Lung function				
- FEV% predicted (mean)	87%	82%	75%	**<0.0001**
- FEV/FVC (mean)	75%	71%	68%	**<0.0001**
- % COPD (GOLD 1-4)	21%	33%	48%	**<0.0001**
- Restrictive pattern	15%	17%	16%	0.15
- Healthy lungs	62%	49%	34%	**<0.0001**
Baseline comorbidities (%)				
- Hypertension	34%	35%	38%	**0.0004[b]**
- Stroke	1.8%	3.0%	4.0%	**<0.0001**
- Heart Disease	11%	13%	16%	**<0.0001**
- Diabetes	10%	10%	10%	0.99
- Chronic Bronchitis	4.7%	11%	18%	**<0.0001**
- COPD	2.2%	6.0%	13%	**<0.0001**
- Emphysema	2.3%	6.6%	18%	**<0.0001**
- Adult Asthma	6.0%	7.0%	8.2%	**0.0012[b]**
- Pneumonia	24%	25%	32%	**<0.0001**
- Any cancer history	1.3%	3.3%	7.8%	**<0.0001**
Mortality after 6.5 yrs (% or /100 persons)				
Total deaths†	105 (15%)	197 (28%)	395 (57%)	**<0.0001**
Cause-Specific deaths‡				
- Lung cancer deaths	21 (0.63%)	45 (1.34%)	123 (3.67%)	**<0.0001**
- CVS deaths	28 (0.84%)	42 (1.25%)	96 (2.86%)	**<0.0001[b]**
- Respiratory deaths	5 (0.15%)	12 (0.36%)	44 (1.31%)	**<0.0001[b]**
- Other cancers deaths	23 (0.69%)	48 (1.43%)	79 (2.36%)	**<0.0001**
- External deaths[#]	9 (0.29%)	16 (0.48%)	18 (0.54%)	0.19
- Other deaths	19 (0.57%)	34 (1.01%)	35 (1.04%)	0.051
Mortality rate/1,000 pt years (95% confidence Limits)†	21042.42	20769.5	20111.57	
Total death rate	5.0 (4.1, 6.0)	9.5 (8.2, 10.9)	19.6 (17.8, 21.7)	
P value		<0.0001	<0.0001	
Cause-Specific death rates				
- Lung cancer deaths	1.0 (0.7, 1.5)	2.2 (1.6, 2.9)	6.1 (5.1, 7.3)	
- P value		0.0034	<0.0001	
- CVS deaths	1.3 (0.9, 1.9)	2.0 (1.5, 2.7)	4.8 (3.9, 5.8)	
- P value		0.09	<0.0001	
- Respiratory deaths	0.24 (0.1, 0.6)	0.6 (0.3, 1.0)	2.2 (1.6, 2.9)	
- P value		0.095	<0.0001	
- Other cancers deaths	1.1 (0.7, 1.6)	2.3 (1.7, 3.1)	3.9 (3.2, 4.9)	
P value		<0.0001	0.0038	
- External deaths	0.4 (0.2, 0.8)	0.8 (0.5, 1.3)	0.9 (0.6, 1.4)	
P value		0.16	0.07	
- Other deaths	0.9 (0.6, 1.4)	1.6 (1.2, 2.3)	1.7 (1.2, 2.4)	
- P value		0.04	0.02	

† % of all deaths, deaths/100 persons, ‡ % of total in each tertile, # Self harm and post-operative death.

[a] tertile 1 significantly different from 3, 2 significantly different from 3, [b] tertile 1 significantly different from 3.

If no superscript all pairwise comparisons are significant on *post hoc* test.

†P values are adjusted for row-wise false discovery rate and are compared to tertile 1, with no adjustment for age and pack years as accounted for in the $PLCO_{M2012}$ risk model.

Table 5.5 Demographic variables and outcomes for the lung cancer cases according to risk of lung cancer ($PLCO_{M2012}$ Model) by risk tertiles

	Lung Cancer Risk ($PLCO_{M2012}$ Model)			
Lung Cancer (LC) Cases	**Tertile 1** N=59	**Tertile 2** N=115	**Tertile 3** N=221	**P**
Demographics				
- Age (mean, yrs)	59 yrs	62 yrs	66 yrs	**<0.0001***
- Gender (% males)	70%	50%	52%	**0.042**
- Cigs/day (mean)	26/day	28/day	29/day	0.10
- Years smoked (mean)	36 yrs	41 yrs	48 yrs	**<0.0001***
- FHx Lung cancer	10.2%	14%	38%	**<0.0001**
- Self-reported COPD	11.9%	27%	37%	**<0.0006**
- BMI (mean)	28.5	27.3	25.8	**<0.0001****
Lung function				
- FEV% predicted (mean)	83%	79%	70%	**<0.0001****
- FEV_1/FVC	73%	69%	64%	**<0.0001****
- % COPD (GOLD 1-4)	25%	37%	63%	**<0.0001**
Lung Cancer- Histology				
- Small cell	6 (10%)	18 (16%)	25 (11%)	**0.050**
- Squamous cell	15 (25%)	24 (21%)	53 (24%)	
- Adenocarcinoma	23 (39%)	40 (35%)	77 (35%)	
- BAC	10 (17%)	14 (12%)	13 (6%)	
- Non-Small cell	3 (5%)	16 (14%)	48 (22%)	
- Large cell/Other	2 (3%)	3 (2%)	5 (2%)	
Lung Cancer –stage	N=59	N=115	N=221	0.77
- 1	31 (52%)	53 (46%)	93 (42%)	
- 2	3 (5%)	4 (3.5%)	11 (5%)	
- 3	13 (22%)	20 (17%)	47 (21%)	
- 4	11 (19%)	35 (30%)	66 (30%)	
- Unknown	1	3	4	
Patient (pt) years	310.5	604.9	1115.4	
Lung Cancer – detection#	N=59	N=115	N=221	0.12
- Screen n(%)	39 (66%)	56 (49%)	111 (50%)	
- Interval	7 (12%)	13 (11%)	21 (10%)	
- Post-screening (PS)	13 (22%)	46 (40%)	89 (40%)	
LC rate/1000 pt years (95% CI)	41.9 (24.3, 72.1)	76 (57.0, 101.5)	79.8 (64.8, 98.2)	
P value		0.057	0.030	
Screening Outcomes – Lung Cancer (Lg Ca)				
Lg Ca surgical rate in CT vs CXR arm (%)†	26/35 vs 14/24 (74% vs 58%)	39/63 vs 24/52 (62% vs 46%)	64/118 vs 42/103 (54% vs 41%)	0.95****
Risk Difference (95% CI)	16.0 (55.1, 78.4) 0.22	15.8 (45.7, 63.6)	13.5 (0.4, 26.7)	
P value		0.10	0.047	
Lg Ca death rate in CT vs CXR arm (%)ϕ	9/35 vs 12/24 (26% vs 50%)	18/63 vs 27/52 (29% vs 52%)	59/118 vs 64/103 (50% vs 62%)	0.71****
Risk difference (95% CI)	-24.3 (-49.0, 0.41)	-23.4 (-40.9, -5.8)	-12.1 (-25.1, 0.87)	
P value	0.07	0.01	0.07	
Lg Ca deaths averted in the CT arm (range)‡ N (range) %	3 (1-8)	9 (5-16)	5 (2-11)	
Deaths averted /1000 pt years	5.1%	7.8%	2.3%	
(95%CI)#	9.6 (3.1, 30.0)	14.9 (7.7, 28.6)	4.5 (1.9, 10.8)	
P value		0.52	0.29	

#Method of lung cancer detection during the NLST-ACRIN study.

‡ P=0.044.*All significantly different from each other (Tukey post hoc test); **T3 different T1/T2; **** interaction tertile;*and group*lung cancer death.

† Overall there is a significant difference between surgical rates in CT and CXR (P=0.0095) and in survival rates across the tertiles (P=0.034) but not significantly different in each tertile (P=0.95).

ϕ Overall the difference in LC death rates is significant between CT and CXR arms (P<0.0001) but not across tertile, P=0.062.

False discovery rate - adjusted comparisons for deaths averted /1000 pt years, tertile 1 v 2 (P=0.52), tertile 1 v 3 (P=0.29), tertile 2 v 3 (P=0.032). 95% CI= 95% Confidence limits.

Table 5.6 Cox-Proportional Hazard Ratios in a multivariate analysis for cause-specific death

Risk variable	Hazard Ratios of death with Standard Error (SE) and P value			
	Lung Cancer	Cardiovascular	Other Cancer	Respiratory
Healthy Lungs	ref	ref	ref	ref
Age†	1.05	1.06	1.08	1.10
SE	(0.01)	(0.02)	(0.02)	(0.03)
P value	0.0003	0.0004	<0.0001	0.0003
Gender (Male)	0.95	1.30	1.80	1.10
SE	(0.15)	(0.18)	(0.19)	(0.30)
P value	0.84	0.12	0.003	0.67
Pack Years‡	1.01	1.01	1.01	1.01
SE	(0.01)	(0.003)	(0.003)	(0.004)
P value	<0.0001	0.005	0.32	0.15
BMI	0.93	0.98	1.00	0.89
Standard Error	(0.02)	(0.02)	(0.02)	(0.05)
P value	<0.0001	0.30	0.93	0.01
COPD (GOLD 1-4)	2.02	1.77	1.23	5.14
SE	(0.18)	(0.18)	(0.19)	(0.37)
P value	<0.0001	0.002	0.22	<0.0001
Restrictive (GOLD U)	2.30	1.87	0.98	2.90
SE	(0.21)	(0.22)	(0.26)	(0.50)
P value	0.0001	0.005	0.95	0.03
PHx-Stroke	0.81	1.94	0.88	0.81
SE	(0.42)	(0.29)	(0.45)	(0.73)
P value	0.70	0.02	0.84	0.83
PHx-Heart D	1.08	2.20	1.02	1.42
SE	(0.21)	(0.18)	(0.23)	(0.33)
P value	0.64	<0.0001	0.85	0.26
PHx-Hypertension	1.17	1.29	1.32	0.86
SE	(0.16)	(0.18)	(0.18)	(0.29)
P value	0.30	0.13	0.12	0.60
PHx-Diabetes	1.01	1.86	2.00	1.68
SE	(0.28)	(0.20)	(0.23)	(0.41)
P value	0.84	0.002	0.003	0.18

5.4 Discussion

This is the first study I am aware of that examines the relationship between airflow limitation and lung cancer risk score in a lung cancer screening cohort and has four important findings. First, worsening airflow limitation (across GOLD grades 1-4), is associated with an increased risk of developing lung cancer, an increased prevalence of comorbid disease and an increased risk of dying from a non-lung cancer related cause, particularly respiratory and cardiovascular diseases. This means that while worsening airflow limitation identifies smokers with the greatest risk of lung cancer,[34] these highest risk smokers are also at greatest risk of dying of something other than lung cancer.[120] Second, increasing risk of lung cancer ($PLCO_{M2012}$ risk score) is associated with more airflow limitation, more aggressive histological subtypes and lower surgical rates.[93] This means a greater risk of developing lung cancer may also be associated with a greater risk of dying of lung cancer, mediated in part by the presence of airflow limitation.[80] Third, when NLST-ACRIN screening participants are stratified according to their overall risk of lung cancer, the reduction in lung cancer death in the CT arm relative to the CXR arm is attenuated in those at greatest risk,[123] most likely due to the effects of underlying airflow limitation described above. Fourth, because the risk of lung cancer is closely correlated with the presence of airflow limitation and non-lung cancer related deaths, the benefits of screening are attenuated with increasing risk of lung cancer (according to the $PLCO_{M2012}$). These findings have implications for risk-based screening as increasing risk of lung cancer does not correlate with increased benefits of screening (the relationship is curvilinear not linear, see Figure 5.7).[113] This analysis supports the view that there exists a “sweet spot” for screening smokers who have an elevated risk of lung cancer but not of dying from other diseases.[117,118,142]

The findings from this study indicate that smokers in the NLST who are at greatest risk of lung cancer do not benefit as much from CT screening relative to smokers at intermediate risk. Such a suggestion is corroborated by four relevant observations. First, the presence and severity of airflow limitation is associated with all-cause mortality, not just respiratory-related deaths.[20] Second the clinical variables used to assign lung cancer risk are closely aligned with those that underlie the risk of having airflow limitation.[102] This means both the risk of lung cancer and the risk of having COPD are closely aligned,[219] and difficult to separate with clinical variables alone. Third, the lung cancer-specific mortality reduction in

the NLST-ACRIN sub-group with airflow limitation is only half that seen for those with normal lung function (15% vs 28% respectively).[80] Fourth, airflow limitation (COPD) confers an increased risk of dying of lung cancer (HR=2.0), respiratory disease (HR=5.4) and cardiovascular disease (HR=1.7), where for the latter a comparable risk is conferred by the presence of diabetes (HR=1.9) (Table 5.6). We suggest that this "diluted" benefit from CT screening over CXR screening may result in part from a "competing risk" effect where smokers at very high risk of developing lung cancer are also at higher risk of dying of a non-lung cancer cause with shortened life expectancy.[102,120,121] This current study examines this relationship by showing divergence in the rate of lung cancer death versus non-lung cancer death as COPD grade worsens (Figure 5.1). A second important reason for a "diluted" benefit in smokers with COPD and at high risk of lung cancer is the lower surgical rate in these groups (Figure 5.5). The implication of this finding is that, screening smokers in the NLST with the greatest risk of lung cancer makes a smaller contribution to the overall benefit in terms of reducing lung cancer specific mortality.[80] It might also explain why, in the NLST, reduction in all-cause mortality (7%) is so much lower than reduction in lung cancer specific mortality (20%).[90] Two recent studies by Tanner et al., and Caverley et al., also suggest that competing mortality may diminish the benefits of screening in those at greatest risk,[112,142] and the results of our study support this suggestion. Importantly, in a recently published cost-effectiveness analysis it was observed that while risk targeting improves efficiency of screening, the gains from this efficiency are attenuated due to reduced quality life-years gained in those at greatest risk (shortened life expectancy).[122] This means the benefit of CT screening for lung cancer (surviving lung cancer) is not shared equally across all eligible smokers[96] and that "overtreatment",[233,234] and competing risks,[121] might be important in the "shared decision making" process mandated in lung cancer screening guidelines.[110,115]

Another important observation from this study is that for both the CXR and CT arms, the lung cancer mortality (%deaths) increased in those at greatest risk (tertile 3), particularly for those in the CT arm (i.e., convergence, Figure 5.7). While this finding might be explained by a "competing risk" effect,[121] it may also stem in part from the observations that worsening airflow limitation and increasing lung cancer risk are associated with more aggressive histological subtypes of lung cancer, that is, Small Cell, Squamous Cell and Non-small Cell.[93,219] Increasing lung cancer risk is also associated with a trend towards reduced early

stage lung cancer, less screen detected cancers (Table 5.5), and a corresponding reduction in surgical rates for lung cancer (Figure 5.5). We note that pre-existing COPD is associated with shorter doubling times in screen-detected lung cancer suggesting a more aggressive biology.[208,225] Additionally, lung cancer patients with severe COPD may have difficulties tolerating treatment (e.g., chemotherapy or radiotherapy induced pneumonitis). Collectively this might explain why smokers with co-existing severe or very severe COPD (GOLD 3-4) have greater lung cancer-specific deaths than those with normal lung function or only mild-to-moderate disease (GOLD 1-2). It might also underlie the halving in lung cancer specific deaths with CT screening relative to CXR in those with COPD compared to no COPD (15% vs 28%).[80]

We and others have previously suggested that the same factors promoting airflow limitation following decades of smoking may also promote coronary artery disease, other cancers and respiratory disease.[20,130] The presence of systemic inflammation, accelerated aging, heightened innate immunity and/or aberrant tissue remodelling from smoking (e.g., advanced emphysema) may be relevant here.[20] In such a setting, where inflammatory cytokines such as interleukin-1 and interleukin-6 are chronically elevated, accelerated ageing effects promote the development of premature (or unstable) coronary artery disease, susceptibility to chronic airways infection (pneumonia) and/or respiratory failure and development of cancers in sites other than the lung (Figure 5.2).[20,102,130] As currently accepted clinical risk variables for lung cancer appear to overlap with the risk of COPD and that COPD confers a high risk of death from either aggressive forms of lung cancer or other causes of death, we suggest that gene-based biomarkers reflecting lung cancer biology may be of use here.[49,118,235]

5.4.1 Strengths and limitations

There are several limitations to this study. First, although comparable with other CT screening studies,[213,107] spirometry in the NLST was performed as a pre-bronchodilator measurement rather than post-bronchodilator as recommended for clinical/diagnostic purposes. This means a proportion of subjects we have classified as having airflow limitation may have had asthma or COPD-asthma overlap rather than COPD (i.e., full, partial, or minimal reversible airflow limitation respectively). A further limitation of the study is that

people with a life expectancy of less than 5 years were excluded from the NLST.[90] This might explain the relatively low prevalence of GOLD grade 3 and 4 subjects in this study (6.6%) and due to under-powering, limit the findings in this COPD subgroup. A third limitation of this study is that we have compared mortality prevalence and mortality rates, combining those in both screening arms, but accept the findings may be biased by a survival effect related to duration of follow-up. That said, a strength of this study was that screening was undertaken in the first 3 years of the study followed by 3-4 years of outcome follow-up, so that screening effects on mortality are likely to be minimal (17/189 or 9% of lung cancer deaths were averted with CT screening relative to CXR in 10,054 NLST participants). Lastly, we have not further subtyped the specific (ICD-based) cause of death for each participant who died but rather grouped them by "organ" system.[178] This is for reasons of powering. It remains possible that someone dying of a "respiratory cause" may have had undiagnosed lung cancer, especially in the presence of severe airflow limitation. However, as death was a primary outcome of the NLST,[90] we feel that best attempts were used to correctly assign cause of death.

5.4.2 Conclusion

In summary, this study found that those in the NLST-ACRIN cohort at greatest risk of lung cancer have the highest prevalence of airflow limitation (spirometry defined COPD) and the greatest risk of dying of diseases other than lung cancer. Moreover, I found that those at greatest risk of lung cancer (Table 5.5), developed lung cancers that were more aggressive and less likely to undergo surgery (Figure 5.5). We believe this might explain the lower reduction in lung cancer specific deaths in the highest risk groups where COPD was most prevalent. We suggest that increasing risk of lung cancer does not always translate into an increased benefit of screening and that the presence of comorbid disease, especially airflow limitation (COPD), may attenuate the benefits of screening for lung cancer in some smokers eligible for screening.[102,113] Future modelling will need to address this issue, by better reflecting risks and benefits (not just lung cancer risk) with respect to outcomes.

Chapter 6 Airflow Limitation, Comorbid Disease, and Risk of Lung Cancer ($PLCO_{M2012}$): Effects on Lung Cancer Screening Outcomes in the National Lung Screening Trial

It is no use saying, 'We are doing our best.'
You have got to succeed in doing what is necessary.
-- Winston Churchill

6.1 Introduction

Annual computed tomography (CT) screening for lung cancer is widely recommended and funded in the United States.[90,218] This recommendation was based on the findings of the National Lung Screening Trial (NLST), which showed a 20% relative reduction in lung cancer specific mortality in those randomised to annual CT compared to annual chest-x-ray (CXR) screening.[90] Recent attention has turned to deciding who is best to screen for lung cancer in order to maximise the benefits and minimise the harms.[218] Although several screening models have been described,[86,123,236-237] only the $PLCO_{M2012}$ has been validated in the NLST.[123] While the efficiency of lung cancer screening can be enhanced by a targeted risk-based approach to selecting current and former smokers for screening, the gains are relatively modest.[236-238] There has been little consideration given to the relationship between lung cancer risk, lung cancer screening outcomes and co-existing comorbidity, in particular from airflow limitation that characterises chronic obstructive pulmonary disease (COPD).[102]

We have previously reported that lung cancer screening in the NLST involves older heavy smokers with 3-5 fold higher rates of pre-morbid cardiovascular and respiratory diseases, compared to screening populations for breast and colon cancer.[102] We have shown that spirometry defined airflow limitation, is a marker for premature death from all causes[20] and was present in about 35% of NLST participants.[93] In studies of both unscreened and screened lung cancer cases, worsening airflow limitation increases the risk of developing lung cancer in a linear relationship.[219] In a post-hoc analysis of a sub-group of the NLST, airflow limitation at baseline was associated with a two-fold greater risk of lung cancer, it was also associated with an almost halving in CT-based lung-cancer specific mortality relative to NLST participants with normal lung function (15% versus 28% respectively).[80] It is therefore

appropriate to propose that as the risk of lung cancer increases, so does the prevalence of airflow limitation, respiratory morbidity and deaths from both lung cancer and non-lung cancer causes.[102] The impact of comorbid disease on outcomes from lung cancer screening, particularly for those at greatest risk, are largely unknown but might be relevant to lung cancer treatment (especially surgery with curative intent) and lung cancer mortality. Two recent studies that have examined outcomes from lung cancer screening according to baseline risk, concluded that the benefits of screening are curvilinear where those at lowest (Quintile 1) and highest (Quintile 5) risk have attenuated benefit from screening relative to the middle three quintiles.[122,142] In this context "overtreatment" in lung cancer screening occurs when harms of work-up or treatment outweigh the benefits.[233,234] This current study examines the relationship between lung cancer risk, airflow limitation, comorbidity, competing mortality and surgery in the context of screening in the NLST. The preliminary results of this analysis have been published in abstract form.[239]

Hypothesis: In the context of high risk smokers, those at either end of the risk spectrum according to the $PLCO_{M2012}$ risk model, show reduced benefit from screening. Those at greatest risk had greater comorbid disease, worse lung function, reduced surgery, and reduced outcomes from screening.

Aim: To examine the relationship between risk of lung cancer and outcomes from CT-based screening.

Objective: To stratify the participants of the NLST at baseline according to their risk of lung cancer ($PLCO_{M2012}$ risk model), and then compare the outcomes from screening before and after screening randomisation.

6.2 Methods

6.2.1 Subjects

This is a secondary analysis of data from the National Lung Screening Trial (NLST) which has been described in detail elsewhere.[90] In the American College of Radiology Imaging Network (ACRIN) cohort of the NLST, participants from 23 centres agreed to undergo baseline pre-bronchodilator spirometry and blood sampling for biomarker analysis (N=10,054). From this cohort, 395 histology-confirmed lung cancer cases were diagnosed over the study period of

7.5 years although another 14 people were diagnosed with lung cancer at post-mortem.[239] Demographic data including pre-morbid diseases were collected through an extensive questionnaire and pulmonary function tests were performed (see below).

6.2.2 Pulmonary function testing

In the NLST-ACRIN cohort, pre-bronchodilator spirometry was measured at baseline screening (T0) in the majority of participants meeting the following criteria; no chest infection in the preceding 3 weeks and no use of a short-acting bronchodilator inhaler in the preceding 6 hours or long-acting bronchodilator in the preceding 24 hours. Those not meeting these criteria were rescheduled for spirometry testing at a later visit (prior to randomisation). Of the 10,054 subjects, 174 (2%) did not perform spirometry to a sufficient standard to be included in this study (N=9,880). The spirometry was measured by trained staff using a Spiropro spirometer (eResearchTechnology, GmbH, Germany). In the current study the severity of airflow limitation was defined according to the Global Initiative on Chronic Obstructive Lung Disease (GOLD) criteria grades 1-4 (www.GOLD.org accessed May 24, 2017).

6.2.3 Lung cancer and cause-specific mortality

Lung cancer cases included all those diagnosed during the trial (N=395), whether screen or non-screen detected (interval), or prevalent (diagnosed at T0 or during the first year) or incident lung cancers (diagnosed during subsequent years T1 to T6).[90] All lung cancer cases were confirmed on histological sampling according to accepted international classification criteria. Lung function results and mortality outcomes were available for 388 of the 395 lung cancer cases (98% of total). Cause of death was a primary outcome and ascertained through review of clinical records and death certification.[90]

6.2.4 Lung cancer risk, competing cause of death and CT outcomes

We stratified our NLST-ACRIN participants according to the $PLCO_{M2012}$ Model,[86,123] to assign quintile of risk and to assess whether there was a relationship between risk of lung cancer and prevalence of COPD and prevalence of doctor diagnosed pre-morbid disease. We also examined whether comorbid disease, airflow limitation and surgical rates were related to $PLCO_{M2012}$ risk and whether they affected the lung cancer-specific mortality reduction conferred by CT screening relative to CXR screening.

6.2.5 Statistical analysis

Differences in demographic variables, lung cancer incidence, and lung cancer outcomes after stratification by clinical phenotype (no airflow limitation, COPD, and lung cancer) or $PLCO_{M2012}$ risk quintiles (6-year lung cancer risk), were compared using chi-square or ANOVA for categorical or continuous variables as appropriate. Significant main effects were further explored as pairwise comparisons against the lowest quintile or the no-airflow limitation group as appropriate. The overall significance level within a table was maintained at the 5% level by calculating false discovery rate protected P values. Planned pairwise differences in lung cancer incidence were compared according to cause of death stratified by $PLCO_{M2012}$ risk, using the Mid-P exact test. Significance was defined as a two tailed $P<0.05$. Statistical analyses were performed using SAS (v9.4, SAS Institute Inc, Cary NC, USA) or STATA statistical software. WWW.Openepi.com was used to calculate confidence intervals for rates per patient year and proportions (Wilson method). Receiver operator analysis was performed using logistic regression and the area under the receiver operating curve estimated from the c-statistic.

6.3 Results

6.3.1 Baseline characteristics, spirometry and mortality in healthy smokers compared to those with airflow limitation or lung cancer

Compared to those with no airflow limitation (N=6,266), screening participants who had airflow limitations (N=3,225), or developed lung cancer (N=395), were older, more likely to be a current smoker, quit smoking for less time, had greater smoking duration and higher pack year exposure. The latter groups also had a marginally lower BMI and a higher rate of self-reported COPD, (i.e., responded “yes” to the question “Has a doctor ever told you that you have any of the following conditions……COPD, emphysema, chronic bronchitis, or adult asthma”), (Table 6.1). Apart from over-representation of self-reported COPD in those with airflow limitation or lung cancer (Table 6.1), the 3 groups were similar in regards to other common pre-morbid conditions, (cardiovascular disease and cancer). When a receiver operator curve (ROC) analysis was applied to the $PLCO_{M2012}$ model for risk of lung cancer (Area under the ROC curve (AUC) =0.67, 95% Confidence Interval (95% CI) =0.65-0.70), it was comparable and overlapping with the predictive utility for the presence of COPD based on airflow limitation (AUC=0.65, 95% CI=0.64-0.67).

Table 6.1 Demographic and Mortality data for NLST-ACRIN participants stratified according to baseline airflow limitation (presence or absence) and subsequent lung cancer (mutually exclusive)

	No Airflow limitation$^{\beta}$	COPD GOLD 1-4	Lung Cancer	P value (β referent)
N=9,886† (% total)	6,266 (63%)	3,225 (33%)	395 (4%)	
Demographics				
Age (mean)	61 (4.9)	63 (5.3)***	64 (5.3)***	<0.0001
Male Gender	3422 (55%)	1970 (61%)***	216 (55%)	<0.0001
Current Smoker	2869 (46%)	1752 (54%)***	226 (57%)	<0.0001
Pack Years (mean)	54 (22)	60 (25)***	63 (26)***	<0.0001
Cigarettes/day (mean)	28 (11)	29 (11)**	28 (11)	0.0010
Years Quit (mean)	4.00 (5.22)	3.10 (4.72)***	2.65 (4.49)***	<0.0001
Smoking duration years (mean)	39.20 (7.31)	42.10 (7.18)***	44.25 (7.18)***	<0.0001
Family history of Lung Cancer	24%	23%	27%	P=0.36
Personal history of COPD#	14%	32%***	30%***	P<0.0001
BMI (mean)	28.6 (5.2)	26.8 (4.8)***	26. (4.5)***	<0.0001
Education Level				
- High school or less	1740 (28%)	1027(32%)	111 (28%)	0.0043
-Post High school training	779 (12%)	397 (12%)	64 (16%)	(LC vs
-Some College	1485 (24%)	692 (21%)	94 (24%)	referent)
-College graduate	1068 (17%)	540 (17%)	55 (14%)	
-Postgraduate/Professional	1038 (17%)	485 (15%)	57 (14%)	
-Other/unknown	156 (2%)	84 (3%)	14 (4%)	
Lung Function				
FEV_1 % predicted (mean)	89.4 (15.5)	66.7 (19.4)***	74.4 (22.1)***	<0.0001
FEV_1/FVC (mean)	77.4 (4.8)	59.5 (9.3)***	66.8 (11.9)***	<0.0001
% COPD (baseline spirometry)	0%	100%	51%	
Pre morbid Disease (self-reported)				
COPD#	851 (14%)	1016 (32%)***	120 (30%)***	<0.0001
Pneumonia	1608 (26%)	994 (31%)***	126 (32%)**	<0.0001
Heart Disease	800 (13%)	444 (14%)	59 (15%)	0.23
Hypertension	2265 (36%)	1156 (36%)	155 (39%)	0.41
Stroke	177 (3%)	102 (3%)	12 (3%)	0.65
Diabetes	656 (10%)	258 (8%)***	38 (10%)	0.0006
Any cancer History	236 (4%)	151 (5%)	18 (5%)	0.09

†Unknown COPD status N=168 (9 Deaths); ‡ 14 lung cancers identified at post-mortem. # Composite of self-reported COPD, emphysema, chronic bronchitis or adult asthma.***P<0.0001, **P<0.001 Dunnetts test compared against no airflow limitation.$^{\Phi}$ Deaths adjusted for age and pack years. $^{\beta}$ % total deaths.

There were 688 deaths from this cohort of 10,054 NLST participants with 191 deaths (27%) occurring in those diagnosed with lung cancer and 238 deaths in those diagnosed with COPD (Table 6.2). Of those diagnosed with lung cancer, 92% died of their cancer. Compared to those with no airflow limitation, we found increased mortality in those with COPD from Cardiovascular disease (1.7-fold), Respiratory disease (5-fold) and Other Cancer (1.5-fold), where cardiovascular mortality accounted for over 30% of all deaths (Table 6.2).

Table 6.2 Mortality data for NLST-ACRIN participants stratified according to baseline airflow limitation (presence or absence) and subsequent lung cancer (mutually exclusive)

	No Airflow limitation$^{\beta}$	COPD GOLD 1-4	Lung Cancer	P value (β referent)
N=9,886† (% total)	6,266 (63%)	3,225 (33%)	395 (4%)	
Mortality (prevalence and rate)				
Patient years/Total deaths (95% CI)	39024/259	19799/238	2031/191	
Total Deaths /1000 Deaths /1000 pt yrs of follow-up$^{\Phi}$ Odds of Death	41 (37, 47) 6.6 (5.9, 7.5) 1 (ref)	74 (65, 83) 12.0 (10.6, 13.6) 1.85 (1.54, 2.22)	484 (435, 533) 94.0 (81.4, 108.1) 21.72 (17.20, 27.42)	<0.0001 <0.0001
- LC Deaths, n (% total deaths$^{\beta}$) -LC deaths /1000	7 ‡ (3% 1.3, 5.5) 1.1 (0.5, 2.3)	7‡ (3% 1.4, 5.9) 2.2 (1.1, 4.5)	175*** (92% 87, 95) 443 (394.8, 492.3)	<0.0001
- Cardiovascular (CVD) Deaths, n (%$^{\beta}$) -CVD deaths/1000	83 (32% 26.7, 38.0) 13.2(10.7, 16.4)	73 (31% 25.2, 36.8) 22.6 (18.0, 28.4)	7 (4% 1.8, 7.4) 17.7(8.6, 36.1)***	<0.0001
- Respiratory (Resp) Deaths, n (%$^{\beta}$) - Resp Death/1000	15 (6% 3.5, 9.3) 2.4 (1.5, 3.9)	43 (18% 13.7, 23.5) 13.3 (9.9. 17.9)***	2 (1% 0.3, 3.7) 5.1 (1.4, 18.3)	<0.0001
- Other Cancer (OC) Deaths, n (%$^{\beta}$) - OC Deaths/1000	79 (31% 25.2, 36.4) 12.6 (10.1, 15.7)	63 (26% 21.3, 32.4) 19.5 (15.3, 24.9)	4 (2% 0.8, 5.3) 10.1 (3.9, 25.7)***	<0.0001
- Other Deaths, n (%$^{\beta}$) - Other Deaths/1000	77 (30% 24.5, 35.6) 12.3 (9.8, 15.3)	52 (22% 17.1, 27.5) 16.1 (12.3, 21.1)	3 (2% 0.5, 4.5) 7.6 (2.6, 22.1)***	<0.0001

†Unknown COPD status N=168 (9 Deaths); ‡ 14 lung cancers identified at post-mortem. # Composite of self-reported COPD, emphysema, chronic bronchitis or adult asthma.***P<0.0001, **P<0.001 Dunnetts test compared against no airflow limitation.$^{\Phi}$ Deaths adjusted for age and pack years. $^{\beta}$ % total deaths.

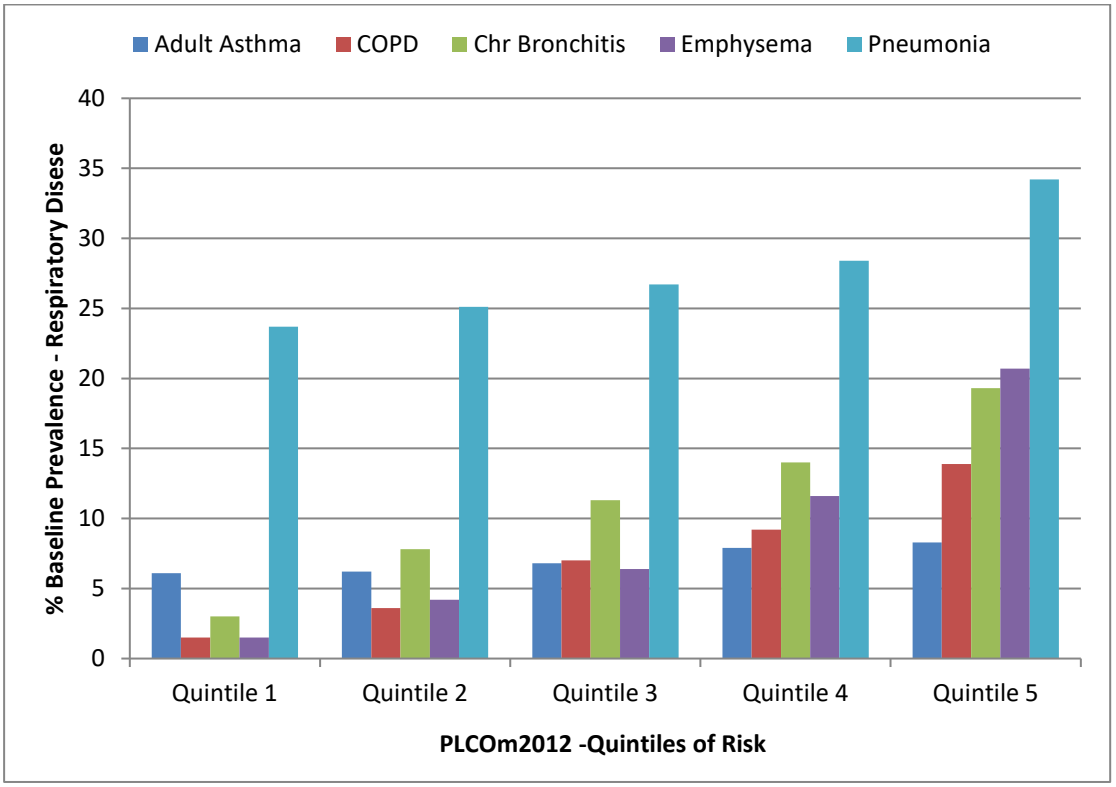
6.3.2 $PLCO_{M2012}$ risk quintiles, lung cancer risk variables, comorbid disease, and mortality

Using the $PLCO_{M2012}$ model, we stratified the NLST-ACRIN cohort according to risk quintiles (Tables 6.3 and 6.4, Figures 6.1-6.3). As expected, with increasing risk quintiles (lowest to highest risk, Quintiles 1 to 5 respectively), lung cancer risk variables increased with respect to age, current smoker status, pack years, cigarettes per day, smoking duration, family history of lung cancer and personal history of COPD. Similarly, the years quit and BMI decreased across quintiles 1-5. When the incidence of lung cancer and lung cancer mortality was compared with the prevalence of COPD according to quintiles of risk, we found a linear relationship (Table 6.3-Table 6.4, Figure 6.2). Specifically, as the risk of lung cancer increases across risk quintiles, lung cancer mortality also increases but not as steeply as for non-lung cancer causes (Figure 6.3). This divergence can be explained by a competing cause of death effect, where lung cancer deaths in Quintile 5 are substituted by premature deaths from cardiovascular disease, respiratory diseases, or other cancers (Table 6.4, Figure 6.3).

When deaths from specific causes are compared according to lung cancer risk, estimated from the $PLCO_{M2012}$ model,[86,123] there is a trend for increasing death rates for lung cancer, other cancers, and cardiovascular disease across each risk quintile (Table 6.4, Figure 6.3). The prevalence of spirometry-based COPD was 19%, 27%, 32%, 41% and 52% in Quintiles 1-5 respectively (Table 6.3). As the lung cancer deaths/1000 persons screened increases with each risk quintile, comparable increases in absolute terms are seen for deaths from cardiovascular disease and other cancers (Figure 6.3).

Figure 6.1 Baseline prevalence (%) of self-reported pre-morbid diseases according to $PLCO_{M2012}$ lung cancer risk quintiles for (a) Respiratory and (b) Other pre-morbid diseases (data Table 6.3)

(a) Respiratory pre-morbid disease at baseline



(b) Other pre-morbid disease at baseline

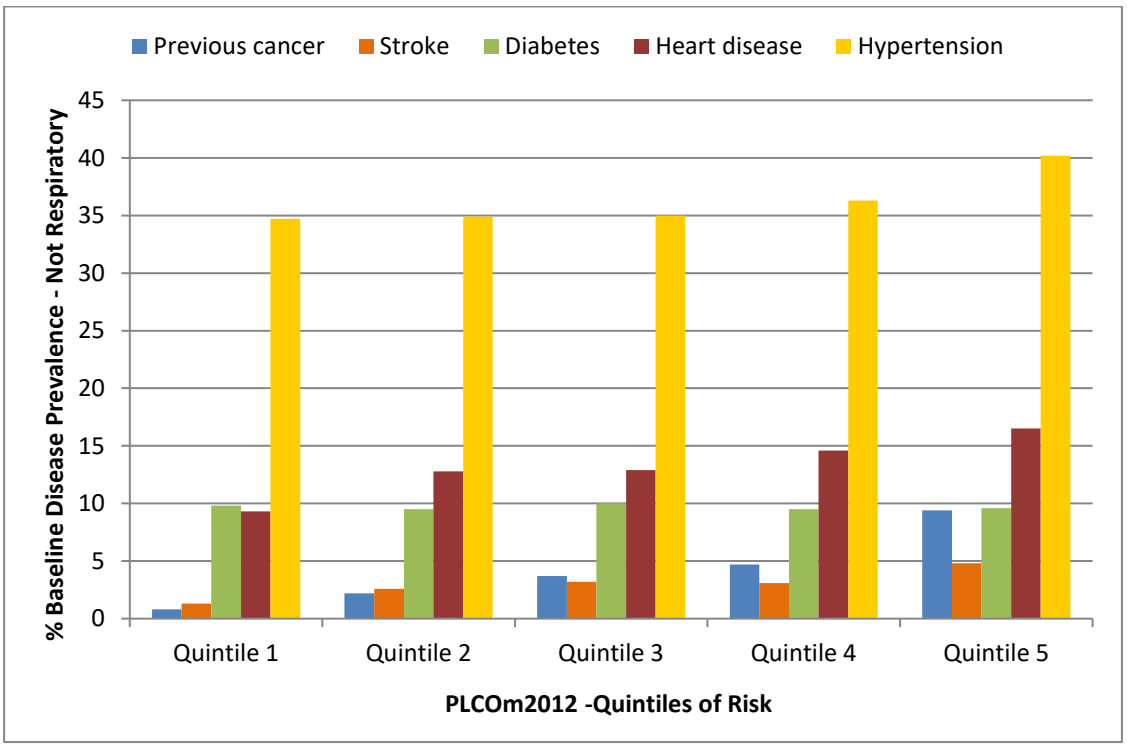


Figure 6.2 Comparison of COPD prevalence, lung cancer prevalence, and mortality from lung cancer, and non-lung cancer causes, in NLST-ACRIN participants stratified according to $PLCO_{M2012}$ quintiles (data from Tables 6.3 and 6.4)

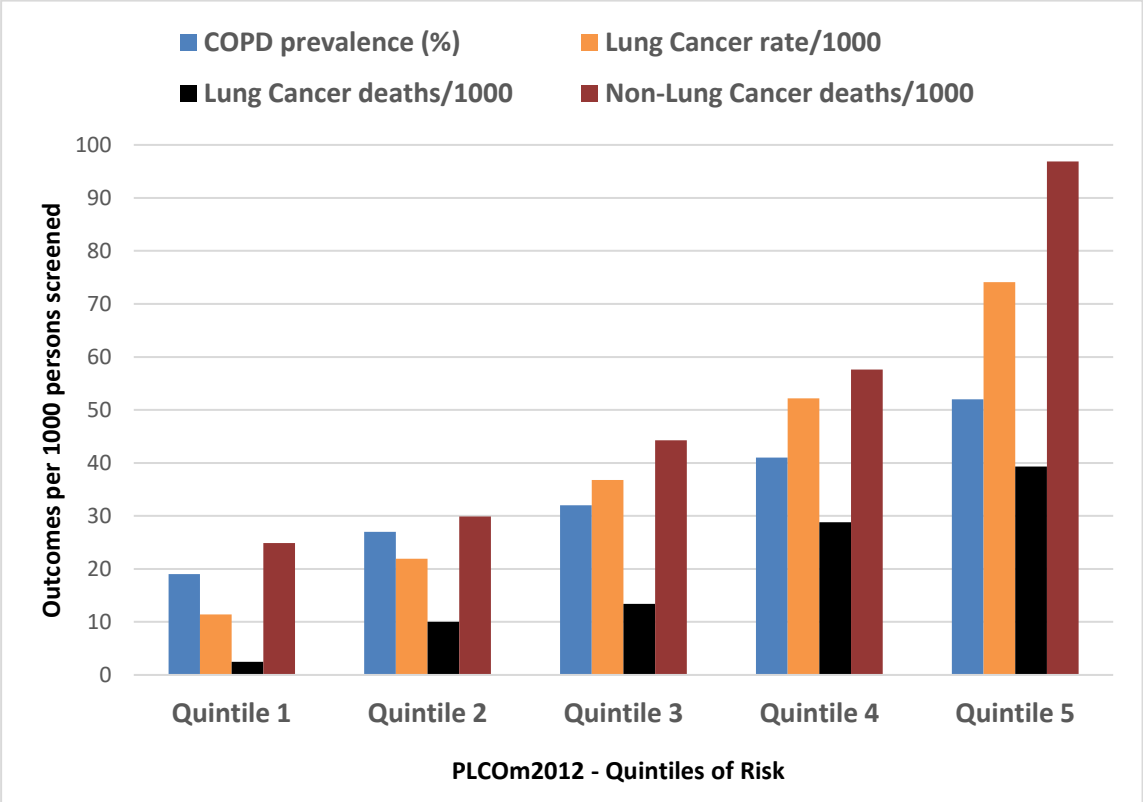
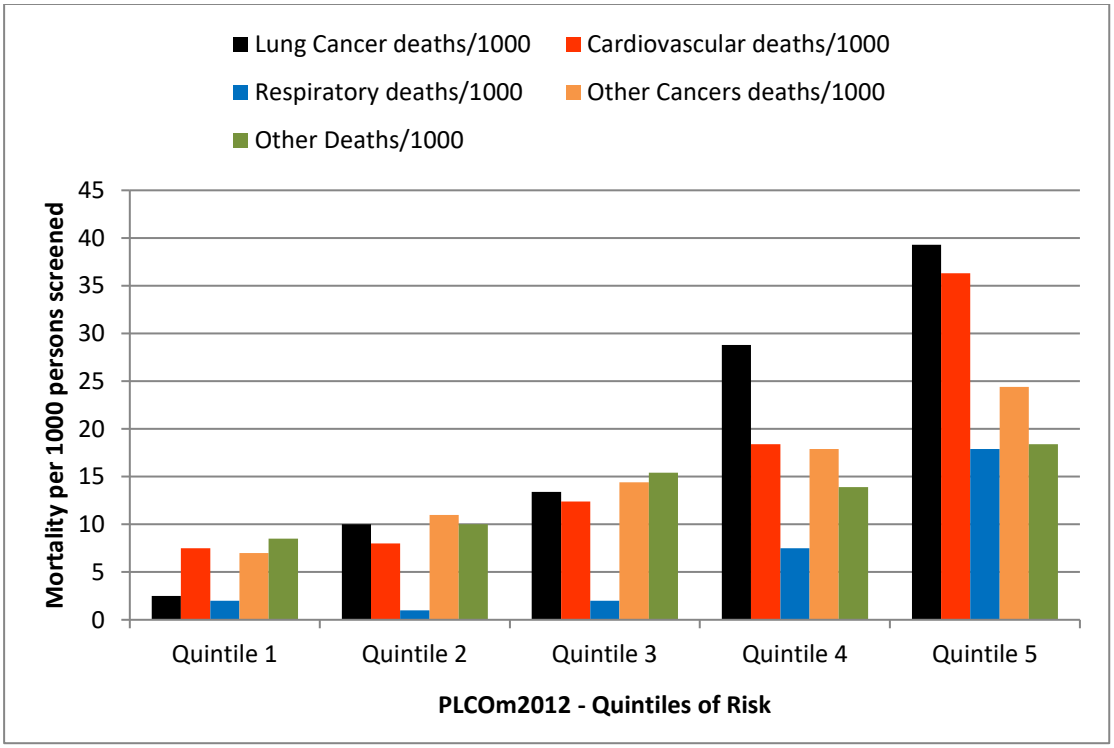


Figure 6.3 Lung cancer and cause-specific mortality per 1000 persons screened according to $PLCO_{M2012}$ risk quintiles (data from Table 6.4)



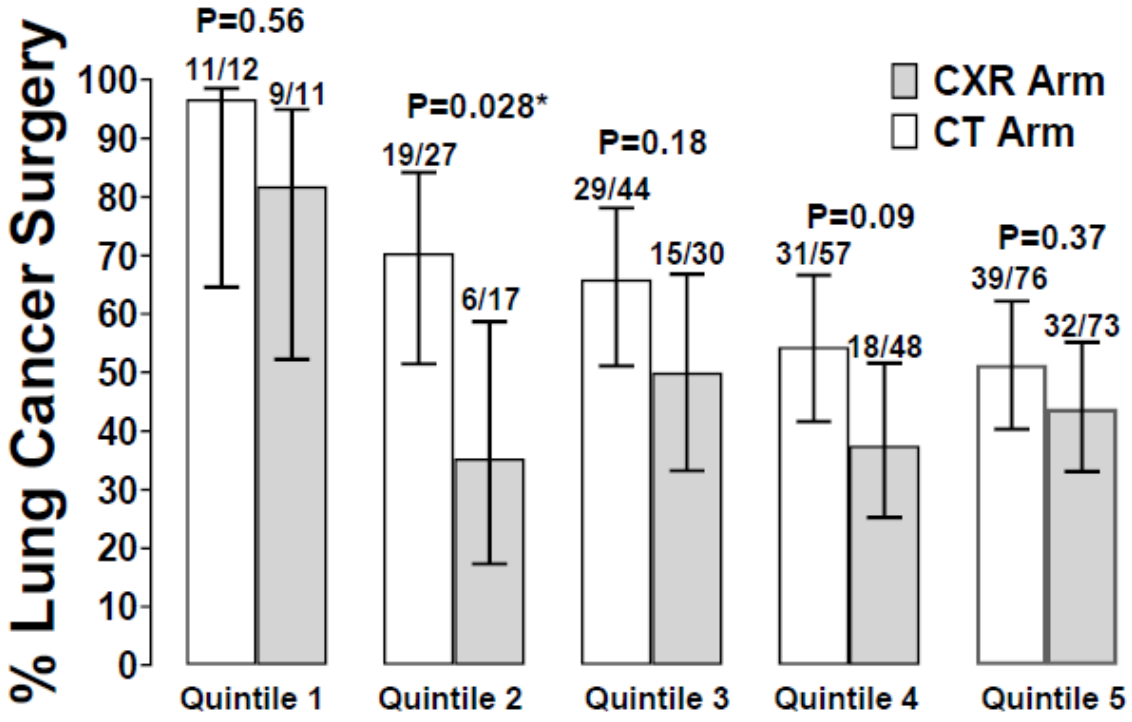
6.3.3 Lung cancer characteristics according to lung cancer risk ($PLCO_{M2012}$) quintile.

In Table 6.5 the diagnosis of lung cancer increases across the quintiles, but the surgical rate for lung cancer decreases while the lethality of lung cancer (lung cancer deaths/total lung cancers identified) increases. Relative to the other quintiles, in Quintile 1, lung cancer lethality is lowest (22%), is associated with less aggressive cancers of small cell and squamous cell subtype (26%), and more indolent cancers of Adenocarcinoma/BAC subtype (61%). In this quintile, the surgical rate was highest (87%); early stage 1-2 disease greatest (70%) and the majority of cancers were diagnosed by screening (78%), (Table 6.5). The vast majority of cancers diagnosed in Quintile 1 (87%) were identified during the screening interval (T0-T2) with a minority (13%) diagnosed during the follow-up period (Table 6.5). In contrast for Quintile 5, the prevalence of COPD is greatest (52%), respiratory comorbid disease is highest (Table 6.3), the non-small cell/large cell histology subgroup was greatest (29%) and the majority of lung cancers diagnosed at post-mortem was observed (9/14 or 64%).

6.3.4 Lung cancer diagnoses, surgical rate, and mortality according to screening arm

When we compared the outcomes of CT versus CXR screening in our NLST-ACRIN participants, stratified by quintile of lung cancer risk according to the $PLCO_{M2012}$ model,[86,123] we found differences in lung cancer surgical rates and lung cancer specific mortality conferred by CT screening (Table 6.6). As described above, with increasing lung cancer risk ($PLCO_{M2012}$ quintiles 1-5), the lung cancer diagnosis rate increased, the surgical rate decreased and lung cancer mortality increased. The differential surgical rates (absolute and relative, Figure 6.4) in quintiles 1 and 5 were not different between the CT and CXR arms. This was in contrast to quintiles 2-4 where randomisation to CT was associated with higher surgical rates than in the CXR arm (Figure 6.4). In Quintile 5, a comparison of lung cancer deaths according to risk quintile and screening arm showed no differences (Table 6.6), regardless of how lung cancer deaths were represented, between those in the CT and CXR arms (Figure 6.5). We note 76% of the lung cancer deaths averted in the CT arm (13/17) were from quintiles 2-4 (Table 6.6). Lung cancer deaths averted in the CT arm compared to CXR in quintile 1 (Table 6.6) was high in relative terms despite the small difference in surgical rate (Table 6.6, Figure 6. 4).

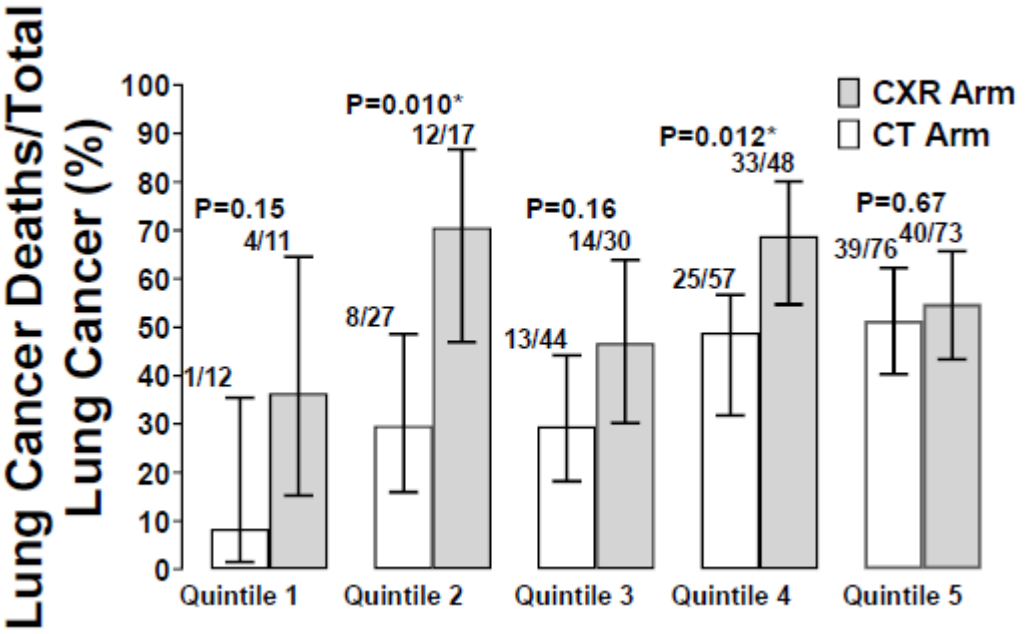
Figure 6.4 Comparison of Lung cancer surgical rates (%) in the computed tomography (clear) and chest –X-ray (shaded) screening arms, according to $PLCO_{M2012}$ risk quintiles (data from Table 6.6)



Data are % undergoing surgery for lung cancer by screening arm (95% CI), with absolute numbers and P for the within quintile pairwise comparison at the top of each quintile

Figure 6.5 Comparison of Lung cancer deaths by computed tomography (clear) and chest –X-ray (shaded) according to $PLCO_{M2012}$ risk quintiles (data from Table 6.6)

(a) Lung Cancer deaths as a % of total lung cancers



(b) Lung Cancer deaths as a % of total deaths

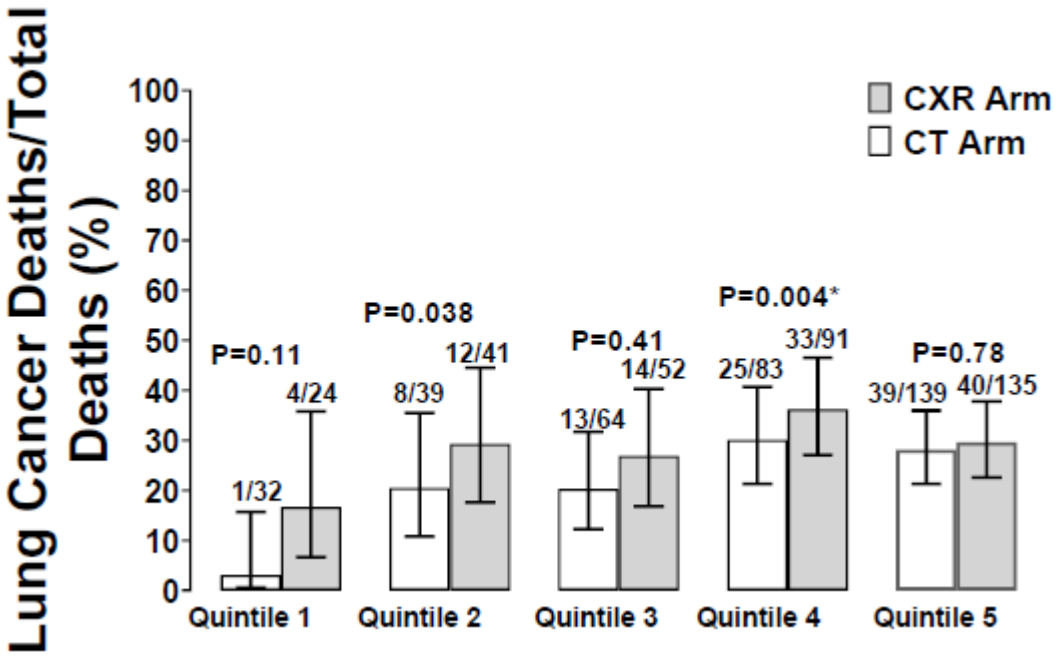


Table 6.3 Lung cancer risk variables, lung function and comorbid disease according to lung cancer risk ($PLCO_{M2012}$) quintile

$PLCO_{M2012}$ Quintiles	**Quintile 1**	**Quintile 2**	**Quintile 3**	**Quintile 4**	**Quintile 5**	**P value**
N=10,054	**N=2010**	**N=2010**	**N=2010**	**N=2012**	**N=2012**	
6-year lung cancer risk (%)	0.27-1.44	1.45-2.19	2.20-3.21	3.22-5.19	5.20-40.97	
Demographics						
Age, mean (SD)	58.0 (3.0)	59.6 (3.9)***	61.2 (4.4)***	63.3 (4.6)***	66.5 (4.6)***	<0.0001
Race -European	1886 (94%)	1896 (94%)	1869 (93%)	1878 (93%)	1844 (92%)**	0.011
Male Gender	1140 (56.7%)	1142 (56.8%)	1117 (55.6%)	1121 (55.7%)	1175 (58.4%)	0.38
Current Smoker	438 (21.8%)	894 (45%)***	1026 (51%)***	1169 (58%)***	1411 (70%)***	<0.0001
Pack Years, mean (SD	43 (15)	49 (18)***	55 (20)***	61 (23)***	71 (28)***	<0.0001
Cigarettes/day, mean (SD)	27 (10)	27 (11)	28 (5)	29 (11)***	30 (11)***	<0.0001
Years Quit, mean (SD)	6.7 (5.5)	3.9 (5.1)***	3.4 (4.9)***	2.7 (4.9)***	1.4 (3.3)***	<0.0001
Years Smoked, mean (SD)	33.2 (5.6)	37.4 (5.2)***	40.0 (5.1)***	43.0 (5.2)***	48.2 (5.8)***	<0.0001
Family history of Lung Cancer	132 (7%)	286 (14%)***	473 (24%)***	606 (30%)***	880 (44%)***	<0.0001
History of COPD# (%)	204 (10%)	317 (16%)***	431 (21%)***	558 (28%)***	813 (40%)***	<0.0001
BMI, mean (SD)	30.1 (5.9)	28.3 (5.0)***	27.9 (4.8)***	27.0 (4.6)***	26.1 (4.3)***	<0.0001
Lung Function						
FEV_1 %predicted, mean (SD)	88.6 (16.9)	84.0 (18.8)***	82.0 (20.0)***	78.6 (20.5)***	73.7 (21.3)***	<0.0001
FVC %predicted, mean (SD)	90.4 (15.5)	87.9 (16.8)***	87.3 (17.3)***	85.7 (18.0)***	82.6 (18.6)***	<0.0001
FEV_1/FVC, mean (SD)	75.2 (8.5)	72.9 (9.5)***	71.5 (10.9)***	69.2 (11.1)***	66.7 (11.8)***	<0.0001
Airflow limitation (COPD 1-4)	373 (18.6%)	532 (26.5%)**	642 (31.9%)**	824 (41.0%)**	1045 (51.9%)**	<0.0001
Pre-morbid Diseases‡ (%)						
Adult asthma	122 (6.1%)	124 (6.2%)	136 (6.8%)	158 (7.9%)*	167 (8.3%)**	0.015
Chronic Bronchitis	61 (3.0%)	156 (7.8%)***	227 (11.3%)***	282 (14.0%)***	389 (19.3%)***	<0.0001
COPD	30 (1.5%)	73 (3.63%)***	140 (7.0%)***	186 (9.2%)***	279 (13.9%)***	<0.0001
Emphysema	30 (1.5%)	84 (4.2%)***	128 (6.4%)***	234 (11.6%)***	416 (20.7%)***	<0.0001
Pneumonia	477 (23.7%)	504 (25.1%)	536 (26.7%)*	571 (28.4%)**	687 (34.1%)***	<0.0001
Heart Disease or Heart attack	186 (9.3%)	257 (12.8%)**	259 (12.9%)**	294 (14.6%)***	331 (16.5%)***	<0.0001
Diabetes	196 (9.8%)	190 (9.5%)	201 (10.0%)	191 (9.5%)	194 (9.6%)	0.98
Hypertension	698 (34.7%)	702 (34.9%)	703 (35.0%)	731 (36.3%)	808 (40.2%)**	0.0011
Stroke	26 (1.3%)	53 (2.6%)**	64 (3.2%)***	62 (3.1%)***	97 (4.8%)***	<0.0001
Past cancer	15 (0.8%)	45 (2.2%)**	74 (3.7%)**	94 (4.7%)***	189 (9.4%)***	<0.0001

12 subjects had missing height and incomplete spirometry. # Composite of self-reported COPD, emphysema, chronic bronchitis, or adult asthma. ‡Self-reported on baseline questionnaire

Table 6.4 Lung cancer diagnosis and cause-specific mortality according to lung cancer risk ($PLCO_{M2012}$) quintile

$PLCO_{M2012}$ Quintiles	**Quintile 1**	**Quintile 2**	**Quintile 3**	**Quintile 4**	**Quintile 5**	**P value**
N=10,054	**N=2010**	**N=2010**	**N=2010**	**N=2012**	**N=2012**	
6-year lung cancer risk (%)	0.27-1.44	1.45-2.19	2.20-3.21	3.22-5.19	5.20-40.97	
Screening randomisation						
Randomised CT	998 (49.7%)	1038 (51.6%)	989 (49.2%)	1006 (50.0%)	1003 (49.8%)	0.60
Lung cancer diagnosis						
Patient years	12636.8	12588.1	12465.5	12255.5	11977.6	
Lung cancer, n (/1000) (% 95% CI)	23 (11) (8, 17)	44 (22) (16, 29)*	74 (37) (29, 46)***	105 (52) (43, 63)***	149 (74) (63, 86)***	<0.0001
Lung cancer rate per1000 person years$^{\Phi}$ (% 95% CI)	1.82 (1.18, 2.69)	3.50 (2.54, 4.69)**	5.94 (4.66, 7.45)***	8.57 (7.01, 10.4)***	12.4 (11.5, 27.3) ***	<0.0001
Mortality						
Lung cancer, n (/1000) (% 95% CI) Rate per 1000 person years$^{\Phi}$ (95% CI)	5 (2.5) (0. 9, 6.0) 0.40 (0.13, 0.92)	20 (9.9) (6, 15)** 1.59 (0.97, 2.45)**	27 (13) (9, 20)*** 2.17 (1.43, 3.15)***	58 (29) (22, 37)*** 4.73 (3.59, 6.12)***	79 (39) (32, 49)*** 6.60 (5.22, 8.22)***	<0.0001 <0.0001
Non-lung cancer, n (/1000) (% 95% CI)	50 (25) (19, 33)	60 (30) (23, 33)	89 (44) (36, 54)**	116(58) (48, 69)***	195 (97) (85, 111)***	<0.0001
Respiratory, n (/1000) (% 95% CI)	4 (1.9) (0.6, 5.3)	2 (0.9) (0.02, 3.9)	4 (1.9) (0.6, 5.3)	15 (7.5) (4.4, 12.4)**	36 (18) (13, 25)***	<0.0001
Circulatory, n (/1000) (% 95% CI)	15 (7.5) (4.4, 12.4)	16 (8.0) (4.8, 13)	25 (12) (8, 18)	37 (18) (13, 25)**	73 (36) (29, 45)***	<0.0001
Other cancer, n (/1000) (% 95% CI)	14 (7.0) (4.0, 17.8)	22 (11) (7, 17)	29 (14) (10.0, 20.7)*	36 (18) (12.9, 24.8)**	49 (24) (18, 32)***	<0.0001
Other cause, n (/1000) (% 95% CI)	17 (8.5) (5.2, 13.6)	20 (10) (6.4, 15.4)	31 (15) (10.8, 21.9)*	28 (13.9) (9.6, 20.1)	37 (18.4) (13.3, 25.3)**	0.0001

$^{\Phi}$ Mortality adjusted for age and pack years.

Table 6.5 Lung cancer characteristics and mortality according to lung cancer risk ($PLCO_{M2012}$) quintile

$PLCO_{M2012}$ Quintiles	**Quintile 1**	**Quintile 2**	**Quintile 3**	**Quintile 4**	**Quintile 5**
Lung Cancer (LC)					
LC Diagnosis n (/1000) (% 95% CI)	23 (11) (7.5, 17.2)	44 (22) (16, 29)*	74 (37), (29, 46)***	105 (52) (43, 63)***	149 (74) (63, 86)***
Surgery, n (% LC by quintile) (% 95% CI)	20 (87%) (67, 96)	25 (57%) (42, 70)*	44 (60%) (48, 70)*	49 (47%) (37, 56)**	71 (48%) (40, 56)**
LC Death n (% of total LC) (% 95% CI)	5 (22%) (9, 42)	20 (45%) (32, 60)	27 (36%) (26, 48)	58 (55%) (46, 64)**	79 (53%) (45, 61)**
Histology					
1. Small Cell, n (% by quintile) (% 95% CI)	2 (8.7%) (1.2, 28.0)	5 (11.4%) (4.5, 24.4)	11 (14.9%) (8.3, 24.9)	14 (13.3%) (8.0, 21.3)	17 (11.4%) (7.2, 17.6)
2. Squamous Cell, n (%) (%, 95% CI)	4 (17.4%) (6.4, 37.7)	12 (27.3%) (16.2, 42.0)	16 (21.6%) (13.7, 32.4)	29 (27.6%) (19.9, 36.9)	31 (20.8%) (15.0, 28.1)
3. Adenocarcinoma/BAC, n (%) (%, 95% CI)	14 (60.9%) (40.8, 77.8)	24 (54.5%) (40.1, 68.3)	36 (48.6%) (37.6, 59.8)	45 (42.9%) (33.8, 52.4)	58 (38.9%) (31.5, 46.9)
5. Large/Non-small Cell, n (%) (%, 95% CI)	3 (13.0%) (4.5, 32.1)	3 (6.8%) (2.3, 18.2)	10 (13.5%) (7.5, 23.1)	17 (16.2%) (10.4, 24.4)	43 (28.9%) (22.2, 36.6)
7. Other/Unknown, n (%) (%, 95% CI)	0 (0, 12.2)	0 (0, 6.6)	1 (1.4%) (0, 8.0)	0 (0, 2.8)	0 (0,2.0)-
Lung cancer stage					
Stage 1-II, n (% by quintile) (%, 95% CI)	16 (69.6%) (48.9, 84.6)	22 (50.0%) (35.8, 64.2)	39 (52.7%) (41.5, 63.7)	45 (42.9%) (33.8, 52.4))	73 (49.0%) (41.1, 56.9)
Stage III-IV, n (%) (%, 95% CI)	6 (26.1%) (12.6, 46.5)	22 (50.0%) (35.8, 64.2)	32 (43.2%) (32.6, 54.6)	59 (56.2%) 46.7, 65.3	73 (49.0%) (41.1, 56.9)
Occult Ca /No assessment possible, (%, 95% CI)	1 (4.3%) (0, 23)	0 (0, 6.6)	3 (4.1%) (0.9, 11.7)	1 (0.95%) (0, 5.7)	3 (2.0%) (0.4, 6.0)
Screening detection					
1. Positive screen (T0-T2) (% by quintile) (95% CI)	18 (78.3%) (58, 92)	26 (59.0%) (44,72)	38 (51.4%) (40.2, 62.4)	49 (46.7%) (37.4, 56.2)	75 (50.3%) (42.4, 58.3)
2. Interval cancers (T0-T2) (%) (95% CI)	2 (8.7%) (0, 12)	6 (13.6%) (1.8, 17.4)	7 (9.5%) (2.9, 14.9)	12 (11.4%) (5.6, 17.8)	14 (9.4%) (5.2, 14.4)
4. Post screen (%) (95% CI)	3 (13.0%) (4.5, 32.1)	12 (27.3%) (16.4, 41.9)	29 (39.2%) (29.9, 50.6)	44 (42.0%) (32.9, 51.5)	60 (40.3%) (32.7, 48.3)
Year Lung cancer detected					
T0-T2 – screening (% by quintile) (95% CI)	20 (87.0%) (67.9, 95.5)	32 (72.7%) (58.2, 83.7)	45 (60.8%) (49.4, 71.1)	61 (58.1%) (48.5, 67.1)	89 (59.7%) (51.7, 67.3)
T3-T7 – follow-up 1 (%) (95% CI)	3 (13.0%) (4.5, 32.1)	12 (27.3%) (16.4, 41.9)	29 (39.2%) (28.9, 50.6)	44 (41.9%) (32.9, 51.5)	60 (40.3%) (32.7, 48.3)

Table 6.6 Lung cancer diagnoses, surgical rate, and mortality according to screening arm

	Lung Cancer Risk ($PLCO_{m2012}$)					
	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	P value
N=10,054 (total)	N=2,010	N=2,010	N=2,010	N=2,012	N=2, 012	-
Screening Method						
Randomised CT, N (%)	N=998 (50%)	N=1038 (52%)	N=989 (49%)	N=1006 (50%)	N=1003 (50%)	0.60
Randomised CXR, N (%)	N=1012 (50%)	N=972 (48%)	N=1021 (51%)	N=1006 (50%)	N=1009 (59%)	
Lung Cancer (LC)						
LC Diagnosis, N (/1000) (95% CI)	23 (11) (7.5, 17)	44 (22) (16, 29)*	74 (37) (29 46)***	105 (52) (43,63)***	149 (74) (63, 86)***	<0.0001
LC Surgery, N (% of LC by quintile) (% 95% CI)	20 (87%) (67, 96)	25 (57%) (42, 70)*	44 (60%) (48, 70)*	49 (47%) (37, 56)**	71 (48%) (40, 56)**	0.0026
LC Death, N (% of total LC) (% 95% CI) (lethality)	5 (22%) (9, 42)	20 (45%) (32, 60)	27 (36%) (27, 48)	58 (55%) (46, 64)**	79 (53%) (45, 61)**	0.075
LC Diagnoses by arm						
CT	12	27	44	57	76	
CXR	11	17	30	48	73	
Excess LC in the CT arm	+1 (p=0.84)	+10 (P=0.22)	+14 (P=0.076)	+9 (P=0.42)	+3 (P=0.80)	-
LC Surgical Rate by arm						
CT	11/12 (91.6%)	19/27 (70.4%)	29/44 (65.9%)	31/57 (54.4%)	39/76 (51.3%)	
CXR	9/11 (81.8%)	6/17 (35.3%)	15/30 (50.0%)	18/48 (37.5%)	32/73 (43.8%)	
Pairwise comparison by quintile	P=0.56	P=0.03	P=0.18	P=0.09	P=0.37	-
LC Deaths/ total LC) by arm						
CT	1/12 (8.3%)	8/27 (29.6%)	13/44 (29.5%)	25/57 (43.9%)	39/76 (51.3%)	
CXR	4/11 (36.4%)	12/17 (70.6%)	14/30 (46.7%)	33/48 (68.8%)	40/73 (54.8%)	
Pairwise comparison by quintile	P=0.15	P=0.01	P=0.16	P=0.012	P=0.67	-
Absolute benefit with CT						
LC deaths averted with CT	3	4	1	8	1	
LC Deaths/ total deaths by arm						
Total Deaths CT	31	39	64	83	139	
CT	1/32 (3.2%)	8/39 (20.5%)	13/64 (20.3%)	25/83 (30.1%)	39/139 (28.1%)	
Total Deaths CXR	24	41	52	91	135	
CXR	4/24 (16.7%)	12/41 (29.3%)	14/52 (26.9%)	33/91 (36.3%)	40/135 (29.6%)	
Pairwise comparison by quintile	P=0.11	P=0.04	P=0.41	P=0.004	P=0.78	-
LC Deaths / Non-LC Deaths by arm						
CT	1/30 (3.3%)	8/31 (25.8%)	13/51 (25.5%)	25/62 (40.3%)	39/107 (36.4%)	
CXR	4/20 (20.0%)	12/29 (41.4%)	14/38 (36.8%)	33/59 (55.9%)	40/103 (38.8%)	

CT=Computed tomography, CXR=chest x-ray.

6.4 Discussion

Our post-hoc analysis of the NLST-ACRIN cohort of 10,054 NLST screening participants shows that the $PLCO_{M2012}$ risk model that predicts lung cancer includes risk variables underlying a greater tendency for having airflow limitation. We also show that those at greatest risk of lung cancer (Quintile 5) have more self-reported respiratory morbidity, (i.e. responded "yes" to the question "Has a doctor ever told you that you have any of the following conditions…. COPD, emphysema, chronic bronchitis, or adult asthma"); greater prevalence of airflow limitation and greater deaths from both cardiovascular and respiratory causes. For screening participants in this high risk group who develop lung cancer, there are lower surgical rates and a lower reduction in lung cancer deaths with CT. This finding challenges the assumption that greatest lung cancer risk confers greatest benefit from CT-based lung cancer screening and that the benefits versus harms of screening must be carefully considered in those at the greatest risk (Quintile 5). In this study the analysis also found that while those at lowest risk (Quintile 1) had high rates of stage 1-2 disease (70%) and screen-detected lung cancer (78%), their lung cancers were found almost exclusively with screening (87%) but not on post-screening follow-up. These findings suggest that the balance of benefit and harm across the lung cancer risk spectrum is not a simple linear relationship and that NLST screening participants of intermediate risk (20-80%, Quintile 2-4), according to risk models like $PLCO_{M2012}$, generally gain more from screening than those at lowest and highest risk.

I propose that estimating the risk of lung cancer through currently accepted risk models,[86,123] does not necessarily accurately reflect the potential benefits of screening because the relationship between the risk of lung cancer and outcomes from screening is not linear. This study shows that the clinical variables underlying risk of lung cancer are comparable to risk variables underlying the presence of airflow limitation (Table6.1).[197] This is important because with a greater prevalence of airflow limitation, the risk of dying of lung cancer flattens relative to the risk of dying of a non-lung cancer cause, which diverges in favour of the latter (Figure 6.2). When applying the $PLCO_{M2012}$ risk model for lung cancer to our NLST-ACRIN sub-cohort, we found its predictive utility, based on a receiver-operator-curve analysis (AUC) was only marginally worse for COPD (AUC=0.65) than it was for lung cancer (AUC=0.67) with overlapping confidence intervals. This is not surprising as age, pack years, smoking duration, BMI, self-reported COPD, and current smoking levels were very similar in

those with spirometry-defined airflow limitation compared to those who developed lung cancer (Table 6.1). These similarities highlight just how closely a smoker's risk for COPD and lung cancer overlap.[34,233,239] This observation means those at greatest risk of lung cancer (Quintile 5) also have the highest prevalence of COPD, greatest level of comorbid disease and the greatest likelihood of dying of a non-lung cancer death, relative to those with normal lung function (Table 6.4, Figures 6.1-6.3). Those at greatest risk have less surgery and an attenuated reduction in lung cancer deaths with CT screening (Table 6.5). These findings would explain our past observation that NLST participants with normal lung function gain a greater reduction in lung cancer mortality than those with airflow limitation (COPD), (28% versus 15% respectively).[80] These findings challenge the current view that using a risk-based approach alone, or a targeted case-finding approach, represent the only ways to improve screening efficiency.[238]

6.4.1 Screening and stage shift

The benefits of CT screening for lung cancer are based on much more than achieving a favourable stage shift in lung cancers where surgery confers a significant likelihood of long-term survival.[102] We have previously suggested that much of the excess cancers detected by CT screening compared to chest x-ray reflects "histology-shift" favouring more indolent forms of adenocarcinoma (formerly bronchioloalveolar carcinoma).[93] These cancers have been linked to over-diagnosis and of note is that in the Quintile 1 risk group 20/23 or 87% of lung cancers were found during the screening interval (T0-T2) with the remaining 13% diagnosed over the following 5 years of follow-up (Table 6.5). We have previously suggested that the presence of airflow limitation (or COPD) plays a part in a propensity to more aggressive lung cancer (unspecified non-small cell).[93] Some investigators have shown that COPD is associated with lung cancers of shorter doubling time and histology of small cell and squamous cell subtype.[208,212,240] Not only does co-existing airflow limitation affect lung cancer biology, but also the likelihood of surgery. We have shown that in the NLST-ACRIN cohort, the surgical rate for early stage cancer is dramatically affected by the presence of COPD (reducing it).[102] Given that COPD is related to an increase in all-cause mortality,[20] it is not surprising that non-lung cancer deaths are much greater in those with airflow limitation (Table 6.4, Figure 6.2-6.3). Collectively, this suggests that airflow limitation is the most important comorbid disease in NLST participants undergoing lung cancer screening. This is

because COPD is associated with greater risk of lung cancer, more aggressive lung cancer and less surgery. This affect from COPD is not recognised in comorbidity scores like the Charlson Comorbidity Score where the presence of COPD is based on history only (not based on spirometry assessment) and is scored equally with other diseases.[114] We suggest that the process of predicting lung cancer risk, without any consideration to lung cancer biology or underlying airflow limitation, makes predicting outcomes from lung cancer screening unreliable.[97]

Two recent studies have examined the benefits of CT-based lung cancer screening according to risk and concluded that there may be an attenuation of benefits from screening those at greatest risk.[122,142] The first by Kumar and colleagues was a cost-effectiveness analysis and concluded that although risk-based targeting in screening may improve screening efficiency, the benefits in those at greatest risk are attenuated when expressed in life-years saved.[122] They suggested this was because those at greatest risk were older, with greater smoking exposure and a greater prevalence of COPD. Moreover, lung cancer death reduction was curvilinear across the $PLCO_{M2012}$ risk quintiles favouring quintiles 2 to 4 (Table 6.6). This effect was also shown by Kovalchik and colleagues, where the relative reduction in lung cancer deaths prevented with CT were 5%, 29%, 29%, 43% and 18% across quintiles 1 to 5 respectively.[123] The second study by Caverly and colleagues showed that the "Lifetime QALY" gains with CT-based lung cancer screening were curvilinear across the lung cancer risk spectrum, again highlighting that for those at greatest risk (top 20% risk group equivalent to quintile 5), competing mortality from non-lung cancer deaths would attenuate the benefits from screening.[142] Neither of these studies considered that increasing frequency of COPD, differing lung cancer characteristics (histology and stage), and lower surgical rate might exist across the risk quintiles (Table 6.4, Figure 6.4). In this current study I show that reductions in lung cancer deaths averted with CT in quintile 5 (Figure 6.5) are modest relative to the number of lung cancers diagnosed and that this is associated with low surgical rates (Table 6.5, Figure 6.4), and highest rates of interval or post-screen cancers, (Table 6.4, Table 6.5). As previously suggested, airflow limitation is associated with a '"histology-shift" in the NLST favouring more aggressive lung cancers less amenable to early detection with CT screening.[93] Based on the results of these studies, it appears that lung cancer screening is most beneficial in NLST screening participants who are in quintiles 2 to 4 (20-80% risk) as they have both an

elevated risk of lung cancer and the greatest likelihood to undergo surgery with long-term survival.[97] We advocate an outcomes-based approach be considered in the shared-decision making discussions, particularly for those in risk quintile 5 where the benefits of screening may be attenuated due to comorbid disease and less treatable lung cancer. An outcomes-based approach can be examined through post-hoc analyses of the NLST where biomarker data or clinical phenotyping (see study in Chapter 4) might better help determine who would benefit most from screening.[102,134,241]

6.4.2 Limitations

There are several limitations to this study. First, although comparable with other CT screening studies,[107,213] spirometry in the NLST was performed as a pre-bronchodilator measurement rather than post-bronchodilator as recommended for clinical or diagnostic purposes. A second limitation is that although the cohort size of 10,000 subjects is large, differences in outcomes, stratified by quintiles and screening arm, are small in absolute terms. A further limitation of the study is that follow-up post active screening rounds, was for only 3 years on average.[90] Lastly, no further subtyping of the specific (ICD code-based) cause of death for each participant who died was done, but rather cause of death was grouped by "organ" system.[178] This is for reasons of powering. It remains possible that someone dying of a "respiratory cause" may have had undiagnosed lung cancer, especially in the presence of severe airflow limitation. However, as death was a primary outcome of the NLST,[90] best attempts were used to correctly assign cause of death.

6.4.3 Conclusion

In summary, this post-hoc analysis of an NLST subgroup suggests that lung cancer risk and the outcomes of screening is not a simple linear relationship.[113,118] We suggest that while clinical risk models identify screening participants at greatest risk of lung cancer they do not adjust for the negative consequences of comorbid disease, in particular COPD, on the outcomes of screening.[113,118] The primary outcome of interest is both a reduction in lung cancer deaths and a reduction in all-cause mortality. This involves a complex relationship between airflow limitation, cancer biology, operability, and long-term survival.[102] To conclude, due to the attenuating benefits of screening on those at greatest risk, the "sweet spot" of screening may be found in those of intermediate risk.[118]

Chapter 7 Airflow Limitation and Survival after Surgery for Non-Small Cell Lung Cancer: Results from a Systematic Review and Lung Cancer Screening Trial

This paper has been published and is included with permission of the journal, 'Lung Cancer'.

RJ Hopkins, J Ko, GD Gamble, RP Young. 2019. 'Airflow limitation and survival after surgery for non-small cell lung cancer: results from a systematic review and lung cancer screening trial (NLST-ACRIN sub-study). *Lung Cancer* 2019; Volume 135: 80-87.'

7.1 Introduction

Lung cancer is one of the leading causes of cancer-related deaths in the world. In 2012, lung cancer was responsible for approximately 1.6 million deaths worldwide.[242] Lung cancer may be broadly divided into small cell lung cancer (SCLC) or non-small cell lung cancer (NSCLC), with many distinct histology subtypes in the latter.[125] In contrast to SCLC, where chemotherapy provides the mainstay of treatment with mostly palliative intent, surgery is routinely performed in early stage NSCLC with a curative intent.[125,243] Despite this, patients with lung cancer typically have a poor prognosis, reflected by an overall 5-year survival rate as low as 15%.[125] This low rate of lung cancer survival can be attributed in the main to the advanced stage at diagnosis, advanced age with associated comorbidities, and with certain lung cancer histological types, characterised by aggressive biology and rapid progression despite therapy (e.g. SCLC).

While smoking is widely recognised as the leading risk factor for the development of lung cancer, studies show between 50-70% of lung cancer patients also suffer from chronic obstructive pulmonary disease (COPD) characterised by airflow limitation.[34,244] It has been shown that the presence of COPD confers a significant and independent risk for the development of lung cancer in both non-screened[53] and screened lung cancer cases.[245] While smokers with underlying COPD are at greater risk of getting lung cancer, they may also be at greater risk of dying following surgery where overall 5-year survival rates have been reported to vary between 35-70%.[246,247] This may be due to death from diseases other than lung cancer such as cardiovascular disease, respiratory disease, or other cancers (i.e. competing risk of death).[20] Alternatively, it may be due to higher peri-operative risk and

poorer outcomes specific to lung cancer surgery in patients with COPD.[243] The question arises: "To what degree does pre-existing COPD lead to worse survival outcomes following surgery for NSCLC?"

To answer this question for unscreened lung cancers, we have undertaken a systematic review of published studies to examine the effect of pre-existing COPD on post-surgical survival in NSCLC. We have also looked at survival following surgery for NSCLC following screening in a subset of the National Lung Screening Trial (NLST), where baseline spirometry was routinely measured in screening centres affiliated with the American College of Radiology Imaging Network (ACRIN-NLST cohort, N=10,054 subjects).[93,219] The findings of these analyses have relevance to the differential benefits and harms of treating "operable" lung cancer in both the unscreened and screened clinical setting.

Hypothesis: The survival of subjects with screen-detected and non-screen detected lung cancers who undergo surgery for early stage cancer are not different in those with and without COPD.

Aim: To examine the relationship between underlying COPD and survival following surgery for early stage lung cancer in those with and without COPD.

Objective: To undertake a systematic review of lung cancer case series where pre-operative lung function was known and survival after surgery for early stage lung cancer. To compare these results with those from the NLST where baseline lung function was measured and outcomes following surgery were known.

7.2 Methods

For this study there were 2 reviewers and 1 adjudicator.

7.2.1 Search strategy and inclusion criteria for the systematic review

In order to correctly report the methods of this study, we used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[181] A systematic literature search was performed in Ovid MEDLINE, EMBASE and the Cochrane Database of Systematic Reviews as well as Google Scholar. The search terms were

group 1 "non-small cell lung cancer/lung cancer, adenocarcinoma",

group 2 "chronic obstructive pulmonary disease/COPD/chronic bronchitis/pulmonary emphysema",

group 3 "spirometry/bronchospirometry", lung function, airflow limitation,

group 4 "pneumonectomy/lobectomy/resection/removal" and

group 5 "mortality/survival".

Our searches were limited to a publication date from 1946 until January 2017 and no language restriction was applied in these initial searches.

The abstracts of all publications found by the literature search were then read and assigned a score from 1 to 5 which represented the number of keyword groupings that each abstract contained. Publications which included keywords from at least 3 groups and which mentioned all of "*non-small cell lung cancer/lung cancer/ adenocarcinoma*", "*pneumonectomy/lobectomy /resection/removal*" were then further analysed by reading the full-text article. Eligibility for further analysis was checked by two reviewers (RJH, JK) and a third reviewer (RPY) resolved any discrepancies. Any reasons for exclusion of studies based on their abstracts were recorded in a spread-sheet.

Any retrospective, prospective or randomised controlled study that met all of the following 6 checklist criteria were eligible for inclusion into the systematic review. The criteria were "*non-small cell lung cancer (stages I-II)*", "*spirometry criteria for COPD*", "*results stratified by COPD/lung function*", "*preoperative lung function*", "*lung surgery*" and "*mortality/survival (1-10 years)*". When we encountered the same patient population in separate publications, only the publication with the most recent data was included. Finally, for non-English publications, we (RJH) attempted to contact the authors for an English translation. If an English translation was unattainable, this study was excluded from the final analysis. Eligibility for the systematic review was assessed by two reviewers (RJH and JK) and a third reviewer (RY) resolved any discrepancies. The reason for the final exclusion of a study not to be included was recorded in a database which comprehensively recorded the full-text articles using a summary and checklist-based approach (summarised in Figure 7.1).

7.2.2 Data extraction for the systematic review

Data extraction was performed using a standardised record form. Two reviewers independently extracted information from each study included in the analysis (RJH, JK). Any disagreements or disparities that arose were resolved by a third reviewer (RPY). In terms of study characteristics, we extracted the lead author name, year of publication, country in which the study was performed, study design, follow-up period, number of patients (including the number of cases and controls), how the cases of COPD were defined, the overall survival rates for cases and controls and the proportion of patients with stage I-II NSCLC. We also included the odds ratio and 95% confidence interval for COPD and its effect on survival post-lung cancer surgery which was calculated on quantitative analysis. The key data extracted is shown in Table 7.1.

7.2.3 Statistical analysis for the systematic review

The data was analysed according to standard methods for a systematic review.[181] For published studies meeting the eligibility criteria, survival according to COPD status was summarised as an odds ratio with 95% confidence interval calculated for each study and meta-analysed using a random effects DerSimonian and Laird model. The results of this analysis was plotted using a forest plot with study heterogeneity reported as the I^2 statistic and publication bias formally assessed by inspecting funnel plots for asymmetry and Egger's Test. Any p-value less than 0.05 was considered statistically significant. Meta-analysis was performed using STATA version 10.1 (Metan, Statacorp, College Station, Texas, USA). Fisher's Exact test and confidence intervals for proportions were calculated using www.openepi.com (accessed 17/6/2019).

7.2.4 Outcomes from a subset of the NLST-ACRIN sub-study

We used data from a subset of the NLST, where subjects attending NLST screening sites affiliated to the American College of Radiology, Imaging Network (ACRIN) underwent routine pre-bronchodilator spirometry assessment.[93] In this sub-cohort of 10,054 high risk smokers randomised to CT or CXR screening, 395 lung cancers were identified after a mean follow-up of 6.4 years. Of these 328 were diagnosed within 5 years and 289 were NSCLC based on histology and 182 with NSCLC underwent surgical resection. All participants in this screening trial were followed for outcomes including death and cause of death.[93]

7.3 Results

7.3.1 Literature search and study characteristics in the systematic review

Using the aforementioned search strategy (Figure 7.1), we identified 288 papers including 89 duplicates that were removed from further analyses. The abstracts of these 199 papers were screened for the presence of at least 3 of the keyword groupings which excluded a further 116 papers. The full-text articles of the remaining 83 papers were acquired, scrutinised, and assessed for further inclusion by applying our 6 requirement criteria outlined above. A further 73 papers were excluded on the grounds of not meeting these inclusion requirement criteria, insufficient data for specified endpoints, duplication of data and non-English language. Ten papers published between 2001 and 2015 were included into the final systematic review[246-255] with follow-up ranging between 3 and 10 years (Table 7.1). This included four papers reporting 5-year survival. A flowchart summarising the process for including papers into the systematic review is shown in Figure 7.1.

Of these 10 studies, 7 were retrospective, 3 were prospective and there were no randomised controlled trials. The combined study population from all 10 studies was 6,899 operable non-small cell lung cancer cases. Due to variation in the spirometry criteria defining COPD, a sensitivity analysis was performed where studies using comparable criteria were grouped according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD criteria, FEV1/FVC<0.70, n=7 and other non-GOLD criteria, n=3).[256] Survival in only prospective studies (N=3) was also compared.

Figure 7.1 Flow chart of the selection process for studies included in the systematic review

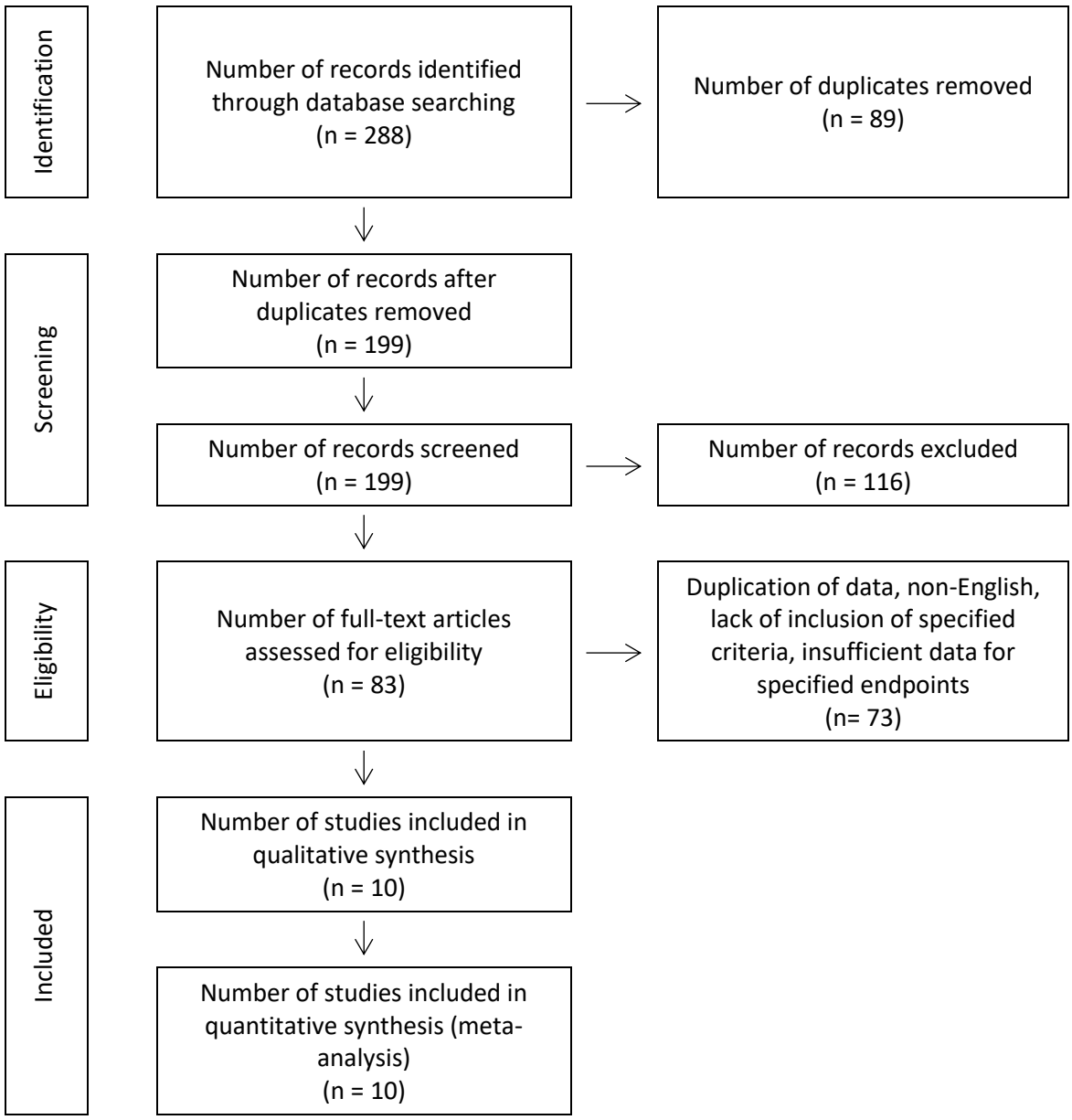


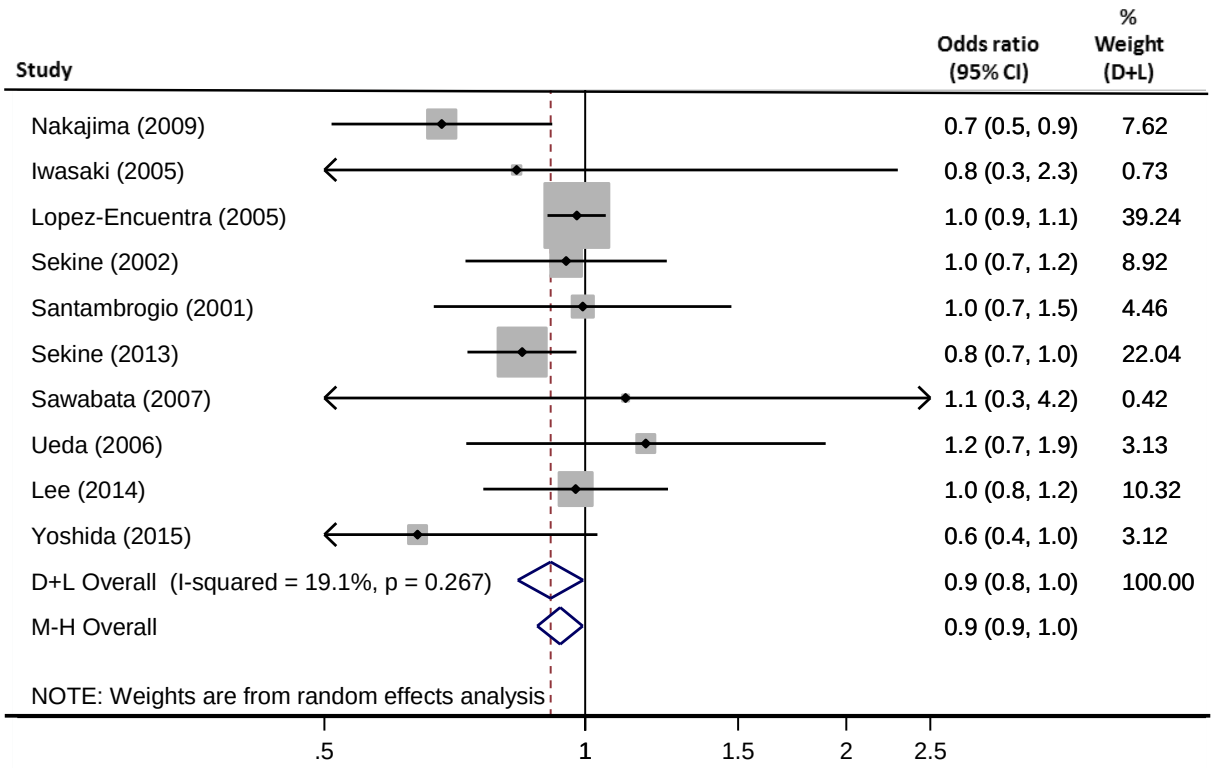
Table 7.1 Data extraction for each eligible study comparing cases (COPD) with controls

	Year	Country	Study Design	Follow-Up (Years)	Number of Patients (Cases/Controls)	Definition of Cases	5-year Overall Survival for Cases/Controls	Proportion of Patients with Stage I-II NSCLC	Odds Ratio (95% CI)	Comment
Iwasaki et al. [246]	2005	Japan	Retrospective	5	50 (12/38)	COPD defined by GOLD criteria*	70.0%/72.4%	72.0%	0.83 (0.30, 2.29)	No significant difference in 5-year survival for patients with and without COPD
Lopez-Encuentra et al. [248]	2005	Spain	Prospective	10	2994 (1370/1624)	COPD defined by GOLD criteria*	43.0%/45.0%	63.7%	0.98 (0.90, 1.06)	No significant difference in 10-year survival for patients with and without COPD
Sekine et al. [247]	2002	Japan	Retrospective	5	244 (78/166)	COPD defined by GOLD criteria*	36.2%/41.2%	68.9%	0.95 (0.73, 1.24)	No significant difference in 5-year survival for patients with and without COPD
Sekine et al. [249]	2013	Japan	Retrospective	10	1461 (363/1098)	COPD defined by GOLD criteria*	51.0%/61.5%	62.1%	0.85 (0.73, 0.98)	COPD patients have significantly poorer long-term survival
Lee et al. [250]	2014	Korea	Retrospective	7	221 (111/110)	COPD defined by GOLD criteria*	29.7%/32.7%	33.0%	0.97 (0.76, 1.24)	No significant difference in 7-year survival for patients with and without COPD
Santambrogio et al. [251]	2001	Italy	Prospective	3	88 (43/45)	COPD defined by FEV_1 %predicted between 40-79%	67%/67%	100%	0.99 (0.67, 1.47)	No significant difference in 3-year survival for patients with and without COPD
Ueda et al. [252]	2006	Japan	Prospective	5	100 (43/57)	COPD defined by FEV_1 %predicted <70%	58.4%/48.3%	>72.0%	1.17 (0.73, 1.89)	No significant difference in 5-year survival for patients with and without poor lung function (FEV_1 <70%)
Nakajima et al. [253]	2009	Japan	Retrospective	6	1461 (36/1425)	Severe COPD defined by FEV_1 %predicted 30-50% and controls defined by normal lung function or FEV_1 %predicted >50%	24.4%/58.6%	63.0%	0.68 (0.51,0.92)	COPD patients have poorer prognosis but are still candidates for surgical resection.
Sawabata et al. [254]	2007	Japan	Retrospective	5	103 (11/92)	Cases defined by FEV_1 %predicted 30-50% and controls defined by normal lung function or FEV_1 %predicted >50%	79.0%/79.0%	100%	1.11 (0.29, 4.24)	No significant difference in 5-year survival for patients with and without COPD
Yoshida et al. [255]	2015	Japan	Retrospective	8	243 (62/181)	COPD defined using GOLD criteria*	66.9%/80.7%	88.9%	0.64 (0.40, 1.03)	COPD patients have significantly poorer long-term survival

7.3.2 Analysis of overall survival in the systematic review

Overall survival (N=10 studies) was 950 (44%) in 2,144 people with COPD and 2,597 (55%) from 4,755 controls (unadjusted P value <0.001). The overall random effects odds ratio for overall survival following lung cancer surgery, comparing those with and without COPD was 0.91 (95% CI [0.84, 1.00]; heterogeneity testing p = 0.27; I^2 = 19.1%). Survival in those studies reporting 5-year data (N=4 studies) was 80 (50%) in 159 people with COPD and 189 (56%) from 338 controls (unadjusted P=0.25). The odds ratio for 5-year survival based on these 4 studies[246-247,252,254] was 0.99 (95% CI [0.79,1.24], P=.87). Survival in the prospective studies (N=3) was comparable (43% or 637/1,471 with COPD and 45% or 736/1,645 with no COPD, P=0.43). This suggests that the presence of COPD had a negligible impact on medium-term survival after lung cancer surgery (Figure 7.2). Although under-powered, a similar finding was made in the sensitivity analyses where no significant difference in overall mortality could be found. There was no evidence of publication bias (Eggers tests p = 0.19).

Figure 7.2 Forest Plot of odds ratio of all cause death within 3-10 years of lung cancer surgery for unscreened NSCLC for those with and without spirometry-based COPD

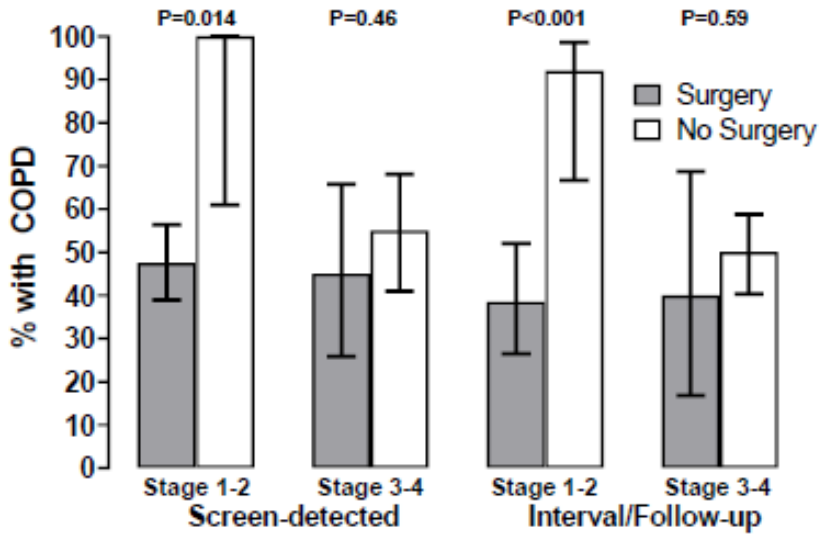


Legend: Random (D+L=DerSimonian-Laird) and fixed (M-H=Mantel-Haenszel) effect models are shown. Box size is proportional to the random effects weight and results below 1 favour reduced odds of all cause death in those without spirometry-based COPD.

7.3.3 Survival in the NLST-ACRIN sub-study

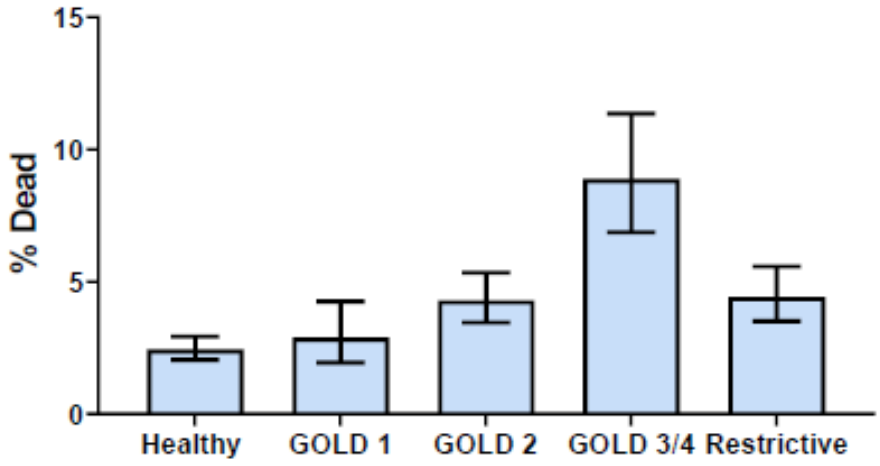
The outcomes from the NLST-ACRIN sub-study censored to 5 years are summarised in Table 7.2. Of the 328 lung cancer cases identified during the screening study, 289 were NSCLC with 153 (53%) being stages 1 and 2. The overall 5-year post-surgery survival in those with and without COPD was not significantly different (84 and 81% respectively, P=0.65). These results were not affected by screening arm nor was there a difference when the analysis was limited to stage 1-2 NSCLC cases (Table 7.2). The prevalence of COPD in early stage cancers that did not undergo surgery was 92-100%, regardless of detection pathway, and was significantly higher than the expected 40-50% who underwent surgery (Figure 7.3). The overall mortality for those not diagnosed with lung cancer was linearly related to COPD severity, where mortality rates were 3-4 times higher in those with GOLD 3-4 grade COPD compared to those with normal lung function (Figure 7.4). These findings were also independent of screening arm.

Figure 7.3 Prevalence of COPD (%) in early and late stage lung cancer detected during screening or during Interval/Follow-up, stratified by surgery in the NLST-ACRIN sub-study



Legend: Bars represent % spirometry-defined COPD (± 95% Confidence Interval). Fisher's Exact P values are shown for pairwise comparisons.

Figure 7.4 Overall % mortality for non-lung cancer-related deaths according to severity of airflow limitation in the NLST-ACRIN screening sub-cohort, censored at 5 years follow-up



Legend: Bars represent % dead after 5 years (± 95% Confidence Interval).

Table 7.2 Summary of outcomes (5-year survival) for operable lung cancer in the NLST-ACRIN sub-study sub-grouped by COPD status and clinical stage

Lung Cancer Characteristics	CT arm		CXR arm		Total		P value
Total Lung cancer cases	180		148		328		
SCLC	19	(11%)	20	(14%)	39	(12%)	
NSCLC	161	(89%)	128	(86%)	289	(88%)	0.42
NSCLC undergoing surgery	116	(64%)	66	(45%)	182	(55%)	
Survival in NSCLC by COPD status							
COPD	51		34		85		
COPD Survival to 5 years	44	(86%)	27	(79%)	71	(84%)	0.81
Non-COPD	62		32		94		
Non-COPD Survival to 5 years	55	(89%)	21	(66%)	76	(81%)[#]	0.38
Unknown status	3		0		3		
Stage 1-2 NSCLC undergoing surgery	101	(56%)	52	(35%)	153	(47%)	
Survival in stage 1-2 NSCLC by COPD status							
COPD	46		31		77		
COPD Survival to 5 years	44	(96%)	27	(87%)	71	(92%)	0.78
Non-COPD	62		32		94		
Non-COPD Survival to 5 years	55	(89%)	21	(63%	76	(81%)	0.38
Unknown status	3		0		3		

#The overall 5-year post-surgery survival in those with and without COPD was not significantly different (84 and 81% respectively, P=0.65).

7.4 Discussion

From our systematic review of the literature, we found no statistically significant decrease in overall and 5-year survival in early stage lung cancer cases following surgery for patients with COPD compared to those without COPD. This is reflected by a random effect's odds ratio for COPD post-lung cancer surgery of 0.91 (N=10, 95% CI [0.84, 1.00) for overall survival and 0.99 (N=4, 95% CI [0.79, 1.24]) for 5-year survival. Therefore, COPD has a negligible impact on medium-term survival post-lung cancer surgery for early stage disease. There was little heterogeneity in our results (N=10), with an I^2 result of only 19.1%, and minimal publication bias (Egger's test was negative, P = 0.19). These findings were confirmed in our analyses of all-cause mortality in early stage NSCLC in the NLST-ACRIN sub-study, where the mortality rates censored to 5 years were comparable in those with and without baseline COPD. These findings suggest that for smokers with predominantly mild-to-moderate COPD, outcomes following lung cancer surgery are comparable to those for non-COPD.

Of the 10 studies included in this review, 8 studies showed little difference in mortality while 2 studies (Nakajima and Yoshida)[253,255] observed that COPD was associated with a 30%-40% greater mortality (see Figure 7.2). For the sake of thoroughness of our review, we decided to further analyse these studies to determine if there were any differences in the methodology or patient group in these studies that may have led to these contradictory results. Nakajima et al. have an odds ratio of 0.68 (95% CI [0.51, 0.92]), however they split their patients into a severe COPD group and a control group containing people with normal lung function, mild COPD, and moderate COPD.[253] This is not comparable to the other studies in this review which have split patients simply into those with COPD and those without COPD. In 2013, Sekine et al. reported an odds ratio of 0.84 (95% CI [0.73, 0.97]), however this study was primarily concerned with the severity of COPD.[249] In our analysis, we calculated an approximate overall survival rate for COPD of 51.0% in the Sekine study[249] by calculating actual survival numbers combining patients with mild, moderate, and severe COPD. In the multivariate analysis of long-term overall mortality carried out by Sekine et al., only severe COPD was associated with poorer long-term outcome post-lung cancer surgery.[249] This finding accords with other studies suggesting that COPD (especially GOLD 3-4) is associated with more aggressive forms of lung cancer,[219,257] shorter volume doubling time[129,258] and greater mortality from non-lung cancer related deaths.[259] This latter finding was strongly

born-out by our findings in the NLST-ACRIN sub-study showing that while increasing severity of COPD is associated with an increased risk of lung cancer,[219] the effect on other causes of deaths is even greater.[80] This means that while the presence of airflow limitation increases a smoker's risk of lung cancer in a simple linear relationship,[219] it also increases the likelihood that smokers will die prematurely of other smoking related complications (Figure 7.4).[102,259] In the context of CT screening for lung cancer, this raises an important consideration with regards to the benefits of screening in this very high risk co-morbid group.[97,103,260] It also suggests that spirometry should be routinely performed in high risk smokers undergoing screening so that the relative benefits versus harms of lung cancer screening can be correctly assessed.[217]

7.4.1 Study strengths

We believe there are several strengths to this study. Firstly, we have followed the recognised methodology for undertaking a systematic review in accordance with the PRISMA statement.[181] In our screening process, we used a strict 5 group keyword search of recognised databases such as Ovid MEDLINE, EMBASE and Cochrane Database of Systematic Reviews. We adopted extensive and repeated searches to attain all possible abstracts relevant to our topic. For those not available in English language, we contacted the authors for an English translation of the papers. We then applied a stringent set of 6 checklist "requirement" criteria to include only those studies with appropriate parameters and correct follow-up data that could be compared. Secondly, our systematic review was both multi-centre and multi-national; including studies from several countries in both Europe, USA, and Asia. Yet, in spite of this diverse study pool, our systematic review had minimal heterogeneity. Third, an important strength of our study was the requirement to define COPD on spirometry terms.[256] This is relevant because "self-reported" COPD is a very unreliable way to assign COPD status, particularly in a study of lung cancer cases where between 50-70% of cases have pre-existing COPD yet less than one half of these individuals will have been "diagnosed".[48] In our NLST-ACRIN study we show that "self-reported" COPD was strongly linked to increasing severity of airflow limitation, best reflecting severe or very severe COPD (GOLD 3-4).[219] In the study by Zhai et al.[261] which relied on self-reported COPD, the authors concluded the overall mortality after surgery for early stage lung cancer was much higher in those with COPD but no spirometry was performed to confirm the diagnosis.

We suggest that this increased mortality is primarily based on severe COPD. Lastly, although for completeness our search dates back to 1946, the 10 studies in our analysis date from 2001 to 2015, and thus represent recent approaches to the surgical management of lung cancer in the non-screening setting.

While inclusion into our systematic review was regulated by our set of 6 rigid checklist "requirement" criteria, we identified several studies that highlight important points regarding the effect of COPD on outcomes other than death.[261-270] An alternative outcome of interest is how lung cancer surgery affects the lung function of patients with and without COPD. Several studies have shown that COPD patients actually have a significantly smaller loss in FEV_1 post-surgery than patients without COPD and, in some cases, COPD patients may actually have improved lung function post-surgery.[261,262] This is important because quality of life after surgery is also a crucial factor to take into consideration. In addition, the potential for postoperative complications should influence the decision for lung cancer surgery as these complications negatively impact the patient's quality of life.[264-268] While some studies have shown that patients with COPD have a comparable postoperative complication rate to patients without COPD, other studies show the presence of COPD confers a significant risk for pulmonary complications such as prolonged air leak, pneumothorax, pneumonia and atelectasis.[264-266] Perioperative mortality is another outcome of interest that has been widely reported throughout the literature, however the results vary. Many studies report that perioperative mortality is not significantly different between patients with and without COPD with some reporting no perioperative mortality at all.[246,249,251,261,267] Other studies, report that perioperative mortality is significantly increased in patients with COPD compared to those without COPD, making COPD an independent risk factor for perioperative mortality.[250,251,269,270] These disparities in results indicate that a systematic review of the evidence, examining the effect of COPD on postoperative mortality in early stage lung cancer is indeed required.

Given the recent interest in CT screening for lung cancer,[97,260] there is a growing interest in the outcomes of lung cancer surgery especially with regards to "competing cause of death". In his recent review on CT screening, Mazzone expands the current concept of "over-diagnosis" to encompass the screening of high risk smokers who die of other complications of smoking.[260] This "overtreatment" is highly relevant to those with co-existing COPD as

airflow limitation is associated with premature death from cardiovascular disease, respiratory disease and other cancers.[38,39,124,271] Mortality from competing risk of death therefore undermines the value of CT screening in those at risk of death from other diseases. For this reason, we re-examined the "COPD effect" on outcome after surgery in the NLST-ACRIN study where spirometry had been routinely performed at baseline. We found that although outcomes for those with COPD were comparable to those with no COPD, severity of airflow limitation was directly related to death from other causes. This finding means those at greatest risk of lung cancer (GOLD 3-4) also have a greater risk of dying from non-lung cancer smoking complications.[272]

7.4.2 Study limitations

This systematic review has several limitations. First, all of the studies included in the systematic review were observational studies as we found no randomised controlled studies. We accept that by analysing observational studies, the strength of the evidence is much weaker than if data from randomised controlled studies were analysed. However, the question we pose cannot be ethically answered by a randomised controlled trial. Second, the classification of COPD varied among the studies considered in this systematic review. There currently exists a widely-accepted Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification system for COPD established in 2001.[256] Although GOLD classification is the current gold standard for diagnosing COPD and grading its severity, prior to 2004 it was not consistently used. Third, most of our studies were hospital-based case series studies from Japan, which have very small study sizes; in fact, all but 3 have a study population of less than 1000. We understand that smaller study populations lead to a reduction in statistical power which therefore decreases the chance of detecting a true effect. While the studies themselves had limited power, in doing a systematic review we aimed to increase this power by pooling the collective populations to create a sufficiently large study size. Fourth, the type of surgical treatment was not standardised across the studies included, where treatment ranged from pneumonectomies to lobectomies. The type of procedure that a patient underwent was usually chosen at the operating surgeon's discretion therefore there was poor consistency across study populations and even within cohorts. This, along with the small size of individual studies prevented us from looking independently at the association between the procedure type and long-term survival. Fifth, there are several

instances of intrinsic bias within our study. The studies included in our systematic review look at a non-screening population, from many centres where the role of surgery in the treatment of "early stage" lung cancer is varied. This reduces the generalisability of our findings and leads to poor external validation. Furthermore, our main aim was to analyse patients with stage I-II NSCLC, however most of the studies analysed included a wide range of patients suffering from stages I-IV NSCLC (Table 7.1). For this reason, we examined the "COPD effect" on a sub-group of the NLST where post-surgical outcomes were limited to early stage lung cancer and this verified the findings of our systematic review. Importantly, we show in the NLST study, that the prevalence of COPD in un-operated early stage lung cancer was 100% and 92% in screen and non-screen detected lung cancer respectively, compared to the expected 40-50% prevalence seen across all lung cancers (Figure 7.3). While the basis of this observation remains unclear, this finding provides further understanding as to why those with COPD had only half the benefit from screening as those with no COPD.[80]

7.4.3 Future implications

Our finding that COPD patients have similar medium-term survival post-lung cancer surgery to those without COPD has several future implications. First, this evidence supports the premise that the presence of COPD should not be a factor in any exclusion criteria for lung cancer surgery or screening.[273] However, while the presence of COPD alone is not enough to create a significant survival disparity, the presence of severe COPD is another question entirely. Both Nakajima et al. and Sekine et al. have shown that GOLD III COPD (FEV_1/FVC <0.7 and FEV_1 %predicted 30-50%) is associated with significantly reduced medium-term survival compared with non-COPD patients, post-lung cancer surgery.[249,253] This would support the use of routine pre-operative lung function testing in order to minimise the number of GOLD III/IV COPD patients undergoing lung cancer surgery.

7.5 Conclusion

The results of our systematic review show that there is no significant difference in medium-term survival between NSCLC patients with and without mild-to-moderate COPD undergoing surgery for early stage disease.

Chapter 8 Are New Zealand Māori More Susceptible to Smoking Related Lung Cancer? A Comparative Case-case Study

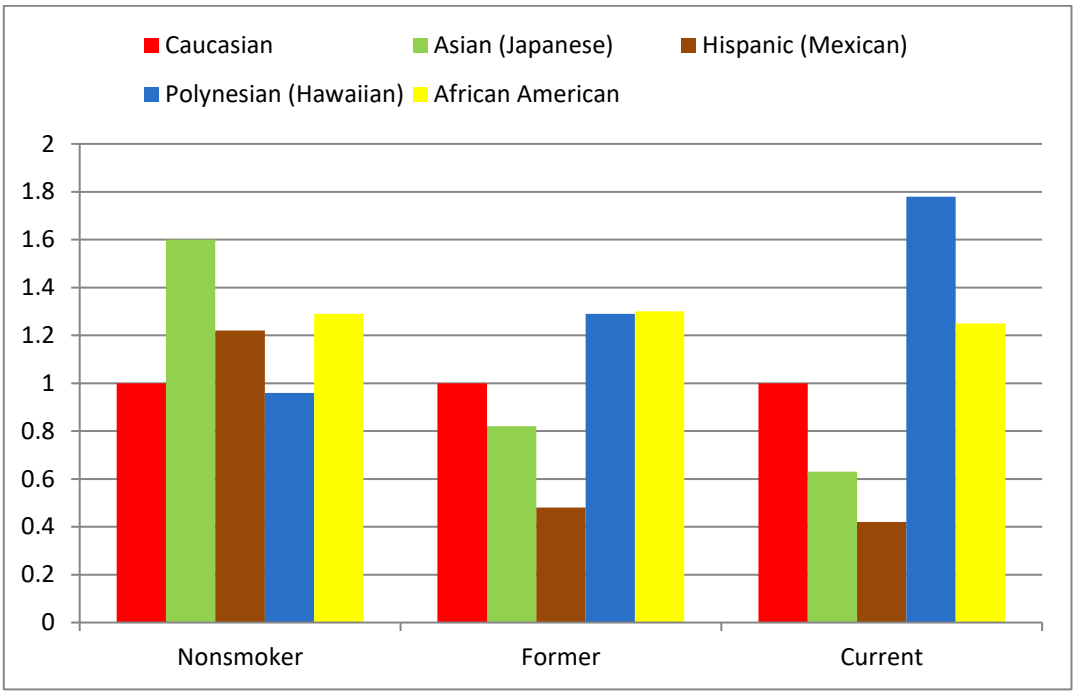
This paper has been published and is included with permission of the official journal EC Pulmonology and Respiratory Medicine

RJ Hopkins, Kendall C, GD Gamble, RP Young. Are New Zealand Maori More susceptible to smoking related lung cancer: - A comparative case-case study. *EC Pulmonology and Respiratory Medicine* 2019; 8.1:72-91.

8.1 Introduction

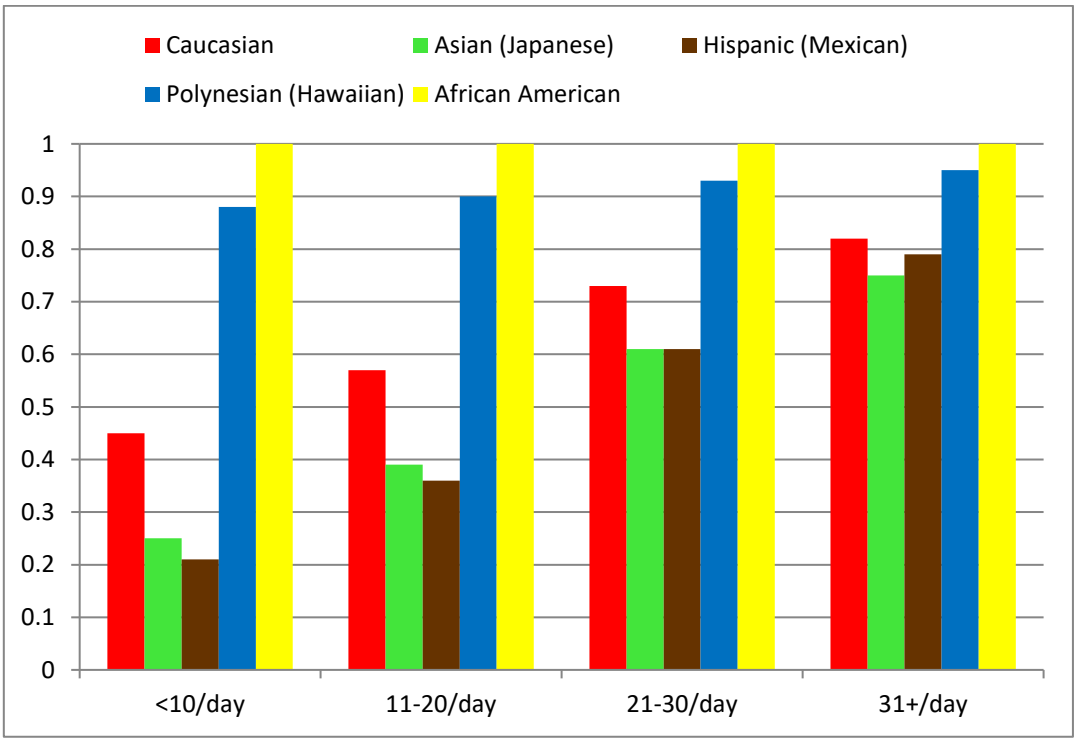
Lung cancer is the leading cause of cancer deaths among men and women, attributed in the main, to smoking.[274] Epidemiological studies show that with increasing smoking exposure, the risk of lung cancer increases in a linear or curvilinear dose-response relationship.[83,198] Other factors known to contribute to an increased risk of lung cancer include; increasing age, the presence of chronic obstructive pulmonary disease (COPD), exposure to other aero-pollutants (asbestos, other occupational dusts, air pollution and radon), maternal smoking, lower socio-economic status (SES), poor diet and ethnicity.[34,83,145-147,198] In a large case-series, comparing five different ethnic groups in the United States (US), it was found that, relative to Caucasians, age adjusted lung cancer incidence in ever smokers was greatest for those of Polynesian (Native Hawaiian) or African American ancestry, while lowest in those of Hispanic (primarily Mexican) and Asian ancestry,[146,147] (Figure 8.1).

Figure 8.1 Risk of lung cancer according to smoking status and ethnicity in the US (Caucasian referent) (data taken from ref.147)[148]



These findings could not be explained by differences in diet, smoking exposure, and high school education, where educational level provides a surrogate of SES.[275] A striking finding of the US study was that while the expected dose-response relationship between increasing cigarette exposure dose (up to 30 cigarettes/day) and lung cancer risk was shown for most ethnicities, it was absent in Hawaiians and African-Americans,[148] (Figure 8.2). In a large prospective study of lung cancer, after adjustment for important risk variables, Hawaiian ancestry conferred the greatest risk of lung cancer relative to Caucasians.[86] This suggests that although smoking generates ethnicity-based differences in lung cancer incidence, these "disparities" appear to be due to factors other than cigarette exposure dose.

Figure 8.2 Risk of lung cancer according to smoking intensity (cigarettes/day) and ethnicity in the US (African-American referent), (data taken from ref.146)



Māori are the indigenous Polynesian people of New Zealand (NZ) and have about 3-4 times higher age-adjusted lung cancer incidence compared to Caucasians.[158,160,276,277] Māori, like Hawaiians, originate from the Pacific Islands and share their genetic ancestry with Polynesians.[278] This Polynesian ancestry is genetically quite distinct from South-East Asians through admixture between Aboriginal Taiwanese and Micronesians.[277] Although smoking rates are two-fold higher in Māori compared to NZ Caucasians, this does not explain the 3-4 fold greater lung cancer incidence.[279] In New Zealand large disparities for lung cancer mortality specifically,[158,160,277] and smoking-related mortality overall,[159,280] have been reported between Māori and NZ Caucasians. After a detailed examination of this disparity in smoking-related mortality, investigators concluded that only about 50% of the difference could be explained by the higher smoking rates and lower SES found in Māori.[159] This raises the possibility that like Hawaiians, ethnicity-specific factors among Māori confer greater susceptibility to lung cancer.[276,281] Our group have previously shown that airflow limitation (or COPD) is associated with an increased risk of lung cancer[34] and propose that this might be relevant for the development of lung cancer in Māori, where COPD rates are two-fold higher.[158] The current study compares the clinical characteristics of Māori and NZ Caucasian

diagnosed with lung cancer. Preliminary results have been previously reported in abstract form.[157]

Hypothesis: The characteristics of lung cancer cases in Maori are similar to those in Caucasians in New Zealand.

Aim: To compare the clinical and lung cancer characteristics between Maori and Caucasian New Zealanders.

Objective: Using a case-case study design to compare the clinical and lung cancer characteristics between Maori and Caucasian New Zealanders.

8.2 Methods

This study involved a retrospective review of case records from the same geographical region of the greater Auckland metropolitan area which included a referral base that encompassed the upper half of the North Island of New Zealand (approximately 40% of the population, 1.5 million).

Lung cancer cases were identified using local district health board databases where lung cancer was reported as the 1st, 2nd, or 3rd diagnosis according to the International Classification of Diseases coding system (ICD 10 C34). "Māori ancestry" was assigned according to self-reported ethnic identification in hospital records. Four hundred and seventy-two Māori patients diagnosed with lung cancer between January 1st, 2004 and December 31st 2014 were retrospectively identified. Four hundred and fifteen NZ Caucasian patients diagnosed with lung cancer were prospectively identified as part of a separate epidemiological study,[34] between January 1st 2004 and December 31st 2008. Data collected on basic demographic characteristics included; age at diagnosis, gender, ethnicity, smoking history, histological subtype, stage at diagnosis, months of survival from diagnosis, spirometry and date of diagnostic CT or diagnostic histology. Never smokers with lung cancer were included in this study.

Date of diagnosis of lung cancer was defined by date of the most definitive investigation, with a descending preference given for histology result or radiological imaging. Staging was defined according to lung cancer staging I-IV in accordance with current guidelines.[183]

Histological subtype was identified from pathology reports and classified as small cell, adenocarcinoma, squamous cell, non-small cell (for cancers with no more precise classification), or other which encompassed the remainder of cancers but excluded carcinoid tumours, secondary tumours, or benign tumours of the lung. The presence of chronic obstructive pulmonary disease (COPD) was defined according to sighted spirometry reports or lung function results, performed prior to or during the early work up of suspected lung cancer, using Global Obstructive Lung Disease (GOLD) criteria.[34] Smoking histories were exclusively based on recorded pack years of smoking exposure. If not otherwise specified, the average number of cigarettes smoked per day for Māori was calculated using the reported pack years smoked, age at diagnosis (current smokers) or age quit smoking (ex-smokers) and the previously published mean age of smoking uptake in Māori (16 years).[282]

In two sensitivity analyses, we conducted an age-gender matched comparison (N=331 Māori cases) and we restricted Māori lung cancer cases to those with Māori surnames (N=249 Māori cases), to potentially enrich them for Māori ancestry. The demographic characteristics were also compared after stratification according to gender, smoking status, and presence of COPD, where the latter has been associated with more aggressive lung cancer histology.[93] A stratification approach was preferred over more complex multivariate analysis as most of the variables of interest were discrete and because linear relationships could not be assumed. However, for those diagnosed between 2004-2007, we compared survival and factors underlying survival in a Cox-proportional model to examine the relative importance of ethnicity and airflow limitation in the development of lung cancer.

8.2.1 Statistical analysis

Demographic variables were compared by unpaired t-test for normally distributed continuous variables and chi-squared test for discrete variables with Mid-P Exact test on a two-tailed analysis (www.openepi.com access 18/12/2016). Survival was compared using Cox-Proportional Hazards regression model, the assumption of proportionality was tested by including log time dependent variables in the model SAS statistical package (v 9.4, Cary, NC, USA was used. All tests were two tailed and $P < 0.05$ was considered significant.

8.3 Results

8.3.1 Demographics

Demographic characteristics are presented in Table 8.1. Māori lung cancer cases were younger with a mean age at diagnosis of 61 years old, compared to 67 years old in NZ Caucasians. In comparison with Caucasians, Māori lung cancer cases were ≈2 fold more likely to be current smokers (67% vs 36%, P<0.05) and less likely to be a never smokers (2% vs 7% respectively, P<0.05), although Māori and Caucasians had similar pack years of smoking (39 vs 41 pack years respectively), irrespective of gender. The estimated average number of cigarettes consumed per day for Māori (17/day)[282] was comparable to that recorded for Caucasian lung cancer cases (16/day). The distribution of age at lung cancer diagnosis for Māori and NZ Caucasian is presented in Figure 8.3, and shows that nearly one half (45%) of cases in Māori were diagnosed at 60 years of age or less, compared to 25% in NZ Caucasian cases (P<0.05). Distribution according to smoking exposure by pack years is shown in Figure 8.4, and suggests Māori with lung cancer were generally lighter smokers than Caucasians.

Figure 8.3 Distribution of age at diagnosis of lung cancer cases according to ethnicity

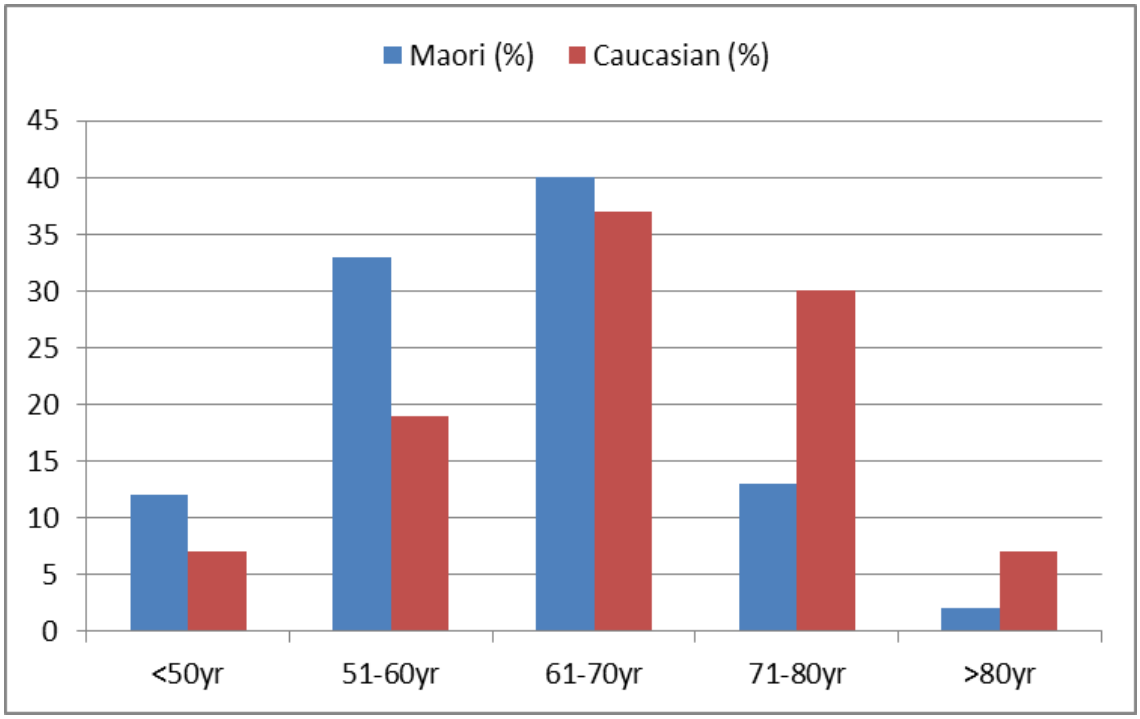
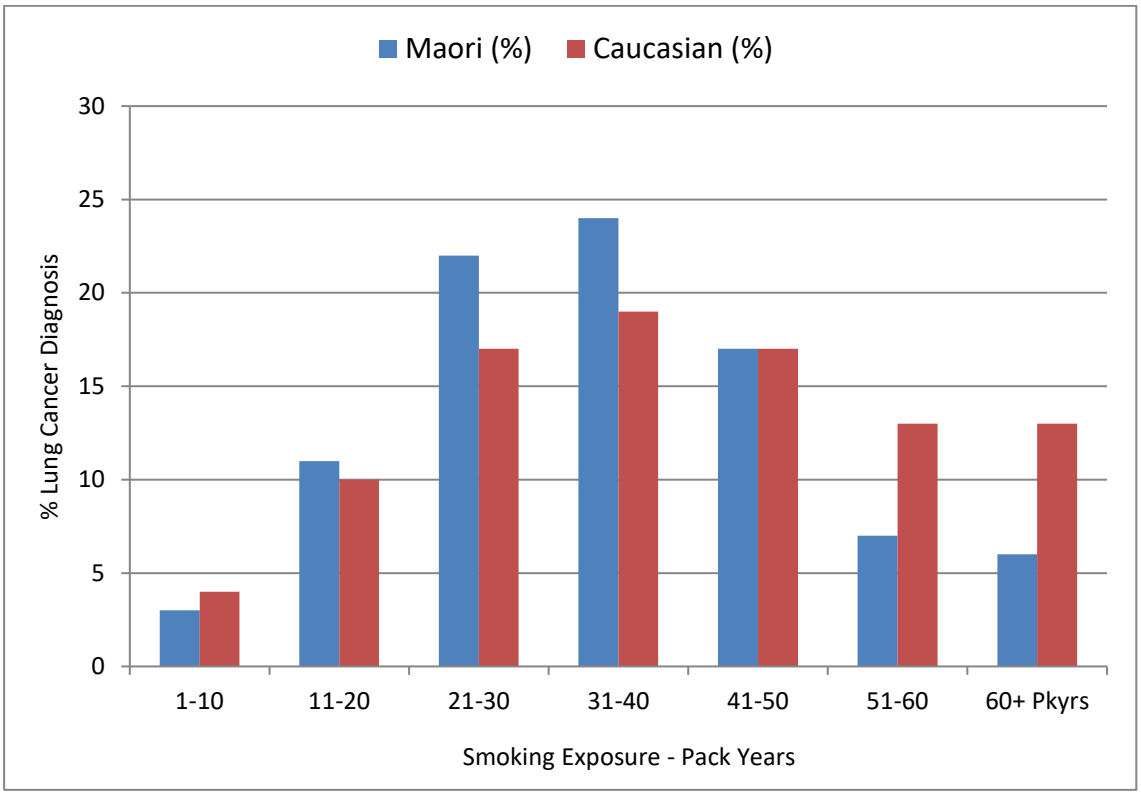


Figure 8.4 Distribution of pack year smoking exposure at diagnosis of lung cancer cases according to ethnicity



8.3.2 Lung function

Spirometry results performed during usual clinical work-up were available for 69% of Māori lung cancer cases and 62% of Caucasian cases (Table 8.1), reflecting that spirometry is not routinely done during the work-up and management of all lung cancer cases. The mean ratio of FEV_1/FVC was identical at 0.64 but the mean per cent predicted FEV_1 was significantly lower in Māori compared to Caucasian (64% vs 71% respectively, $P<0.05$). Across increasing pack year exposures, the distribution of lung function test results, and COPD prevalence, showed several differences when comparing Caucasian results with Māori, where the same normal reference values were used as previously suggested in studies in Polynesian.[283] For FEV_1% predicted, at every level of accumulated smoking exposure (pack years), I found the FEV_1% predicted was lower in Māori compared to Caucasian lung cancer cases, (Figure 8.5). This was significant at low exposures (1-10 and 11-20 pack years). For the FEV_1/FVC ratio, Māori had a significantly lower ratio than Caucasians (67% vs 75%) for those who had smoked 1-10 pack years, (Figure 8.6). When COPD prevalence is compared, I note for smokers of 1-10 pack years, Māori have two fold greater COPD prevalence than Caucasians

(64% vs 30%, P<0.05) and the linear "dose-response" relationship observed in Caucasians is lost in Māori (Figure 8.7 and Figure 8.8), shows a comparable distribution for GOLD severity.

Figure 8.5 Distribution of lung function mean FEV_1%predicted according to smoking exposure (pack years) and ethnicity

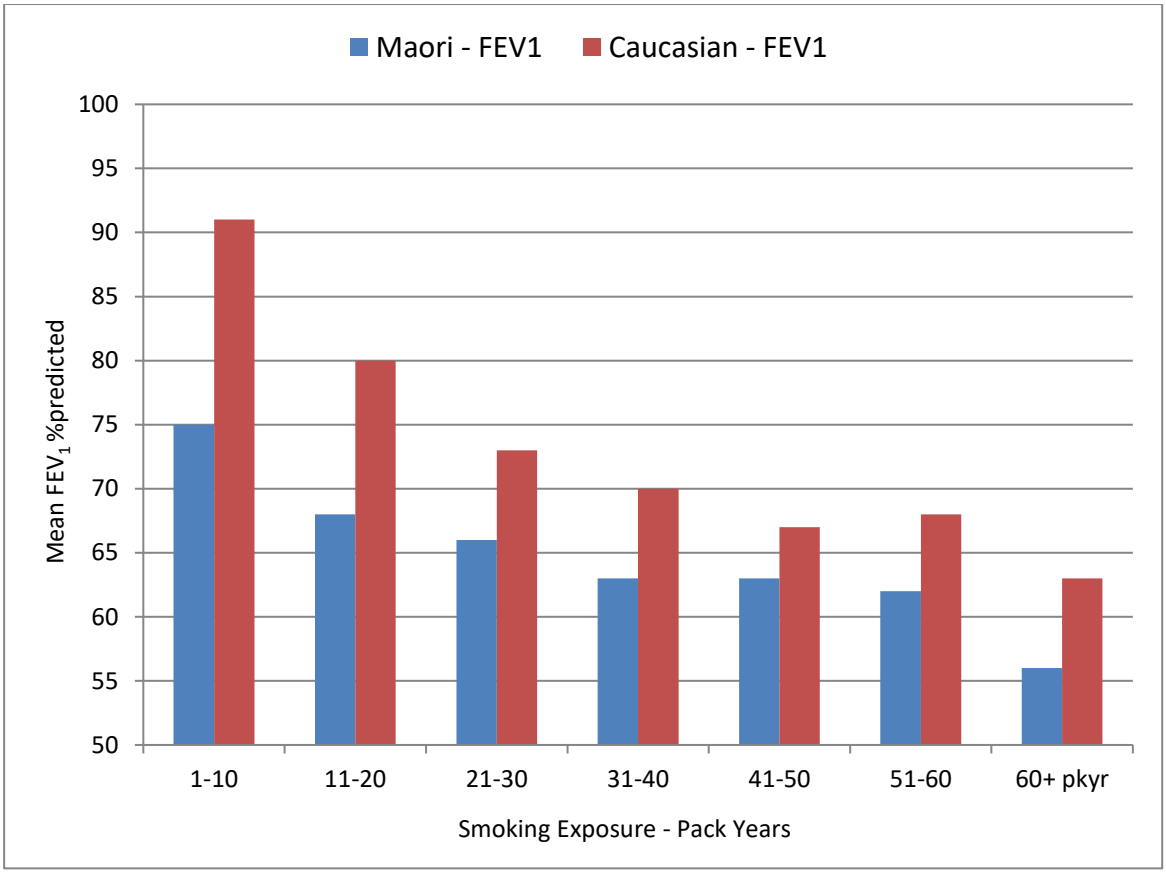


Figure 8.6 Distribution of lung function mean FEV_1/FVC ratio according to smoking exposure (pack years) and ethnicity

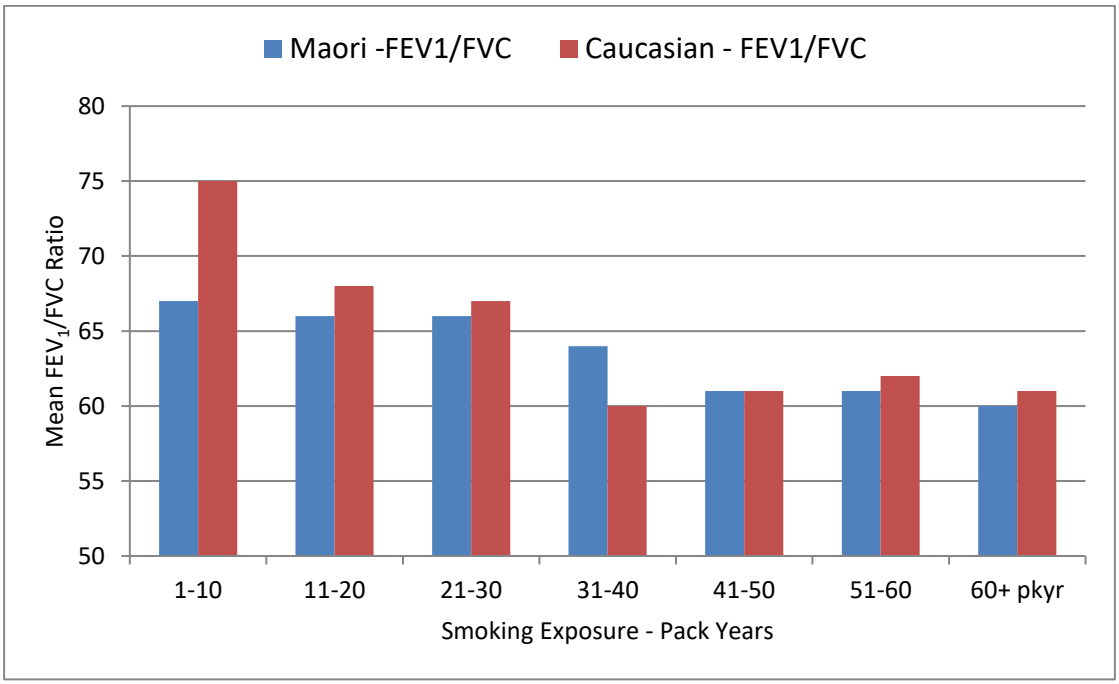


Figure 8.7 Distribution of lung function and prevalence of COPD, according to smoking exposure (pack years) and ethnicity

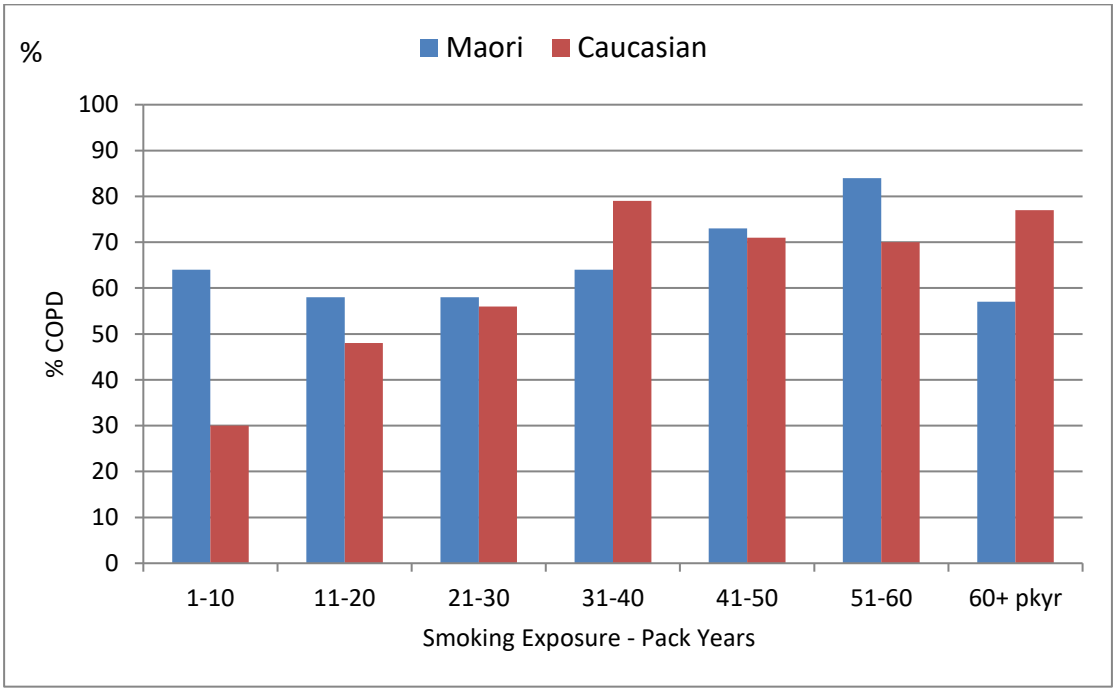
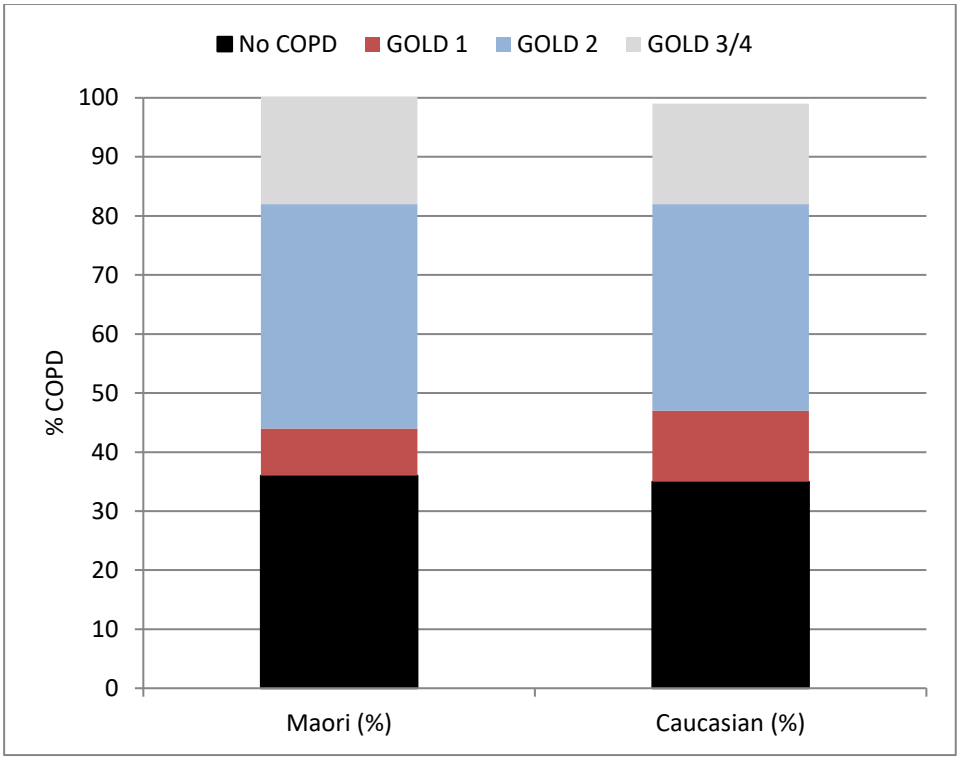


Figure 8.8 Distribution of COPD by GOLD severity according to ethnicity



8.3.3 Histology, staging and survival

When the prevalence of each histological group was compared, Māori were found to have less adenocarcinomas than was found in Caucasian lung cancer cases (32% vs 42%) and marginally more of the other more aggressive histological subgroups (62% vs 52%, P=0.02) (Table 8.1 and Table 8.2). Table 8.2This relative under-representation of adenocarcinoma histology in Māori compared to Caucasian was also found in the age-gender matched comparison (P=0.07), (Table 8.2) and the Māori-enriched comparison where this under-representation was greatest (P=0.002), (Table 8.3).

Figure 8.9 Distribution of lung cancer histology according to ethnicity

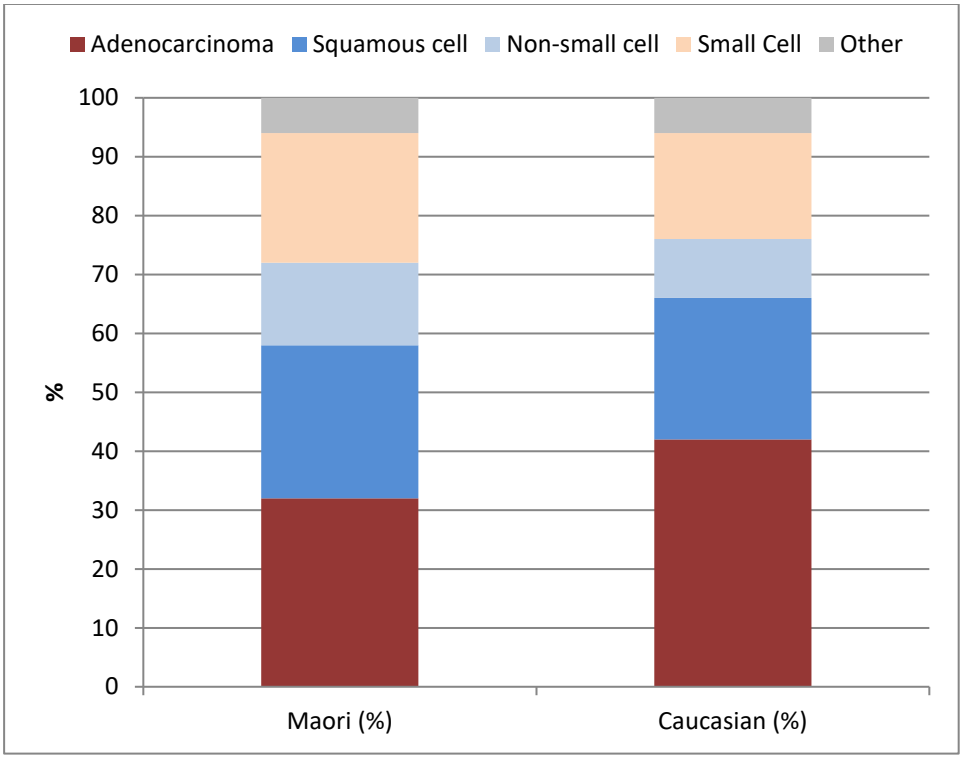
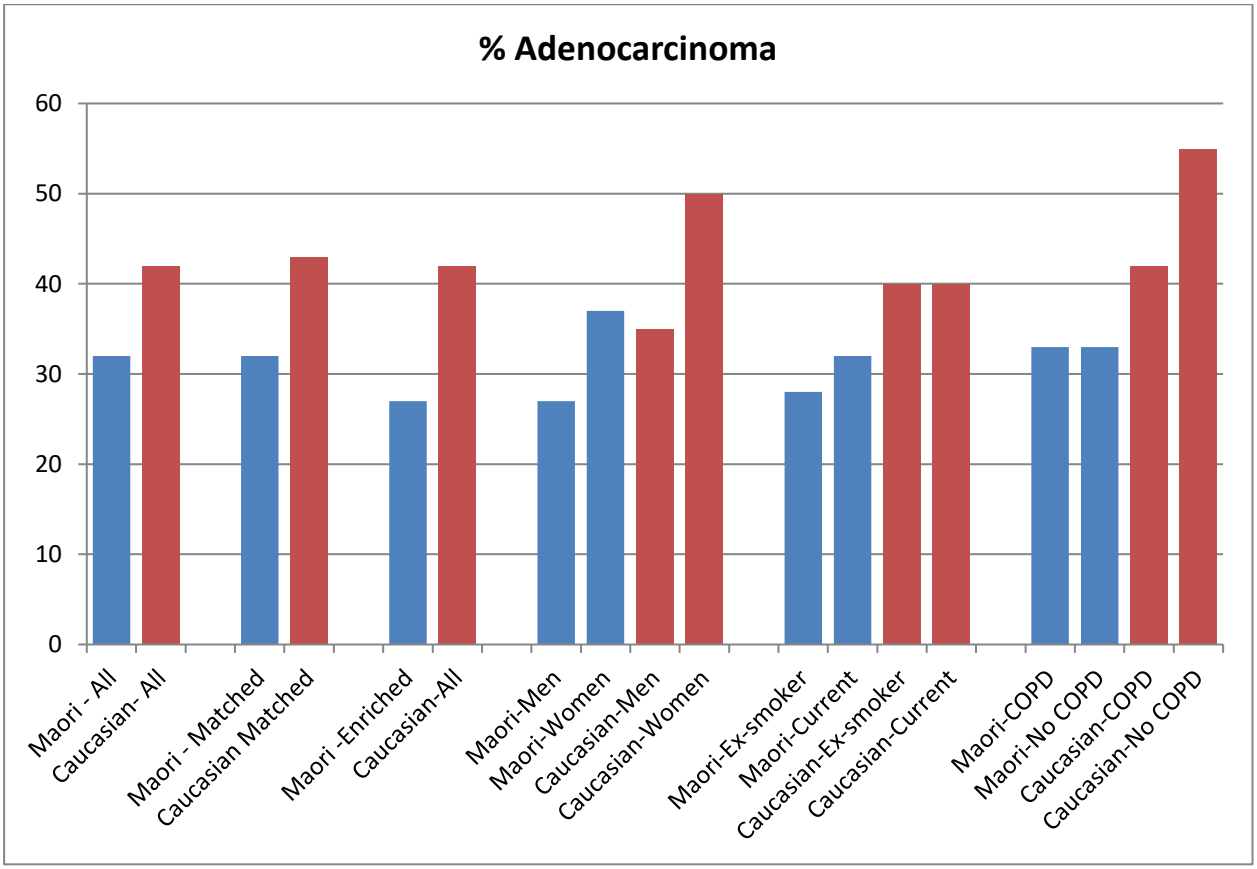


Figure 8.10 Proportion of lung cancer cases having Adenocarcinoma histology subtype according to ethnicity, before and after stratification by gender, smoking status, and presence of COPD



This under-representation of adenocarcinoma in Māori persisted after stratification for gender (female>male, P=0.00002), smoking status (ex-smoker>current, P=0.01) and presence of COPD (absence>presence, P=0.01), (Figure 8.10, Table 8.4, Table 8.5 and Table 8.6). When the distribution of lung cancer stages were compared, we found only a marginally lower proportion of early stage (1-2) disease in Māori compared to Caucasian (28% vs 33% respectively, P>0.05, Table 8.1). This difference was significant in those with COPD (37% vs 51% respectively, P<0.05, Table 8.6). Using Cox-proportional survival curves, when the 1 year, 2 year, 5 year and 10 year survivals were compared, I found that Māori (N=81) had 1.4 fold lower survival than Caucasians (N=312), (Figure 8.11), and that stage, age, gender and histology were relevant to mortality but not COPD or smoking status (see Discussion and Figure 8.11).

Figure 8.11 Cox-proportional hazard ratio analyses for lung cancer mortality in Māori and Caucasian lung cancer cases series diagnosed between 2004 and 2007

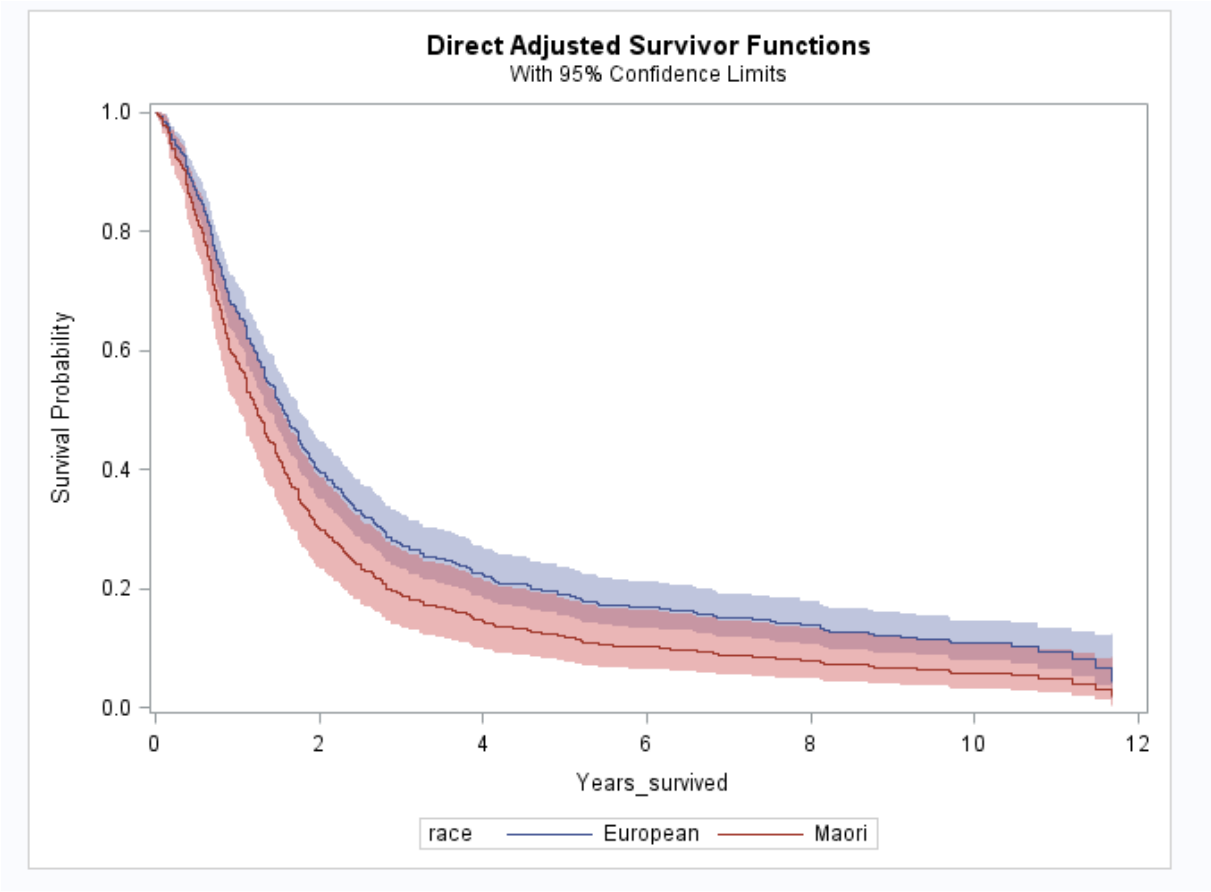


Table 8.1 Demographic variables in the Māori and Caucasian lung cancer cases series (Unmatched comparison)

Demographic Variable	Māori (N=472)	Caucasian (N=415)	P value
Mean age at diagnosis (SD)	61 yrs (9)	67 yrs (10)	<0.001
Range	32-87 yrs	40-92 yrs	
Male (%)	205 (43%)	213 (51%)	0.019
Smoking History			
Smoking at diagnosis (%)			
- Current	316 (67%)	151 (36%)	
- Ex-smoker	136 (29%)	236 (57%)	
- Never smoker	8 (2%)	28 (7%)	
Unknown	12 (3%)	0 (0%)	<0.001
Mean Pack years			
- Men	42 (10)	45 (25)	0.16
- Women	37 (17)	36 (17)	0.84
- Total	39 (18)	41 (22)	0.14
- Estimated Cigs/day	17*	16	-
Lung function			
% available	325 (69%)	258 (62%)	0.043
Mean FEV1 (SD)	1.71 (0.63)	1.81 (0.71)	0.067
Mean FEV1% predicted[#](SD)	64% (19)	71% (24)	<0.0001
Mean FEV1/FVC (SD)	0.64 (0.12)	0.64 (0.14)	0.83
COPD status			
No COPD	118 (36%)	91 (36%)	-
- GOLD 1	27 (8%)	32 (12%)	
- GOLD 2	122 (38%)	90 (35%)	
- GOLD 3	51 (16%)	32 (12%)	
- GOLD 4	8 (2%)	12 (5%)	0.21
Total COPD	208 (64%)	166 (64%)	-
Histology			
Adenocarcinoma	153 (32%)	176 (42%)	
Squamous Cell	124 (26%)	100 (26%)	
Non-small Cell	64 (14%)	43 (10%)	
Small Cell	102 (22%)	73 (18%)	0.02
Other/Unknown	29 (6%)	23 (6%)	-
Lung Cancer Stage			
NSCLC Stage 1	86 (18%)	93 (22%)	
NSCLC Stage 2	46 (10%)	45 (11%)	
NSCLC Stage 3	92 (19%)	96 (23%)	
NSCLC Stage 4	146 (31%)	105 (25%)	0.13
Small Cell – Limited	41 (9%)	23 (6%)	
Small cell - Extensive	61 (13%)	50 (12%)	0.24
Unknown	-	3 (1%)	-

predicted values for Caucasians were used for Māori (24). *Ref 282 – based on age of initiation of smoking.

Table 8.2 Demographic variables in the Māori and Caucasian lung cancer cases series (Matched comparison)

Demographic Variable	Māori (N=331)	Caucasian (N=330)	P value
Mean age (year) at diagnosis (SD) Range	64yrs (9) 40-87yrs	64yrs (10) 40-88yrs	0.99
Gender - Male (%)	149 (45%)	167 (51%)	
Smoking History			
Smoking at diagnosis (%) - Current - Ex-smoker - Never smoker - Unknown	 214 (65%) 102 (31%) 6 (2%) 9 (3%)	 130 (39%) 179 (54%) 21 (6%) 0 (0%)	 P<0.001
Mean pack years - Men - Women - Total	 42 (20) 38 (17) 40 (18)	 44 (23) 36 (17) 38 (20)	 0.45 0.39 0.76
Lung Function			
% available	238 (72%)	202 (61%)	0.005
Mean FEV_1 (absolute) (SD)	1.68 (0.59)	1.86 (0.72)	0.004
Mean FEV_1 % predicted (SD)	65% (19)	70% (23)	0.008
Mean FEV_1/FVC (SD)	0.64 (0.12)	0.61 (0.14)	0.95
COPD status			
No COPD	83 (25%)	75 (22%)	0.005
- GOLD 1 - GOLD 2 - GOLD 3 - GOLD 4	18 (7.5%) 94 (28%) 37 (15.5%) 5 (2.1%)	21 (13%) 73 (38%) 20 (15%) 12 (5%)	 0.08
COPD Total	154 (46%)	126 (38%)	0.036
Histology			
Adenocarcinoma Squamous Cell Non-Small Cell Small Cell	106 (32%) 91 (28%) 44 (13%) 65 (20%)	142 (43%) 82 (25%) 33 (10%) 60 (18%)	 0.06
Other/Unknown	25 (8%)	12 (4%)	
Lung Cancer Stage			
NSCLC Stage 1 NSCLC Stage 2 NSCLC Stage 3 NSCLC Stage 4	65 (20%) 33 (10%) 64 (19%) 104 (31%)	61 (18%) 38 (12) 81 (25%) 87 (26%)	 0.40
Small Cell Extensive Small Cell Limited	39 (12%) 26 (8%)	42 (13%) 18 (5%)	 0.24

Table 8.3 Demographic variables in the Māori and Caucasian lung cancer cases - Māori-enriched

Demographic Variable	Māori (N=249)	Caucasian (N=415)	P value
Mean age at diagnosis (SD) Range	61 yrs (9) 32-87 yrs	67 yrs (10) 40-92 yrs	<0.001
Male (%)	130 (52%)	213 (51%)	0.019
Smoking History			
Smoking at diagnosis (%) - Current - Ex-smoker - Never smoker - Unknown	 171 (69%) 70 (28%) 2 (1%) 6 (2%)	 151 (36%) 236 (57%) 28 (7%) 0 (0%)	<0.001
Mean Pack years - Men - Women - Total	 43 37 40 (19)	 45 36 41	 0.14
Lung function			
% available	178 (71%)	258 (62%)	-
Mean FEV_1 (SD)	1.79 (0.64)	1.81 (0.71)	0.054
Mean FEV_1% predicted	65%	71%	<0.001
Mean FEV_1/FVC (SD)	0.69 (0.12)	0.64	0.08
COPD status			
No COPD	66 (37%)	92 (36%)	
- GOLD 1 - GOLD 2 - GOLD 3 - GOLD 4	14 (12%) 72 (64%) 22 (20%) 4 (4%)	32 (19%) 90 (54%) 32 (19%) 12 (7%)	 0.21
COPD Total	112 (63%)	166 (64%)	
Histology			
Adenocarcinoma Squamous Cell Non-small Cell Small Cell	68 (27%) 70 (28%) 35 (14%) 58 (23%)	176 (42%) 101 (24%) 43 (10%) 73 (18%)	 0.0020
Other/Unknown	18 (7%)	23 (6%)	
Lung Cancer Stage			
NSCLC Stage 1 NSCLC Stage 2 NSCLC Stage 3 NSCLC Stage 4	44 (18%) 24 (10%) 49 (20%) 77 (31%)	93 (22%) 45 (11%) 96 (23%) 105 (25%)	 0.23
Small Cell - Limited Small Cell - Extensive	22 (9%) 33 (13%)	23 (6%) 50 (12%)	 0.32

predicted values for Caucasians were used for Māori, (283,284).

Table 8.4 Demographic variables in the Māori and Caucasian lung cancer cases - stratified by gender

Demographic Variable	Māori Men N=205	Caucasian Men N=213	Māori Women N=267	Caucasian Women N=202
Mean age at diagnosis (SD) Range	61yrs 40-81yrs	68yrs 41-89yrs	61yrs 32-87yrs	66yrs 40-92yrs
Male (%)	100%	100%	0%	0%
Smoking History				
Smoking at diagnosis (%) - Current - Ex-smoker - Never smoker Unknown	 133 (65%) 61 (30%) 2 (1%) 9 (4%)	 77 (36%) 129 (61%) 7 (3%) -	 183 (69%) 75 (28%) 6 (2%) 3 (1%)	 74 (37%) 108 (53%) 20 (10%) -
Mean Pack years Range	41.5 (4-100)	44.7 (2-191)	36.6 (8-120)	36.3 (4-88)
Lung function				
% available	138 (67%)	129 (61%)	188 (70%)	128 (63%)
Mean FEV_1 (SD)	2.09 (0.65)	2.03 (0.42)	1.49 (0.38)	1.58 (0.44)
Mean FEV_1% predicted	63%	69%	66%	74%
Mean FEV_1/FVC	0.62	0.63	0.65	0.64
COPD status				
No COPD	43 (31%)	44 (34%)	75 (40%)	47 (37%)
- GOLD 1 - GOLD 2 - GOLD 3 - GOLD 4	15 (11%) 52 (38%) 24 (17%) 4 (3%)	15 (12%) 45 (35%) 18 (14%) 7 (5%)	12 (6%) 70 (37%) 27 (14%) 4 (2%)	17 (13%) 45 (35%) 14 (11%) 5 (4%)
Total COPD	95 (69%)	85 (66%)	113 (60%)	81 (63%)
Histology				
Adenocarcinoma[#] Squamous Cell Non-small Cell Small Cell	55 (27%) 67 (33%) 27 (13%) 46(22%)	75 (35%) 69 (32%) 20 (9%) 38 (18%)	98 (37%) 57 (21%) 37 (14%) 56 (21%)	101 (50%) 31 (15%) 23 (11%) 35 (17%)
Other/Unknown	10 (5%)	11 (5%)	19 (7%)	12 (6%)
Lung Cancer Stage				
NSCLC Stage 1 NSCLC Stage 2 NSCLC Stage 3 NSCLC Stage 4	33 (16%) 15 (7%) 44 (21%) 67 (33%)	44 (21%) 24 (11%) 49 (23%) 56 (26%)	53 (20%) 31 (12%) 48 (18%) 79 (30%)	49 (24%) 21 (10%) 47 (23%) 49 (24%)
Small Cell – Limited* Small Cell - Extensive	14 (7%) 32 (16%)	8 (4%) 30 (14%)	27 (10%) 29 (11%)	15 (7%) 20 (10%)

[#]P=0.00002, *P=0.04.

Table 8.5 Demographic variables in the Māori and Caucasian lung cancer cases - stratified by smoking status

Demographic Variable	Māori Current N=316	Caucasian Current N=151	Māori Ex-smoker N=136	Caucasian Ex-smoker N=237
Mean age at diagnosis (SD) Range	60 yrs (8.8) 32-87 yrs	64 yrs (9.5) 40-84 yrs	64 yrs (8.0) 44-83 yrs	69 yrs (9.6) 41-92 yrs
Male (%)	133 (42%)	77 (51%)	61 (45%)	129 (54%)
Smoking History				
Smoking at diagnosis (%) - Current - Ex-smoker	 100% 0%	 100% 0%	 0% 100%	 0% 100%
Mean Pack years - Men - Women - Total	 44.0 37.8 40.4	 52.7 39.2 46.1	 37.7 36.5 37.1	 40.0 34.3 37.4
Lung function				
% available	219 (69%)	78 (52%)	102 (75%)	163 (69%)
Mean FEV_1 (SD)	1.74 (0.63)	1.76 (0.58)	1.62 (0.48)	1.81 (0.69)
Mean FEV_1% predicted	65%	66%	63%	73%
Mean FEV_1/FVC	0.64	0.61	0.64	0.64
COPD status				
No COPD	84	20	33	59
- GOLD 1 - GOLD 2 - GOLD 3 - GOLD 4	18(13%) 79 (58%) 32 (24%) 6 (4%)	4 (7%) 38 (66%) 12 (21%) 4 (7%)	8 (12%) 41 (59%) 18 (26%) 2 (3%)	26 (25%) 51 (49%) 19 (18%) 8 (8%)
Total COPD	135	58	69	104
Histology				
Adenocarcinoma* Squamous Cell Non-small Cell Small Cell	101 (32%) 85 (27%) 39 (12%) 75 (24%)	61 (40%) 28 (19%) 17 (11%) 38 (25%)	38 (28%) 37 (27%) 24 (18%) 25 (18%)	94 (40%) 70 (30%) 23 (10%) 35 (15%)
Other/Unknown	16 (5%)	7 (5%)	12 (9%)	15 (6%)
Lung Cancer Stage				
NSCLC Stage 1 NSCLC Stage 2 NSCLC Stage 3 NSCLC Stage 4	53 (17%) 29 (9%) 65 (21%) 94 (30%)	27 (18%) 12 (8%) 33 (22%) 41 (27%)	29 (21%) 13 (10%) 24 (18%) 45 (33%)	58 (25%) 33 (14%) 55 (24%) 54 (23%)
Small Cell - Limited Small Cell - Extensive	30 (9%) 45 (14%)	12 (8%) 26 (17%)	9 (7%) 16 (12%)	11 (5%) 24 (10%)

*P=0.01.

Table 8.6 Demographic variables in the Māori and Caucasian lung cancer cases - stratified by COPD status

Demographic Variable	Māori COPD N=208	Caucasian COPD N=166	Māori No COPD N=118	Caucasian No COPD N=91
Mean age at diagnosis (SD)	64yrs	69yrs	59yrs	65yrs
Range	42-83yrs	40-92yrs	32-87yrs	43-89yrs
Male (%)	96 (46%)	85 (51%)	42 (36%)	44 (48%)
Smoking History				
Smoking at diagnosis (%)				
- Current	135 (65%)	58 (35%)	81 (69%)	20 (22%)
- Ex-smoker	69 (33%)	104 (63%)	33 (28%)	59 (65%)
- Never smoker	1 (0.5%)	4 (2%)	2 (1%)	12 (13%)
Unknown	3 (1%)	-	2 (1%)	-
Mean Pack years				
- Men	41	46	44	39
- Women	38	41	34	34
- Total	40	44	37	37
Lung function				
% available	100%	100%	100%	100%
Mean FEV_1 (SD)	1.56 (0.48)	1.59 (0.50)	1.99 (0.85)	2.22 (0.87)
Mean FEV_1% predicted	59%	63%	74%	86.5%
Mean FEV_1/FVC	0.57	0.56	0.76	0.78
COPD status				
No COPD	0	0	118	91
- GOLD 1	27 (13%)	32 (19%)	-	-
- GOLD 2	122 (59%)	90 (54%)	-	-
- GOLD 3	51 (24%)	32 (19%)	-	-
- GOLD 4	8 (4%)	12 (7%)	-	-
Total COPD	208	166	0	0
Histology				
Adenocarcinoma*	68 (33%)	69 (42%)	39 (33%)	50 (55%)
Squamous Cell	71 (34%)	53 (32%)	37 (31%)	26 (29%)
Non-small Cell	22 (11%)	17 (10%)	16 (14%)	6 (7%)
Small Cell	34 (16%)	12 (7%)	14 (12%)	5 (5%)
Other/Unknown	13 (6%)	15 (9%)	12 (10%)	4 (4%)
Lung Cancer Stage				
NSCLC Stage 1#	49 (24%)	58 (35%)	32 (27%)	28 (31%)
NSCLC Stage 2	28 (13%)	26 (16%)	12 (10%)	15 (16%)
NSCLC Stage 3	48 (23%)	41 (25%)	30 (25%)	23 (25%)
NSCLC Stage 4	49 (24%)	28 (17%)	30 (25%)	19 (21%)
Small Cell - Limited	13 (6%)	4 (2%)	9 (8%)	3 (3%)
Small Cell - Extensive	21 (10%)	8 (5%)	5 (4%)	2 (2%)

*P=0.01, # P<0.00001.

8.4 Discussion

In this comparative study, where the demographic characteristics of 472 Māori lung cancer patients were compared with 415 Caucasian lung cancer cases, I found a number of significant differences. Compared to Caucasians, Māori had a younger mean age at diagnosis (61 years old vs 67 years old), fewer never smokers (2% vs 7%), more aggressive histological subtypes and more advanced airflow limitation. These findings could not be explained by differences in mean pack years of smoking or cigarettes per day and in a sensitivity analysis persisted despite close matching for age, gender, and smoking exposure (pack years and cigarettes/day). The differences in COPD prevalence and airflow limitation (FEV_1% predicted and FEV_1/FVC) were most apparent at lower smoking levels of exposure (<20 pack years) and attenuated at higher exposure (Figure 8.5). The differences I reported in mean age at diagnosis, smoking status and smoking exposure replicate those of a smaller previously published study which showed that despite comparable smoking exposure, Māori were younger at diagnosis (63 years old vs 71 years old) and less likely to be never smokers (1% vs 8% in Caucasians).[281] However in this current study, Māori were more likely to have squamous cell, small cell and non-small cell histology and less likely to have adenocarcinoma histology, relative to Caucasians, (Figure 8.9). These inter-ethnic differences in histology were most apparent according to gender and COPD status (Figure 8.10 and Table 8.3 - Table 8.6), where more aggressive histological subtypes have been reported in the latter.[93] Collectively these findings suggest that not only do Māori appear more susceptible to getting lung cancer, they have worse airflow limitation and more aggressive subtypes of lung cancer. The results of these two studies suggest lung cancer in Māori manifests differently to that in Caucasians and raises the questions, “Could this greater susceptibility to lung cancer in Māori be mediated through a greater susceptibility to COPD?” and “Could the greater incidence of lung cancer in Māori stem in part from a greater overall susceptibility to smoking?”

8.4.1 Airflow limitation and smoking by ethnicity

The first novel finding of my study is that at all levels of smoking exposure, Māori with lung cancer have greater airflow limitation (lower FEV_1/FVC and FEV_1%predicted) compared to Caucasian lung cancer cases, (Figure 8.5 and Figure 8.6). More importantly, at low levels of smoking exposure, Māori had a two-fold greater COPD prevalence (Figure 8.7). Indeed,

when smoking dose and prevalence of COPD was compared between Māori and Caucasian lung cancer cases, the expected dose-response relationship evident for Caucasians at lower smoking exposure (<30 pack years) is lost in Māori (Figure 8.7). This finding bares similarities with that found in Hawaiian suggesting an important ethnicity-smoking interaction which is most evident at lower smoking exposure levels (Figure 8.1 and Figure 8.2).[148] We propose that at these low exposure levels, Polynesians (and possibly African-Americans) have a heightened responsiveness to smoking not seen to the same degree in Caucasians.[146-148] In another US-based study, it was found that compared to Caucasians with the same severity of COPD, African-Americans were younger, smoked less and had greater lung function decline.[284] In Caucasians and other ethnic groups, where the dose-response relationship is evident,[146-148] high smoking exposures (>30 pack years or >30 cigarettes/day, (Figure 8.1 and Figure 8.2) presumably overwhelm the innate capacity of smokers to tolerate smoke exposure at these higher exposures.[83,198]

8.4.2 Ethnicity, smoking exposure, and risk of lung cancer

In the introduction to this paper we outline the results of the study by Haiman et al. showing that in contrast to never smokers, where the risk of lung cancer in Polynesians (Hawaiians) and Caucasians is comparable (Figure 8.1), Polynesian current and former smokers had approximately 1.2 and 1.8 fold greater risk of lung cancer respectively than Caucasians.[146] After consideration of other demographic data, Haiman et al. suggest that an important smoking-by-ethnicity effect exists independent of smoking exposure, education and diet. On re-analysis of their data,[148] it has been shown that the expected dose-response effect of increasing lung cancer risk with increased smoking exposure dose seen for Caucasians (and Asians and Hispanics) was lost for Polynesians (Figure 8.2). This suggests a greater propensity to lung cancer in Hawaiians relative to Caucasians, especially at lower smoking exposure. When lung cancer rates were calculated according to histology and gender, in heavy smokers, Polynesian (regardless of gender) had greater squamous cell cancers and small cell cancers than corresponding Caucasians.[146] On this basis we decided to examine the difference in histology according to gender, smoking status and the presence of COPD using a stratified approach (Table 8.4 - Table 8.6). Such an analysis underpins the view of our group that important differences underlying lung cancer susceptibility may be missed if stratification for important contributing variables is not examined.[153]

8.4.3 Effect of gender, smoking status, and the presence of COPD on histology in Māori and NZ Caucasian with lung cancer

The second novel finding of this inter-ethnic case-case study is the relative decrease in the proportion of lung cancer cases in Māori that were adenocarcinomas (generally less aggressive biology), (Figure 8.9 and Figure 8.10), especially in those with no COPD (Table 8.6) and of female gender (Table 8.4). Compensatory increases were seen in Māori for squamous cell, non-small cell, and small cell lung cancer, where the latter two histological subgroups contributed independently to greater all-cause mortality along with Māori ethnicity, age, male gender, and stage, (Figure 8.11). This is in contrast to finding no differences according to severity of COPD (GOLD grade) and clinical stage at diagnosis (Table 8.3 – Table 8.6). We found that the proportion of adenocarcinomas in Māori was less than for Caucasians, and that the proportion of other lung cancers such as squamous cell, non-small cell and small cell were correspondingly greater. This was the case despite comparable smoking exposure history (Table 8.1 - Table 8.6), matching age (Table 8.2) and stratifying by gender (Table 8.4), smoking status (Table 8.5) and the presence of COPD (Table 8.6). In contrast, there was no difference in the severity of COPD (GOLD grade) or the clinical stage, either before or after stratifying for the effects of gender, smoking status, and COPD, or matching for age. This suggests that these differences in histology, where the prevalence of adenocarcinoma is reduced in Māori while other lung cancer histological subgroups are greater (Figure 8.9), is not the result of differences in age, gender, smoking status (current vs ex-smokers), smoking exposure or lung function. Differences in histology extended to small cell cancer where the prevalence was 2-fold greater in Māori relative to Caucasians stratified by the presence of COPD suggesting an ethnicity-by-COPD interaction (Table 8.6). We suggest that these differences in the lung cancer histology, reflects differences in the pathobiology of lung cancer in Māori. This finding that Māori with lung cancer have worse lung function than Caucasians, at low smoking exposure, suggests Māori may be more “susceptible” to the adverse effects of smoking. Indeed, Māori have roughly 2-fold greater rates of COPD than Caucasians.[158] Our group and others have shown that in Caucasians with lung cancer, the presence of COPD is associated with more aggressive lung cancer such as squamous cell, non-small cell, and small cell cancers.[93] Interestingly in the US study,[146] inter-ethnic differences between Polynesians and Caucasians were greatest for adenocarcinomas and small cell cancers. Interestingly, in my study small cell cancer prevalence was 1.5 fold greater

in current smokers compared to ex-smokers (Table 8.5), suggesting an interaction between histology and smoking status as previously described.[285] Based on this current study, our group believe that the higher prevalence of more aggressive lung cancer histological subtypes in Māori may also be related to overlapping pathogenic pathways underlying airway remodelling (or COPD).[34] Although the current study cannot prove this relationship, my robust finding that differences in lung cancer histology exists between Māori and Caucasians requires further analyses. These apparent differences in the biology of lung cancer in Māori compared to Caucasian have important implications. Outlined below are the results of an analysis on all-cause mortality in lung cancers diagnosed during this study (see section 8.4.6).

8.4.4 Differences in age at diagnosis according to Māori or Caucasian ethnicity

The significantly younger age at diagnosis in Māori might also indicate a greater susceptibility although ethnic differences in aging, age structure and age-specific mortality may be relevant here (personal communication with Dr T Blakely, NZ Epidemiologist).[159,279] In the US-based lung cancer study,[146] Hawaiian were noted to have the youngest mean age at diagnosis consistent with our finding in Māori, although this was only a small difference compared with other US-based ethnicities. Assuming that age of initiation of daily smoking is similar between Māori and Caucasian in New Zealand (16 years old),[282] and the pack year exposure history being comparable in our lung cancer case series (Table 8.1), it appears unlikely that differences in smoking exposure per se (i.e., cigarette consumption) account for the differences observed. Further support for our hypothesis comes from New Zealand data showing that after adjustment for age and smoking, Māori are two-fold more likely to get COPD than Caucasians.[158] Indeed, two studies reported several decades ago showed that Māori men who smoked had worse lung function than their Caucasian counterparts suggesting greater susceptibility in Māori.[286,287] Collectively, these findings demonstrate that the greater propensity to lung cancer in Polynesians relative to other ethnic groups is unrelated to lifetime differences in smoking exposure. Instead, these differences result from a greater inherent susceptibility to the adverse effects of smoking on the lungs generally.[146-148] The basis of this hypothesis is discussed below.

As the increased incidence of lung cancer in Māori appears to be associated with earlier onset of disease, greater airflow limitation and a greater propensity to COPD at lower smoking exposure, we are of the view that Māori may be inherently at greater risk of lung cancer through as yet unknown pathogenic or sociocultural mechanisms. Differences in the metabolism of nicotine have been proposed to explain inter-ethnic differences, but the data are inconclusive.[288] In studies including Native Hawaiians, no significant difference in CYP2A enzyme activity was found compared to Caucasians.[289] Hawaiians have shared genetic ancestry with Māori and have the highest incidence of lung cancer in the United States,[86,146,198] suggesting shared genetic factors may in part underlie this increased susceptibility.[146-147] Further studies will be needed to test this hypothesis. Another possibility is the effect of high rates of maternal smoking among Māori women of reproductive age compared to Caucasian women.[290] Maternal smoking rates in Māori have been historically high for several decades relative to Caucasians, a feature of smoking prevalence shared with Hawaiian women. The contribution of marijuana smoking to lung cancer in Māori is also a possibility although this has been shown to be relevant in only a small proportion of lung cancer cases (primarily those <50 yr old, ≈10%) so unlikely to explain our observations.[291] Differences in the innate immune response to smoking and downstream exaggerated inflammation has been suggested as another possible mechanism underlying inter-ethnic differences in susceptibility to lung cancer.[146,148,29] We propose that the most likely factor underlying my observed differences are secondary to the biological difference in susceptibility of Māori to cigarette smoke, where Māori are more sensitive to one or a combination of the addictive, carcinogenic or pro-inflammatory substances in smoke, compared to NZ Caucasians. Collectively, poorer lung function, higher COPD prevalence at low pack years, more aggressive subtypes of lung cancer and earlier age at lung cancer diagnosis,[33,53,131,219] suggest that Māori have a different biological susceptibility to lung cancer than NZ Caucasians (Figure 8.3-Figure 8.10).

We found that mean age at diagnosis was significantly lower in Māori compared to Caucasians (61 vs 67 years old, Table 8.1) comparable to the study by Stevens et al.[281] This difference persisted in our stratified comparisons, first using Māori enriched for Māori ancestry (Māori surname, Table 8.3), and second after stratification for gender and the presence of COPD (Table 8.4 and Table 8.6). Interestingly, after stratification by smoking

status (Table 8.5), we found age at diagnosis was lower in Māori by 4 years in current smokers (60 vs 64 years old) and 5 years in ex-smokers (64 vs 69 years old). It is notable that the mean age at diagnosis is younger in current smokers, consistent with the literature,[293] with Māori current smokers being on average 9 years younger than Caucasian ex-smokers (60 vs 69 years old). However, after stratification by smoking status, we found the difference in age at diagnosis was reduced. Studies comparing Native Hawaiian with "non-Native Hawaiian" have also shown a younger age at diagnosis of about 5-6 years.[294] As previously noted, these age differences may reflect ethnic differences in age structure across each ethnic group. Unexpectedly, age was not a strong predictor of all-cause mortality in this current study and this may be due to the strong independent effects we found for ethnicity (Māori), aggressive histology (Small cell and Non-small cell), propensity to COPD at low smoking exposure and advanced clinical stage (see later discussion). Our matched comparison (Table 8.2), was primarily done to match for age, in addition to gender and smoking exposure, where we found the most notable difference between Māori and Caucasian was for histology despite the smaller numbers (P=0.07). It is acknowledged that while surname has been used to help "enrich" ancestry in studies of indigenous peoples,[295] where varying degrees of genetic admixture has occurred following colonisation, it remains inferior to formal genetic testing.

8.4.5 Effect of demographic characteristics on all-cause mortality in Māori compared to Caucasian lung cancer cases

In order to minimise biases from variation in the investigation and treatment of lung cancer, or variations in temporal trends or institutional policy on lung cancer management, we compared the all-cause mortality of Māori and Caucasian lung cancer cases diagnosed from our single institution between 2004 and 2007. While it is unlikely that Māori lung cancer patients diagnosed in our hospital have been missed, lung cancers diagnosed in the community or on post-mortem may not have been included in this study. However, the same can be said for my comparator population of NZ Caucasian lung cancer patients. Based on the results from 312 Caucasian cases and 81 Māori cases (recruited between 2004-2007), we found the mean survival in months was significantly less in Māori 29.2 months (SE 3.2 months) compared to Caucasians (35.3 months (SE2.3) ($p<0.05$)). In a Cox-Proportional analyses we found that the Hazard Ratios (HR) for all-cause mortality were; ethnicity

(Māori=1.4, 95% Confidence Limits (95% CI) = 1.0-1.9, P=0.03); gender (female HR=0.8, 95% CI=0.6-0.9, P=0.04), advanced clinical stage (NSCLC Stage 4 or Small Cell-extensive = 2.6, 95% CI 2.0-3.4, P<0.0001), and histology (small cell or non-small cell relative to adenocarcinoma=1.4, 95% CI=1.0-2.1, P=0.04). Similar findings have been described for lung cancer mortality in Hawaii where Hawaiian ancestry conferred a hazard risk of 1.4.[294] Importantly in this current study, age, smoking status, and COPD status were not significant contributors to all-cause mortality (Figure 8.11). Of note is that while clinical stage was an important determinant of all-cause mortality, it was not very different between Māori and Caucasian lung cancer cases. We do note that the non-significant deficit in early stage 1 disease in Māori (Table 8.1, 18% vs 22%, P>0.05) is magnified in those with COPD (Table 8.6, 24% vs 35%, P<0.05). This likely reflects that Māori ethnicity, advanced stage and aggressive histology independently obscure any effects of COPD on mortality. This observation casts some doubt on the currently accepted view that Māori lung cancer cases have greater mortality because of demographic variables such as smoking and presenting late due to poor access to doctors.[181] A similar conclusion was proposed from the study in Hawaiians.[294] Our results suggest poor outcomes for lung cancer may be in part related to both differences in the biology of the lung cancers (more aggressive histological subtypes) and greater disposition to airflow limitation, where non-respiratory mortality (cardiovascular disease and other cancers) has been shown to be greater.[20] Such an observation may have implications in the use of computed tomography for lung cancer screening or early case finding.[93]

8.4.6 Basis of biological differences between lung cancer in Māori compared to Caucasian

One reason for the ethnic differences in lung cancer susceptibility may be differences in how nicotine or carcinogens are absorbed or metabolised by different racial groups. With regards to absorption, one explanation may be that Māori smoke cigarettes differently to NZ Europeans, inhaling deeper per cigarette as has been proposed amongst African-Americans.[11,288-289] This behaviour has been associated with genetic polymorphisms of CYP2A enzymes, which are responsible for nicotine metabolism and activation of carcinogenic substances in cigarette smoke.[288-289,296-297] Increased CYP2A enzyme activity has been correlated to increased smoking depth, and consequently an increased dose of the carcinogen nitrosamine 4-(methylnitro-samino)-1-(3-pyridyl)-1-butanone (NNK) per

cigarette.[288,296-297] Differences in smoking depth and nicotine metabolism remain possible explanations for greater susceptibility of lung cancer and COPD in Māori. Differences in nicotine addiction have been examined in different ethnic groups in Hawaii and, based on studies assessing nicotine consumption relative to nicotine metabolism, concluded that Hawaiians were more addicted to nicotine secondary to a higher rate of metabolism.[298] This was supported by an earlier study correlating higher addiction rates in Native Hawaiian compared to Caucasians.[299] Based on these studies, we propose that the one possible factor underlying greater "susceptibility" of Māori to cigarette smoke stems from Māori being more sensitive to the addictive properties of smoking relative to NZ Caucasians. Another possibility has been differences in diet such as fruit and vegetables which have been linked to lower rates of lung cancer.[300] However, in the study by Haiman and colleagues they could find no effect from the intake of fruit and vegetables to account for the ethnic differences they reported.[146] However, it also remains possible that through shared ancestry,[301] Hawaiian and Māori have an exaggerated immune response to smoking in the lungs.[146-148] We conclude that any one, or combination, of these various factors could contribute to the greater rates of current smoking, earlier age of onset of lung cancer, worse airflow limitation and greater tendency to more aggressive forms of lung cancer in Māori relative to Caucasians. Regardless of the basis of these important differences, aggressive tobacco control measures are required if disparities between Māori and Caucasians are to be addressed.

8.4.7 Study limitations

We acknowledge that this study has several limitations. First, while the Caucasian lung cancer cases were collected prospectively from 2004-2008, the Māori lung cancer cases were identified using a retrospective design. However, the Māori cases were identified from the same tertiary hospital, serving the same geographical region and during an overlapping time interval. Second, only about two thirds of the lung cancer case series had spirometry before or around the time of diagnosis. This is likely to introduce some bias but, as discussed above, it likely affects both groups equally and remains the more accurate way to report COPD prevalence compared to using COPD history as stated in medical notes.[281] Third, Māori ancestry was self-reported and was not confirmed (or quantified) on a genetic basis. It is likely that a large proportion of the 'Māori' population has a variable quantity of NZ

Caucasian genetic ancestry (estimated to be about 30% on average). This means our results might actually underestimate these ethnic differences compared to a study using genetically defined ancestry. To test this hypothesis, we re-examined our results using only Māori lung cancer cases with a Māori surname and found the differences were slightly more marked with Adenocarcinoma prevalence of 27% in Māori compared to 42% in Caucasian (Figure 8.9 and Table 8.3). While this approach has strengthened our original findings (rather than diluting them), we cannot be more definitive about this observation in the absence of using genetic ancestry markers to better assign ancestry. Given the retrospective recruitment of Māori lung cancer cases, it was not possible to capture ancestry in their grandparents as was done for Caucasians. That said, it is likely we have captured a strongly Caucasian population and compared them to a Māori lung cancer case series of variable Māori-Caucasian ancestry. Fourth, the case-case study design is inferior to a cohort study where lung cancer cases can be identified prospectively, with relevant demographic data collected prior to the diagnosis of lung cancer.

8.5 Conclusion

In conclusion, the results of this study suggest that Māori who smoke may be at greater risk of lung cancer, and possibly COPD, than their Caucasian counterparts. The finding that Māori have worse airflow limitation and more aggressive lung cancer is novel to this study, as is the finding that Māori ethnicity and histology independently contribute to all-cause mortality. Regardless of the basis of this heightened susceptibility, tobacco control measures aimed at substantially reducing all exposure to smoking are indicated if disparities in outcomes for New Zealand Māori are to be successfully addressed.[147,157]

Chapter 9 Chr15q25 Genetic Variant (rs16969968) Independently Confers Risk of Lung Cancer, COPD, and Smoking Intensity in a Prospective Study of High Risk Smokers

'THINK OF ME GENTLY'
Think of me gently,
You who carry, The bouquet of genes.
For I may be like a bell,
But goes on beyond your range.
If this is so, I may have a need,
Of your sympathetic thoughts
The warmness of your minds'
(Poem by Jock R Mackenzie, 2003, NZ)

9.1 Introduction

Large cross-sectional genome wide association studies (GWAS) have consistently reported an association between the nicotinic cholinergic receptor subunit (*CHRNA5*) gene locus on chromosome 15q25 and lung cancer.[165-167] This locus has also been linked in cross-sectional studies to smoking exposure and addiction to smoking.[12,302,303] These findings have led to some debate as to whether the relationship between *CHRNA5* and lung cancer is mediated through it effects with smoking, where increased risk of lung cancer is related to increased smoking exposure.[304,305] Soon after the original GWAS reported the lung cancer link, we reported that this locus was also associated with chronic obstructive pulmonary disease (COPD).[11] In our case-control study we compared the frequency of the *CHNRA5* variant rs16969968 in healthy smokers (normal lung function), with smokers with COPD and/or lung cancer relatively closely matched for age and smoking exposure history. This variant has been shown to change the structure and function of this subunit, consistent with it being a "causative" disease variant.[306,307] Our findings for COPD were quickly confirmed in a large COPD GWAS and subsequent case-control studies.[164] However, because COPD is also associated with lung cancer, the interactive effects of this locus on the development of lung cancer, independent of its effects on smoking or the development of COPD remain unclear.[304,308]

One way to explore the effect of this locus on the development of lung cancer is through a large prospective study, where current or former smokers at risk of lung cancer are followed

prospectively after their smoking history and lung function were fully documented at baseline.[93,219] This reduces the bias that comes from using lung function performed exclusively for pre-operative assessment; using lung cancers collected cross-sectionally and thus subject to survivor bias; or smoking history collected after the diagnosis of COPD or lung cancer was made.[309] A prospective study design including baseline spirometry also helps address the question of confounding, where biomarker associations made with lung cancer are potentially based on an unrecognised association with COPD.[153] The National Lung Screening Trial is a large prospective study of 53,000 high risk smokers followed over a mean of 6.4 years.[90] Among this cohort there were 9,270 Non-Hispanic whites, including 3,500 current or former smokers with airflow limitation (35%) at baseline, and 380 lung cancers diagnosed during the follow up period. Preliminary results of this study have been reported in abstract form.[15]

Hypothesis: The CHRNA receptor variant rs16969968 increases risk of smoking exposure, COPD, and lung cancer independently.

Aim: To assess the relationship between the CHRNA5 genotype (allele and genotype) in high risk smokers with and without COPD at baseline and the subsequent development of lung cancer.

Objective: Using a post-hoc analysis for the biomarker arm of the NLST-ACRIN study to compare what effect the CHRNA5 variant has on lung cancer relative to its effect on smoking and COPD.

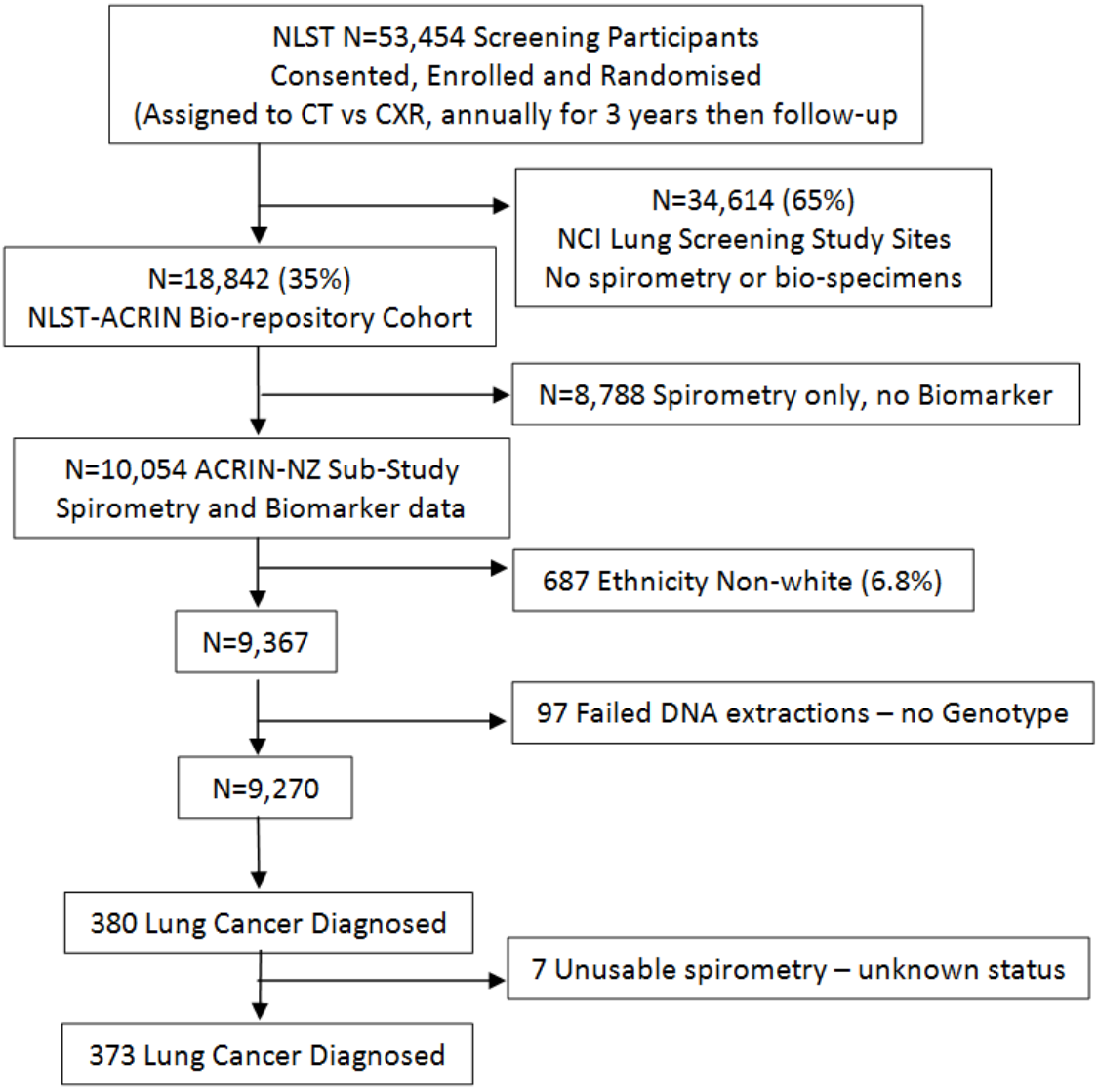
9.2 Methods

9.2.1 Subjects

This is a secondary data analysis of the National Lung Screening Trial (NLST). The recruitment and study design of this trial, involving 53,452 screening participants yielding 2,058 histology confirmed lung cancers, has been described elsewhere.[90] In the ACRIN sub-cohort of the NLST, participants from 23 centres agreed to undergo baseline pre-bronchodilator spirometry (NLST-ACRIN Bio repository Cohort, N=18,842) and, for a subgroup (Figure 9.1), blood sampling for biomarker analysis (N=10,054). Demographic data, including history of pre-morbid disease, were collected through an extensive questionnaire and shows this NLST-

ACRIN cohort to be highly representative of the full NLST cohort.[93,219] From the total group of 10,054 (including all ethnicities), we analysed genomic data for Non-Hispanic whites comprising 9,270 high risk smokers from which 380 lung cancers were diagnosed during the study follow up (Figure 9.1).

Figure 9.1 Consort diagram of the genetic study subgroup from the NLST



Legend: NLST=National Lung Screening Trial, ACRIN= American College of Radiology, Imaging Network, NCI= National Cancer Institute, NZ= New Zealand.

9.2.2 Pulmonary function testing

In the NLST-ACRIN cohort, pre-bronchodilator spirometry was measured at baseline screening (T0) in the majority of participants meeting previously published criteria.[90] The spirometry was measured by trained staff using a Spiropro spirometer

(eResearchTechnology, GmbH, Germany). The severity of airflow limitation was defined according to the Global Initiative on Chronic Obstructive Lung Disease (GOLD) criteria grades 1-4 (www.GOLD.org accessed February 2, 2020). Those with no airflow limitation were further sub-grouped into those with normal lungs ("Resistant Smokers") and Restrictive Spirometry, where the latter is defined as $FEV_1/FVC \geq 0.70$ and FEV_1%predicted <80% (http://www.copdgene.org/study-design accessed February 2, 2020). For comparative purposes in this study,[11,308] COPD was defined as GOLD 2-4 grade.

9.2.3 Lung cancer outcomes

Lung cancer cases included all those diagnosed during the trial (N=380), whether screen or non-screen detected (interval), or prevalent (diagnosed at T0 or during the first year) or incident lung cancers (diagnosed during subsequent years T1 to T6) or at post-mortem.[90] All lung cancer cases were confirmed on histological sampling according to accepted international classification criteria. Lung function results and mortality outcomes were available for 373 of the 380 lung cancer cases (98% of total). The NLST was terminated early when the endpoint of a 20% reduction in lung cancer specific mortality in the computed tomography (CT) arm, relative to the chest x-ray (CXR) arm, was reached with a mean follow-up of 6.4 years. Cause of death was a primary outcome for the NLST, and was ascertained through review of clinical records and death certification.

9.2.4 Genotyping

Genomic DNA was extracted from buffy coat samples using standard salt-based methods and purified genomic DNA was aliquoted (10 $ng \cdot \mu L^{-1}$ concentration) into 384-well plates. Samples were genotyped for rs16969968 of the *CHRNA* 3/5 gene using the Sequenom™ system (Sequenom™ Autoflex Mass Spectrometer and Samsung 24 pin nano-dispenser) by Agena (San Diego, USA). The Sequenom™ sequences were designed in house by Agena with amplification and separation methods (iPLEX™, www.sequenom.com) as previously described.[154]

9.2.5 Statistical analysis

Differences in all outcomes according to *CHRNA5* genotype were examined using both allelic (0, 1 or 2 "A" alleles), genotype (AA vs AG vs GG) and recessive (AA vs AG/GG combined) models. Chi-square tests, test for trend or Fisher's exact test (for small cell counts) were

used to compare categorical variables, and t-tests or Kruskal-Wallis tests (for non-Normally distributed variables) were used to compare continuous variables by genotype. Lung function measures at baseline were assessed using robust linear regression to account for outlying values and adjust for pack years, age, and sex. Prevalence rates and 95% confidence intervals (CIs) for lung cancer and mortality outcomes were calculated per 1,000 person-years. Cox proportional hazards models were used to compare survival adjusted for pack years, with estimates presented as hazard ratios (HRs) and 95% CIs. The proportional hazards assumption was verified. Multiple logistic regression modelling was used to model odds with and without adjustment as indicated. For models of association with COPD status and lung cancer diagnosis, allelic (log additive) and recessive models were compared to the general genotype model (AA vs AG vs GG) using a likelihood ratio test (LRT), AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) metrics, to assess which model provided the best fit. Mediation analysis was performed using the CAUSALMED procedure of SAS (SAS v9.4, SAS Institute Inc, Cary NC), fitting a binomial distribution with logit link function to the binary outcome lung cancer outcome as function of the direct and indirect effect of the binary *CHRNA5* recessive variable with dichotomous GOLD group as a mediator and pack years >50 years as categorical covariate. Total effect decompositions were estimated. Statistical significance was defined as a two tailed $P<0.05$. All analyses were performed using SAS (V 9.4, SAS Institute Inc, Cary NC) or STATA statistical software (version 16, College Station, TX: StataCorp LLC).

9.3 Results

Of the Non-Hispanic Whites (N=9,367, 93%) of the total NLST-ACRIN cohort, genotype data for the *CHRNA5* nicotinic receptor polymorphism (rs16969968) was available for 9,270 subjects (99%). In Table 9.1, the demographic variables and pre-morbid self-reported diseases were compared according to the AA (homozygous minor allele, 12%), AG (heterozygous, 46%) and GG (homozygous major allele, 42%). These genotype frequencies accord with Hardy-Weinberg equilibrium and were consistent with the published frequencies in other Caucasian populations.[11-12,164-167,302,303] Using this method, we have previously shown 100% concordance in genotyping with a SNP in complete linkage disequilibrium.[15,310]

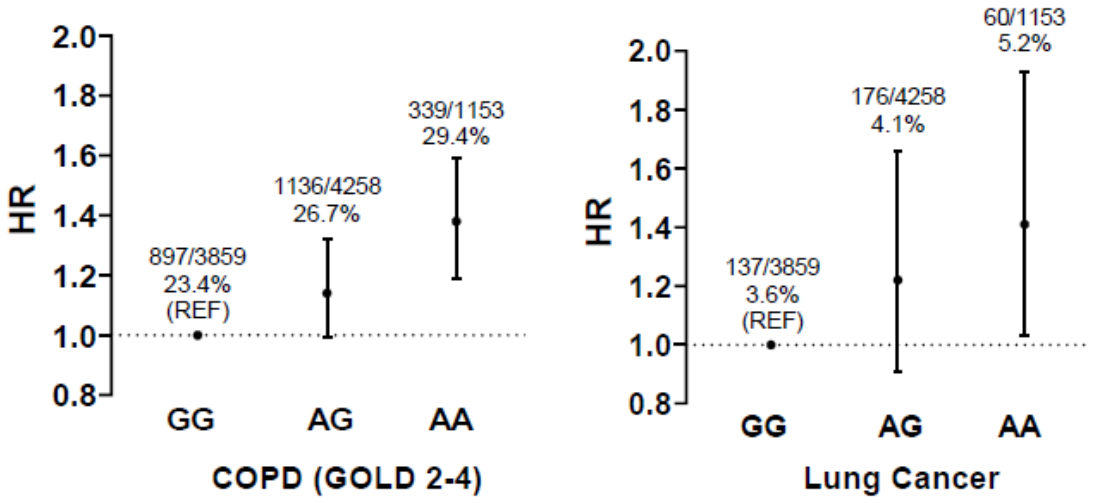
We found that across the 3 genotypes, age, gender, family history for lung cancer, smoking status, years quit, smoking duration, BMI and educational level were not significantly different (Table 9.1). We found the A allele (allelic model) and AA genotype (recessive model) were associated with higher pack years (P<0.001) and higher cigarettes per day (P<0.001) with no differences in years quit or years smoked. A similar finding was found for doctor diagnosed (self-reported) COPD but not for other pre-morbid diseases at baseline (Table 9.1).

Comparable to our first study,[11] on comparing lung function and COPD severity (GOLD grade) in Table 9.2, we found that the AA genotype (recessive model) was associated with lower FEV_1% predicted (P=0.005), greater COPD (GOLD 2-4) frequency (P<0.001) but not GOLD 1 COPD. Indeed, the AA frequency increased as the severity of airflow limitation increased from "Resistant smokers" (11.6%), GOLD 1 (11.5%), GOLD 2-4 (14.3%) and GOLD 3-4 (15.0%), (Table 9.2). Moreover, the fully adjusted effect estimates were greatest in the recessive model and most specific for airflow limitation (least significant for FVC% predicted), (Table 9.3). However, on comparing the odds of COPD (GOLD 2-4) adjusted for smoking intensity, there was evidence that the general genotype model fitted the data better than a recessive model (LRT P=0.0005), and further that the allelic model provided the best fit (lowest AIC and BIC values), with no evidence that the genotype model was an improvement (LRT vs allelic model, p=0.387) (Table 9.4a). Adjusted for smoking intensity, there was a 1.16 (95% CI 1.08, 1.25, P=<0.001) increase in the odds of COPD for each additional A allele (Table 9.2, Figure 9.2). Collectively these findings suggest a log additive (allelic) model is most consistent for COPD (GOLD 2-4) at this locus (Table 9.2, Table 9.3, and Figure 9.2). In those with no airflow limitation, consisting of those with "restrictive" or "resistant" spirometry subgroups, there were no allele or genotype differences (Table 9.2).

Consistent with our previous study,[11] and after adjustment for smoking, the AA genotype was also associated with a greater prevalence and incidence of lung cancer overall (Table 9.5, Figure 9.2). In the current study where lung cancer was identified prospectively the AA genotype relative to the GG genotype had an adjusted HR=1.39 (95 % CI 1.02,1.88, P=0.034) and for the AG genotype an adjusted HR=1.14 (0.91,1.43, P=0.243) (Table 9.5, Figure 9.2). Adjusted for smoking intensity there was a 1.17 (95% CI 1,01, 1.36, P=0.035) increase in the hazard of a lung cancer overall for each additional A allele (Table 9.5, Table 9.4b). On

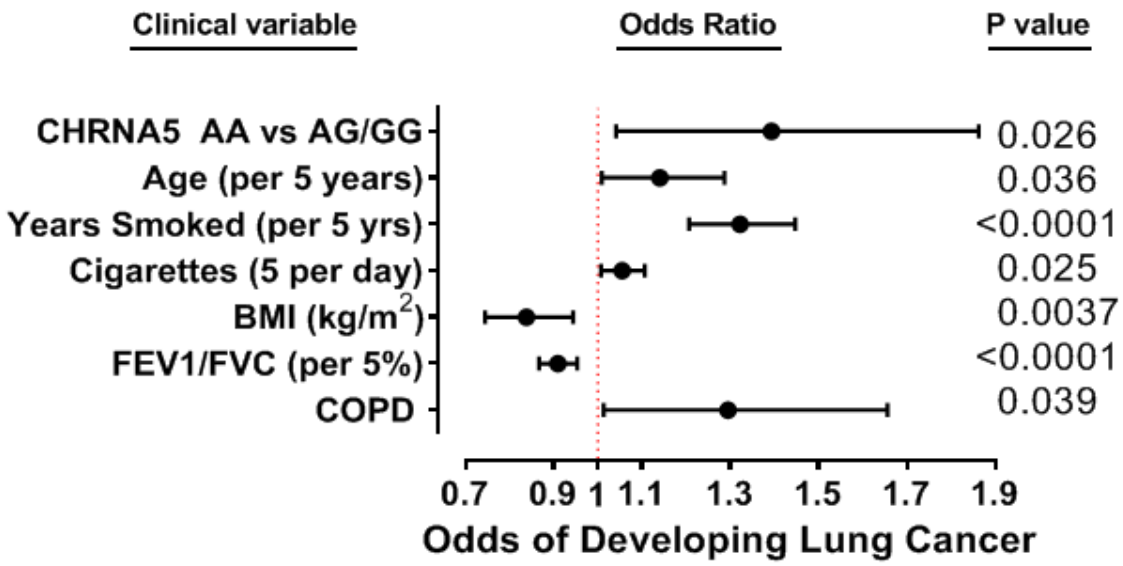
comparing the risk of a lung cancer using the likelihood ratio test, and adjusted for smoking intensity, we found no evidence that the general genotype model fitted the data better than a recessive model (LRT P=0.242) or an allelic model (LRT p=0,770). While AIC and BIC values were slightly lower in the allelic model (Table 9.4.b), we note that after stratifying lung cancer cases for COPD (Table 9.6), the lung cancer with COPD data best fits an allelic model while for lung cancer alone the data best fits a recessive model (based on LRT analysis and AIC results, data not shown). This is consistent with our finding that the AA genotype but not the AG genotype, was significantly associated with lung cancer relative to GG (Figure 9.2). We found no major differences in lung cancer characteristics or demographic variables across the 3 genotypes (Table 9.5). While the allelic log additive model best fits the data for COPD, (Table 9.2, Table 9.4a, Table 9.6, Figure 9.2), we suggest the recessive model appears to best fit the data for lung cancer, especially when those with COPD are excluded (Table 9.4b, Table 9.5, Table 9.6,). In Table 9.6 relative to the referent group (AG=45%), the excess in AG genotype is a feature of COPD GOLD 2-4 (AG=48%), and lung cancer with COPD (50%), but not for those with lung cancer alone (44%).

Figure 9.2 Odds Ratio for COPD (GOLD 2-4) and hazard ratio for lung cancer according to rs16969968 *CHRNA5* genotype after adjustment for smoking pack years (+/- 95% CI) and referenced against the GG genotype



In a multiple logistic regression, where the AA genotype was compared to the other genotypes, age, and smoking duration (grouped in 5-year bands), cigarettes per day (grouped as 5 per day bands) and FEV_1/FVC (grouped in 5% bands), independently contributed to lung cancer risk (Figure 9.3).

Figure 9.3 Multiple logistic regression analysis for risk of lung cancer according to genotype, smoking, BMI, lung function and presence of COPD (GOLD 2-4)



In a mediation analysis we confirmed that smoking (50+ pack years), COPD (GOLD 2-4) and the *CHRNA5* AA genotype all contributed independently to the risk of lung cancer (Figure 9.4). *CHRNA5* AA genotype contributed directly (independently) 9.0% of the effect (P=0.03).

Figure 9.4 Mediation analysis comparing the relative contributions of rs16969968 AA genotype, pack years and presence of COPD (GOLD 2-4) on the risk of developing lung cancer (OR=Odds ratio, CI= Confidence Interval)

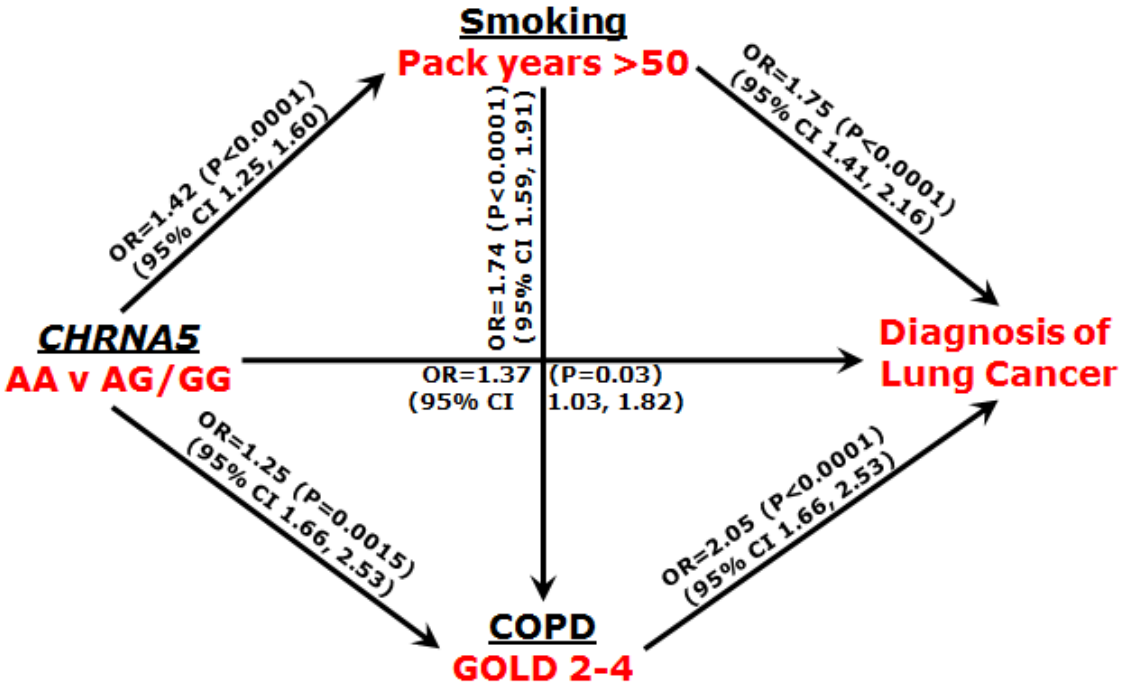


Table 9.1 Non-Hispanic Whites: *CHRNA* rs16969968 genotypes - baseline demographic and co-morbidity data according to Allelic and Recessive regression models

CHRNA rs16969968 N= 9,270[#]	AA	AG	GG	Allelic P value	Recessive P value
Genotype (% total)	1153 (12.4%)	4258 (45.9%)	3859 (41.6%)	-	-
Demographics					
Age (Years)	61.6 ± 5.0	61.8 ± 5.1	61.8 ± 5.1	0.18	0.17
Male	655 (56.8%)	2416 (56.7%)	2198 (57.0%)	0.98	0.98
Family history of Lung Cancer	276 (23.9%)	1049 (24.6%)	881 (22.8%)	0.16	0.91
Self-report COPD (Composite)	255 (22.1%)	897 (21.1%)	752 (19.5%)	0.08	0.16
Smoking History					
Current Smoker	561 (48.7%)	2061 (48.4%)	1839 (47.6%)	0.74	0.70
Pack Years	59.7 ± 25.4	56.6 ± 23.3	54.7 ± 22.5	<0.001	<0.001
Cigarettes per day	30.0 ± 11.8	28.4 ± 11.1	27.5 ± 10.8	<0.001	<0.001
Years Quit	3.6 ± 5.0	3.8 ± 5.1	3.7 ± 5.1	0.86	0.60
Smoking duration (years)	40.2 ± 7.5	40.4 ± 7.4	40.3 ± 7.5	0.69	0.58
Body composition					
Body Mass Index (BMI) (kg/m^2)	27.7 ± 5.0	27.8 ± 5.1	28.0 ± 5.1	0.52	0.31
Weight (kg)	82.0 ± 18.0	82.4 ± 17.7	83.0 ± 18.1	0.45	0.23
Height (cm)	171.6 ± 9.9	171.8 ± 9.9	172.0 ± 9.9	0.89	0.34
Education Level					
High school or less	329 (28.5%)	1185 (27.8%)	1127 (29.2%)	0.27	0.16
Post High School / Some College	389 (33.7%)	1549 (36.4%)	1361 (35.3%)		
College grad/ Postgrad/Professional	412 (35.7%)	1407 (33.0%)	1274 (33.0%)		
Other/unknown	23 (2.0%)	117 (2.7%)	97 (2.5%)		
Pre morbid Disease (self-report)					
COPD	106 (9.2%)	338 (7.9%)	230 (6.0%)	<0.001	0.007
Chronic Bronchitis	123(10.7%)	469 (11.0%)	437(11.3%)	0.80	0.62
Emphysema	116(10.1%)	409 (9.6%)	328 (8.5%)	0.13	0.28
Adult Asthma	85 (7.4%)	298 (7.0%)	253 (6.6%)	0.56	0.46
Pneumonia	336 (29.1%)	1190 (27.9%)	1074 (27.8%)	0.67	0.38
Heart Disease	152 (13.2%)	578 (13.6%)	511 (13.2%)	0.89	0.83
Hypertension	407 (35.3%)	1534 (36.0%)	1365 (35.4%)	0.80	0.78
Stroke	31 (2.7%)	123 (2.9%)	119 (3.1%)	0.75	0.58
Diabetes	111 (9.6%)	379 (8.9%)	354 (9.2%)	0.74	0.51
Any cancer History	54 (4.7%)	181 (4.2%)	151 (3.9%)	0.48	0.35

minus subjects who failed genotyping N= 97 (1%). Allelic AA vs AG vs GG. Recessive AA vs AG/GG. Recessive AA vs AG/GG. CI=Confidence Interval. Composite COPD= "yes" to doctor diagnosed 'COPD', or 'emphysema' or chronic bronchitis' or adult asthma'.

Table 9.2 Non-Hispanic Whites: *CHRNA* rs16969968 relationship with lung function and COPD status according to Allelic, Recessive and Genotype models using logistic regression

Baseline lung function, GOLD grade airflow limitation and airway phenotypes

ACRIN Non-Hispanic Whites					
CHRNA rs16969968	AA	AG	GG	Allelic Model	Recessive Model
Genotype N=9,270 (% total cohort)	N=1153 (12.4%)	N=4,258 (45.9%)	N=3,859 (41.6%)	P value	P value
Lung Function (baseline)					
FEV1/FVC (%, mean ± SD)	70.1 ± 11.0	70.6 ± 10.9	71.7 ± 10.4	<0.001[a]	0.014[a]
FEV1 % predicted (mean ± SD)	79.7 ± 20.9	80.7 ± 20.0	82.5 ± 19.8	<0.001[a]	0.005[a]
FVC % predicted (mean ± SD)	86.0 ± 18.4	86.6 ± 17.3	87.2 ± 17.1	0.023[a]	0.072[a]
Airflow Limitation (GOLD Grade) [b]					
GOLD 1, N=814	94 (8.15%)	405 (9.51%)	315 (8.16%)	0.062	0.95
GOLD 1-4, N=3,186	433 (37.55%)	1,541 (36.19%)	1,212 (31.41%)	<0.001	0.008
GOLD 2-4 (FEV1 ≤80%), N=2,372	339 (29.40%)	1136 (26.68%)	897 (23.24%)	<0.001	<0.001
GOLD 3-4 (FEV1 ≤50%), N=608	91 (7.89%)	296 (6.95%)	221 (5.73%)	<0.001	0.017
COPD GOLD2-4 – N (%by genotype)					
OR unadjusted (95% CI) P value	1.41 (1.21,1.64) P<0.001	1.23 (1.11,1.36) P<0.001	Ref	1.20 (1.12,1.29) P<0.001	1.26 (1.10,1.45) P<0.001
OR adjusted[c] (95% CI) P value	1.32 (1.13,1.54) P<0.001	1.20 (1.08,1.33) P=0.001	Ref	1.16 (1.08,1.25) P<0.001	1.19 (1.04,1.38) P=0.014
No Airflow Limitation (AL)					
No AL total N=5,927 (%)	692 (60.02%)	2655 (62.35%)	2580 (66.86%)	-	-
- Restrictive Spirometry N=1,460	172 (11.78%)	675 (46.23%)	613 (41.99%)	0.280 [b]	0.885 [b]
- Resistant smokers N=4,467	520 (11.64%)	1980 (44.33%)	1967 (44.03%)		

AL= Airflow Limitation. No Spirometry for 152 subjects. OR=odds ratio. Pers yrs=person years. [a]P value adjusted for pack years, age, and gender. [b] comparison is for restrictive spirometry vs resistant smokers within the no AL group. [c]OR adjusted for pack years (see Figure 9.2).

Table 9.3 Non-Hispanic Whites: *CHRNA* rs16969968 relationship with lung function and COPD status according to Allelic, Recessive and Genotype models using logistic regression

Unadjusted and adjusted effect estimates for association with baseline lung function

MODEL	Adjustment	FEV_1/FVC, % Estimate (95% CI)	P-value	FEV_1 % predicted Estimate (95% CI)	P-value	FVC, % predicted Estimate (95% CI)	P-value
ALLELIC	None	-0.80 (-1.08, -0.51)	<0.001	-1.64 (-2.24, -1.05)	<0.001	-0.73 (-1.21, -0.25)	0.003
	Pack years	-0.60 (-0.88, -0.31)	<0.001	-1.24 (-1.83, -0.65)	<0.001	-0.49 (-0.96, -0.02)	0.042
	Pack years, age, sex	-0.67 (-0.95, -0.39)	<0.001	-1.31 (-1.89, -0.73)	<0.001	-0.55 (-1.02, -0.08)	0.023
RECESSIVE	None	-0.91 (-1.50, -0.32)	0.003	-2.20 (-3.42, -0.98)	<0.001	-1.17 (-2.12, -0.19)	0.019
	Pack years	-0.60 (-1.18, -0.02)	0.044	-1.55 (-2.75, -0.35)	0.011	-0.76 (-1.73, 0.21)	0.124
	Pack years, age, sex	-0.72 (-1.29, -0.15)	0.014	-1.71 (-2.91, -0.52)	0.005	-0.89 (-1.85, 0.08)	0.072

Allelic = number of A alleles (0=GG, 1=AG, 2=AA), Recessive = AA vs AG/GG, 95% CI= 95% Confidence Interval.

Table 9.4 Models of best fit using the Likelihood Ratio Test (LR Test)

(a) Odds of COPD (GOLD 2-4) (cross-sectional, N=8,299) compared to controls ("No Airflow limitation") from logistic regression models

Model	Type of model	group	Odds Ratio (95% CI)	p-value	AIC	BIC	LR test p-value
Unadjusted	Genotype (general)	AA vs GG	1.41 (1.21, 1.64)	<0.001	9911.6	9932.6	Reference
		AG vs GG	1.23 (1.11, 1.36)	<0.001			
	Recessive	AA vs AG/GG	1.26 (1.10, 1.45)	<0.001	9925.2	9939.3	<0.0001
	Allelic (log additive)	Per A allele*	1.20 (1.12, 1.29)	<0.001	9910.1	9924.1	0.490
Adjusted for pack years	Genotype (general)	AA vs GG	1.32 (1.13, 1.54)	<0.001	9750.7	9778.8	Reference
		AG vs GG	1.20 (1.08, 1.33)	0.001			
	Recessive	AA vs AG/GG	1.19 (1.04, 1.38)	0.014	9760.8	9781.9	0.0005
	Allelic (log additive)	Per A allele*	1.16 (1.08, 1.25)	<0.001	9749.5	9770.6	0.387

*coded as 0=GG 1=AG 2=AA

(b) Hazard of lung cancer diagnosis (prospective, N=9,270) from Cox proportional hazards models

Model	Type of model	group	Hazard Ratio (95% CI)	p-value	AIC	BIC	LR test p-value
Unadjusted	Genotype (general)	AA vs GG	1.48 (1.09, 2.01)	0.011	6736.9	6751.1	Reference
		AG vs GG	1.17 (0.93, 1.46)	0.177			
	Recessive	AA vs AG/GG	1.36 (1.03, 1.80)	0.028	6736.7	6743.8	0.175
	Allelic (log additive)	Per A allele*	1.21 (1.04, 1.40)	0.012	6735.0	6742.1	0.692
Adjusted for pack years	Genotype (general)	AA vs GG	1.39 (1.02, 1.88)	0.034	6708.6	6730.0	Reference
		AG vs GG	1.14 (0.91, 1.43)	0.243			
	Recessive	AA vs AG/GG	1.29 (0.98, 1.70)	0.070	6708.0	6722.3	0.242
	Allelic (log additive)	Per A allele*	1.17 (1.01, 1.36)	0.035	6706.7	6721.0	0.770

*coded as 0=GG 1=AG 2=AA

Legend: AIC=Akaike Information Criterion, BIC=Bayesian Information Criterion.

Table 9.5 Non-Hispanic Whites: *CHRNA* rs16969968 genotypes and Lung Cancer (LC) outcomes according to Allelic, Recessive and Genotype models using Cox proportional hazard models

ACRIN Non-Hispanic Whites					
CHRNA rs16969968	**AA**	**AG**	**GG**	**Allelic Model P value**	**Recessive Model P value**
Lung Cancer Diagnosed n=373	60	176	137		
Prevalence per/100 screened	5.2%	4.1%	3.6%	0.013	0.029
Incidence/1000 person years (95% CI)	8.5 (6.6,10.9)	6.7 (5.8,7.8)	5.8 (4.9,6.8)	-	-
HR unadjusted (95% CI)	1.48 (1.09,2.01)	1.17 (0.93,1.46)	ref	1.21 (1.04,1.40)	1.36 (1.03,1.80)
P value	0.011	0.177		0.012	0.028
HR adjusted (95% CI)‡	1.39 (1.02,1.88)	1.14 (0.91,1.43)	ref	1.17 (1.01,1.36)	1.29 (0.98,1.70)
P value	0.034	0.243		0.035	0.070
Time to LC diagnosis or end of follow up - mean years (SD)	6.02 (1.33)	6.08 (1.25)	6.09 (1.24)	-	-
Lung Cancer according to COPD status					
- With COPD (GOLD 2-4)	23 (15.2%)	75 (49.7%)	53 (35.1%)	0.630	0.793
- Without Airflow limitation	29 (16.3%)	78 (43.8%)	71 (39.9%)		
Lung cancer Histology (N=% grp LCdx)					
1 Small Cell	7 (11.7%)	25 (14.2%)	17 (12.4%)	0.60	0.26
2 Squamous Cell	16 (26.7%)	38 (21.6%)	34 (24.8%)		
3 Adenocarcinoma	18 (30.0%)	64 (36.4%)	51 (37.2%)		
4 BAC	8 (13.3%)	11 (6.25%)	15 (10.9%)		
5 Large Cell	3 (5.0%)	2 (1.1%)	2 (1.5%)		
6 Non-small Cell	8 (13.3%)	36 (20.5%)	17 (12.4%)		
7 Other	-	-	1 (0.73%)		
Lung cancer Stage at Dx (N=% grp LCdx)					
1 Stage I	35 (58.3%)	72 (40.9%)	57 (41.6%)	0.27	**0.04**
2 Stage II	-	10 (5.7%)	8 (5.8%)		
3 Stage III	10 (16.7%)	36 (20.5%)	31 (22.6%)		
4 Stage IV	14 (23.3%)	56 (31.8%)	37 (22.6%)		
5 Occult carcinoma/ unk	1 (1.7%)	2 (1.1%)	4 (2.9%)		
Surgery					
Surgery LC - Yes (N=% total LCdx)	36 (60%)	84 (47.7%)	77 (56.2%)	0.96	0.22
Screening Interval (N=% grp LCdx)					
T0-T2	40 (66.7%)	105 (59.7%)	87 (63.5%)	0.89	0.47
T3-T6	20 (33.3%)	71 (40.3%)	50 (36.5%)		
Lung Cancer detection (N=% grp LCdx)					
Screen-detected	32 (53.3%)	85 (48.3%)	75 (54.7%)	0.79	0.49
Missed	5 (8.3%)	16 (9.1%)	10 (7.3%)		
Interval	3 (5%)	4 (2.3%)	2 (1.5%)		
Follow up	20 (33.3%)	71 (40.3%)	50 (36.5%)		
Other Demographics of LCdx					
% Male	25 (41.6%)	104 (59.1%)	77 (56.2%)	0.16	**0.02**
% Current smoker	34 (56.7%)	106 (60.2%)	72 (52.6%)	0.39	0.98
Family history of Lung Cancer – (yes)	13 (21.7%)	49 (27.8%)	35 (25.5%)	0.74	0.40
Self-reported COPD (composite)	21 (35%)	64 (36%)	31 (23%)	**0.03**	0.54
Randomised CT	30 (50%)	97 (55.1%)	72 (52.6%)	0.89	0.58

HR=Hazard Ratio, Pers yrs=person years. [a]Model adjusted for pack years. ‡ See Figure 9.2.

Table 9.6 Non-Hispanic Whites only: *CHRNA* rs16969968 genotype frequencies according to airflow limitation, lung cancer (LC), lung cancer sub-phenotypes according to the Allelic and Recessive models using logic regression

ACRIN Non-Hispanic Whites only (N=9,270)					
CHRNA rs16969968	AA	AG	GG	Allelic P-value*	Recessive P-value#
Genotype N=9,270 (% total)	N=1,153 (12.4%)	N=4,258 (45.9%)	N=3,859 (41.6%)	-	-
No Air Flow Limitation (AL)					
No AL, N=5,921 (% total)	692 (11.7%)	2,653 (44.8%)	2,576 (43.5%)	-	-
No AL and no LC, N=5,749 (% total) ‡ (referent)	663 (11.5%)	2,577 (44.8%)	2,509 (43.6)	-	-
% COPD (FEV_1/FVC<0.70,)					
GOLD 1, N=814 (% total)	94 (11.5%)	405 (49.8%)	315 (38.7%)	0.101	0.763
GOLD 2-4 (FEV_1 ≤80%), N=2372 (% total)	339 (14.3%)	1136 (47.9%)	897 (37.8%)	<0.001	0.012
Lung Cancer					
Total Lung cancer diagnosis, N=373 (% total)	60 (16.1%)	176 (47.2%)	137 (36.7%)	0.008	0.033
Squamous Cell lung cancer, N=88 (% total)	16 (18.2%)	38 (43.2%)	34 (38.6%)	0.207	0.126
Adenocarcinoma lung cancer, N=133 (% total)	18 (13.5%)	64 (48.1%)	51 (38.3%)	0.356	0.688
Lung Cancer and COPD status					
Lung cancer and GOLD 2-4, N=151 (% total)	23 (15.2%)	75 (49.7%)	53 (35.1%)	0.077	0.354
Lung cancer and No AL, N=178 (% total)	29 (16.3%)	78 (43.8%)	71 (39.9%)	0.163	0.092

#AA vs AG/GG and * AA vs AG vs GG, AL=Airflow limitation

‡ No AL "healthy smokers" with no LC diagnosis during follow-up is the reference group for all subsequent comparisons. All analyses are adjusted for pack years.

9.4 Discussion

This study confirms our earlier findings that the *CHRNA5* polymorphism (rs16969968) is independently associated with smoking, the presence of airflow limitation (COPD) and lung cancer.[11] We found a dose-response relationship between the *CHRNA5* A allele and AA genotype with pack years, cigarettes per day, lung function (%predFEV_1) and airflow limitation (FEV_1/FVC), with a log-additive allelic model providing the best fit for COPD. We also found evidence for a linear relationship between the A allele and self-reported COPD, the presence of COPD defined by spirometry and the incidence of lung cancer. While the AA genotype (recessive model) was also associated with smoking history, lung function, airflow limitation (COPD) and lung cancer incidence, the allelic and recessive models were a comparable fit for lung cancer. These findings suggest the *CHRNA5* locus is not only associated with smoking exposure but links COPD and lung cancer at a molecular genetic level. Such a hypothesis has been debated for over 40 years based on early genetic epidemiological studies and raises a number of important issues.[7,33]

First the genetic overlap between COPD and lung cancer supports the hypothesis that these diseases share pathogenic pathways.[35] This finding has been made possible by measuring lung function routinely across high risk smokers followed prospectively in a large cohort study.[90] Such an approach is unique in lung cancer genetics where most studies are cross-sectional case-control in design and make no provision for the mediating effects of co-existing airflow limitation.[165-167,311] While large studies confirm the association between this locus and lung cancer (or smoking exposure), the mediating role of COPD has not been assessed in the same study population.[311] The rs16969968 polymorphism has been shown to have functional effects on the activity of the nicotinic acetylcholine receptor found in both the respiratory epithelium where it modulates inflammatory pathways, and in the brain where it modulates addiction pathways (Figure 1.2 and Figure 1.3).[306,307,312] Nicotine replaces acetylcholine in receptor activation indicating that nicotine itself has a pathogenic role in COPD and lung cancer, not just isolated to promoting addiction to cigarette smoking. This is relevant because there is a growing body of literature suggesting exposure to inhaled nicotine, such as from vaping, is directly pathogenic in the lungs.[313,314] It also has relevance in the regulation of nicotine levels in its various inhaled forms. The findings of this study fuel further concerns, about the long-term dangers of chronic nicotine inhalation.

Second, this study highlights the close relationship between COPD and lung cancer from a genetic epidemiological perspective. Given the well-publicised genetic studies of lung cancer make little provision for co-existing COPD,[165-167] there is the possibility that some of the reported associations and subsequent genetic modelling, result from a confounding or mediating effect from the disproportional presence of COPD in the lung cancer cases (versus controls), in these retrospective studies. Furthermore, we suggest that genetic models may be different in lung cancer according to the presence or absence of COPD. We have demonstrated this previously with regards to the Glutathione S-transferase M (*GSTM1*) polymorphism linked to lung cancer in a large meta-analysis but shown to be confounded by unrecognised COPD.[153] This close relationship has relevance to the use of genetic data in risk-based approaches to targeted interventions such as CT screening for lung cancer.[241,315] We have shown that those with normal lung function or only mild COPD gain the most from lung cancer screening.[102] In contrast, those with severe or very severe COPD achieve little or no reduction in lung cancer death from screening.[102] This is because COPD is associated with more aggressive lung cancer, more late stage disease, lower surgical rates, and higher rates of deaths from cardiovascular and respiratory causes.[102,113]

Third, understanding the genetic basis of lung cancer requires some consideration of the contributory effects of the genetic basis of COPD. Indeed, the heritability of COPD, estimated to be 50-70%, is 2-3-fold greater than for lung cancer, estimated to be about 20-30%. Between 50-80% of smokers with lung cancer have COPD depending on how the latter is defined.[11,35,308] This means that the genetic basis of lung cancer may include a large contribution from a genetic overlap with the tendency for COPD. The only way to successfully tease out the genetic basis of lung cancer is to account for the presence of COPD and its contribution to risk, and genetic modelling, as we have done in this study. This is only possible where lung function data is available in the same cohort of people getting lung cancer. The current study is the only cohort study of this type that we are aware of. In a case-control study VanderWeele and colleagues used mediation analysis to confirm this *CHRNA*5 locus was linked to lung cancer independent of smoking exposure.[316] Not only do our mediation results concur with the VanderWeele study, we have expanded their findings by showing the independent pathways linking this locus with lung cancer may in part be related to pathogenetic pathways underlying COPD.[317,318] This is important, as suggested

above, when it comes to using genetic data in assessing risk of lung cancer for targeted interventions such as screening.[102,315] We have shown here that the *CHRNA5* polymorphism (rs16969968) mediated the risk for both COPD and lung cancer independently, albeit by different genetic models (allelic and recessive respectively). We have also previously found that some genetic variants confer risk of COPD only, lung cancer in the absence of COPD[308,154] and greater lung cancer lethality.[319] The latter is of particular relevance because smokers with normal lung function, or only mild COPD, achieve significantly better reductions in lung cancer deaths with CT-based screening relative to CXR screening in the NLST.[102]

9.4.1 Study strengths and limitations

The current study has several strengths and limitations. First, the data linking the *CHRNA5* (rs16969968) polymorphism to airflow limitation is based on cross-sectional data from a cohort of heavy smokers rather than prospective data from a population more representative of the wider non-smoking and smoking community e.g. the HUNT study.[304] That said, while the study population we used was a smoking cohort of high risk heavy smokers, it is representative of screening populations and provides a more stringent test of the smoking-by-gene interaction. This contrasts with previous case-control studies that included lighter smokers of lower overall risk to identify comparable associations.[12,165-167,302-303,311] Second, we have only investigated one polymorphic locus and not considered the relevance of other variants in this chromosomal region (e.g. iron-responsive element-binding protein 2, *IREB*2).[320] That the rs16969968 polymorphism has been shown to have functional effects on gene expression and protein function,[306-307,317-318] and a consistent relationship between the presence of the A allele and increased risk, supports the possibility that this is indeed an important disease-causing polymorphism. Further studies will be required to examine possible epistatic effects with neighbouring or distant variants including from other candidate genes (e.g., *IREB2*, *PAMA5*). While residual confounding and incomplete modelling might remain an issue in our study, we have at least attempted to account for the complex interplay between smoking exposure, COPD, and lung cancer through our mediation analysis. Third due to relatively small sample sizes, type 1 error (false positive result) remains a possible explanation for our findings. However, the consistency of the association between this variant and lung cancer, and consistency in the magnitude of the effect size with other

larger studies, is reassuring. We have confirmed the association using both a genotype (Table 9.5) and phenotype (Table 9.6) approach. Fourth, McKay and colleagues recently reported that in a large case-control study, this variant is associated specifically to the squamous cell histological subgroup of lung cancer.[311] While our study is under-powered to explore subgroup associations, our analysis supports this finding (Table 9.6). However, the finding by McKay might be explained through mediating or confounding effects of COPD.

A major strength of this current study is our validation of a molecular marker for lung cancer risk in a population for whom assessing lung cancer risk has clinical relevance, in this case for lung cancer screening and its outcomes.[321]

9.4.2 Summary

In summary, we have used a prospective cohort study to confirm that the A allele and/or AA genotype of the *CHRNA5* (rs16969968) polymorphism is independently associated with smoking exposure, presence of significant airflow limitation and risk of developing lung cancer. The study design we have used is unique in confirming this complex inter-relationship and highlights the need to consider the mediating effect of COPD in lung cancer genetics. Our findings support those of others suggesting the nicotine receptor, and by inference inhaled nicotine exposure, may have direct effects on the development of COPD and lung cancer. Our findings confirm our previous hypothesis that COPD and lung cancer are linked at a molecular genetic level.

Chapter 10 Discussion and Concluding Remarks

'The truth is incontrovertible.
Malice may attack it, ignorance may deride it, but in the end, there it is.'
– Winston Churchill

10.1 Introduction (Chapter 1)

The overall theme of this thesis is, "in what ways does airway limitation (COPD) effect lung cancer development, diagnosis and outcome." Each chapter examines specific aspects related to this overall theme. To explore the COPD-lung cancer relationship I have used a two data sets involving N=18,640 participants (Chapter 3), and N=10,054 participants (Chapters 4-6), from the ACRIN sites of the NLST. This data represents a subgroup from a large randomised, prospective, population-based lung cancer screening programme of 53,200 high risk smokers – (the ACRIN-NLST-NZ cohort). I have done an in-depth preliminary interrogation of this data-base which has led to some novel and important findings. A number of these findings have been presented as abstracts at conferences and consequently full papers have been published as referenced and acknowledged.

The results of the studies included in this thesis suggest that an outcomes-based approach may be more effective in CT screening programmes rather than continuing to use the risk-based approach centred mainly on age and smoking history. Such an approach represents a paradigm shift in CT-based lung cancer screening by targeting those who would get the greatest benefit from screening, rather than simply those at greatest risk of getting lung cancer (i.e. the intermediate risk group where I show there is a sweet spot for maximising the benefits over the harms). This would also represent a paradigm shift in cancer screening overall, reflecting the very important role that comorbid COPD (airflow limitation) has by increasing the risk of lung cancer but also decreasing the benefit due to worse outcomes in this high risk group.

Although the risk factors for lung cancer are known and generally universally agreed, there is no international consensus on what national screening programmes should look like.

Indeed, within the USA many regional and/or institutional CT screening programmes have been introduced with varying degrees of enrolment success.

Within the group at risk for lung cancer, taking into consideration their coexisting airflow limitation would better identify those most likely to cope with lung cancer treatment and therefore be more likely to benefit from screening. We previously proposed this 'targeted' screening approach (ref 97, 99) and this thesis provides evidence of why this approach is important, (Chapters 3-9).

The individual chapter synopses have been summarised in the introduction to this thesis. As declared at the end of the introduction, (Chapter 1), this is a thesis with publication as permitted by the University of Auckland. Where chapters represent the published work, the citation for the publication is indicated on the front page.

10.2 Methods (Chapter 2)

This chapter outlines the different epidemiological methods I used to explore the COPD – lung cancer relationship.

A strength of this thesis is the use of a variety of methods and study design which has enabled me to unravel novel associations with regards to the development and outcomes in lung cancer in both screening and unscreened cases.

Included are four studies based on a post-hoc analysis of a subgroup from the NLST, a large prospective, randomised study of high risk smokers undergoing annual CT or CXR screening for lung cancer followed for an average of 6.4 years, where outcomes are notably different, (Chapters 3-6).

I also undertook three further studies –

(a) a systematic review of outcomes after surgery in unscreened and screened cases, (Chapter 7).

(b) a case-case comparative study based on a retrospective comparison between European and Māori lung cancer cases in the Auckland region of NZ, (Chapter 8).

(c) a genetic validation study based on the NLST to replicate an earlier NZ case-control study we had published in 2009 (Chapter 9).

Table 2.1 gives an overall outline of the specific studies including the various statistical approaches used. Throughout I have had the guidance and help of an experienced statistician.

10.3 Reduced expiratory flow rate among heavy smokers increases lung cancer risk: results from the NLST-ACRIN cohort (Chapter 3)

I examined the relationship of airflow limitation and lung cancer risk, using a large prospective study of older heavy smokers (N=18,473) from the American College of Radiology Imaging Network sub-cohort of the National Lung Screening Trial (NLST-ACRIN). Participants were followed for a mean of 6.4 years, and airflow limitation was defined by pre-bronchodilator spirometry sub-grouped according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) grade 1-4.

I found 35 % (N=6,436) had airflow limitation and 65% (N= 12,037) had no airflow limitation. In those with no airflow limitation, i.e., the referent group, lung cancer incidence was 3.78/1000 person years (pyrs). In those with airflow limitation lung cancer incidence rates increased in a simple linear relationship: GOLD 1 (6.27/1000 pyrs), GOLD 2 (7.86/1000 pyrs), GOLD 3 (10.71/1000 pyrs) and GOLD 4 (13.25/1000 pyrs) - all $P \leq 0.0001$ vs the referent group. In this large prospective study, the results show there was a strong linear relationship between increasing severity of airflow limitation and increasing lung cancer risk. Moreover, the analysis showed that the effect a reduced FEV_1 (<80% predicted) had on lung cancer risk was greater than for age and pack years in this high risk group. This replicates the findings of Burrows et al in 1977[2]. This prospective study showed that with increasing airflow limitation the risk of lung cancer increases. This finding also supports previous studies showing a similar linear relationship. However, in the subsequent chapters, I report that there exists discordance between risk of lung cancer and benefit from screening and that COPD may in part underlie this observation.

The implications from Chapter 3 are that airflow limitation (AFL) affects 1/3rd of high risk smokers eligible for lung cancer screening in the NLST. Consistent with other studies 65% of these were undiagnosed. In my study AFL affected just over ½ of all those who got lung cancer. People who are smokers and have AFL have a higher risk of lung cancer. This

prospective study showed that with increasing airflow limitation the risk of lung cancer increases.

Implications for change of practice could include performing routine portable spirometry as quantifying airflow limitation identifies those at risk of lung cancer, and those who gain the most from screening.. For those targeting high risk smokers, spirometry helps further define the target groups for CT screening. This represents a paradigm shift by suggesting that spirometry should be routinely undertaken in all high risk smokers not just those presenting with symptoms suggestive of COPD. Lung cancer screening has been shown to reduce lung cancer mortality and as such is an evidence based intervention for smokers that can be further optimised by the wider use of spirometry.

Next steps from Chapter 3

Given the findings reported in Chapter 3, confirming that decreased airflow limitation increases lung cancer risk, the next steps would include further testing to better delineate the airflow limitation vs emphysema effect on lung cancer risk. For example, by looking at the relative contributions of DLCO vs Airflow limitation vs emphysema scored from CT. Further, as stated on page 84 we had insufficient numbers to examine this effect in minority groups so another next step would be to examine this FEV1 - lung cancer relationship in other ethnicities. As noted above I recommend that spirometry be routinely done in all lung cancer screening participants to better identify its presence and severity given the complex relationship between airflow limitation (COPD) and lung cancer outcomes.

10.4 NLST Phenotypes airway disease and airflow limitation: effect on outcomes in high risk smokers screened for lung cancer (Chapter 4)

Self-reported airways disease and spirometry defined airflow limitation (COPD) are both associated with an increased risk for a lung cancer diagnosis. In this study four distinct subgroups were identified based on self-reported history of respiratory disease and the lung function test result. The defining characteristic denoting which group subjects were ultimately assigned to was whether the spirometry measurement showed airflow limitation with or without self-reported airways disease. Using this approach, I defined a control (referent) group consisting of subjects who had neither airflow limitation nor reported airways disease (Group1). The remaining subjects consisted of those who reported airways

disease but without measured airflow limitation (Group 2), those who did not report airways disease but the spirometry measurement showed airflow limitation (Group 3), and lastly subjects with both airflow limitation and self-reported airways disease (Group 4). Using this phenotype approach, it was possible to identify the smokers at greatest risk of lung cancer and those who get the most value from CT screening. In this study a self-reported history of doctor diagnosed airways disease seems to be associated with worse outcomes reducing the expected benefit from CT screening which is in contrast to those who reported no airways disease but have undiagnosed COPD on spirometry testing. This analysis showed that nearly 90% of the lung cancer deaths averted with CT screening come from those with no reported respiratory disease (Group 1 and Group 3). With the goal of screening being earlier detection of cancers that can be treated, an individual's spirometry results could act as both a guide to individualise risk and screening benefits within a lung cancer screening programme. The findings of this study support the use of spirometry screening in asymptomatic smokers entering CT lung cancer screening programmes. The implications of these findings are that spirometry provides a useful assessment of both risk and outcomes from lung cancer screening. Further implications are that this study shows those with undiagnosed COPD seem to gain more from CT screening than those with diagnosed (presumably symptomatic) COPD although the study is underpowered statistically to confirm this. In my analysis, participants with diagnosed COPD appeared to derive less benefit from screening. This was associated with increased non-lung cancer deaths and less surgery. The airway diseases group (the symptomatic), had increased lung cancer risk and decreased benefit although the numbers are small and the study is underpowered to explain this further.

The study is underpowered to change clinical practice so we are undertaking the same analysis in the bigger cohort of 18,000 NLST participants.

Next steps from Chapter 4

Given the findings from Chapter 4 next steps would include examining why the airway disease group do badly despite having normal lung function. For example, further research to see if this effect may be related to ILD, or obesity, or chronic bronchitis which were over-represented in this group. Similar to my recommendations from chapter 3 (10.3 above), it

appears that a prior diagnosis of COPD, together with an assessment of airflow limitation be routinely included in lung cancer screening.

10.5 The effect of airflow limitation and lung cancer risk on lung cancer screening outcomes: all cause and lung cancer specific mortality (Chapter 5)

In the era of personalised medicine, it is prudent to look at the impact comorbid disease may have on outcomes in those at risk of lung cancer. This is especially important in screening programmes where the benefits of screening maybe tempered or attenuated by pre-existing comorbid disease. The aim of this study was to examine lung cancer screening outcomes in the context of comorbid COPD. Specifically, I was interested to see if there were possible associations and relationships between COPD severity, all-cause and cause-specific mortality.

The main finding of this study showed that in a screened population, as the risk of lung cancer increases according to lung function there is a separation (divergence) between deaths due to lung cancer and deaths due to a non-lung cancer cause. This separation happens to correlate nicely with an increase in COPD severity which in turn, correlates with a decrease in surgery and decreasing benefit from CT screening, notably for those of GOLD grade 3 and 4 severities. These results suggest airflow limitation as defined by spirometry has an effect on lung cancer risk, surgical treatment and lung cancer death.

A further finding from this study was that more deaths were attributed to cardio-vascular and respiratory causes than deaths due to lung cancer in those with GOLD grade 3-4 disease. Another finding of note is that worsening airflow limitation confers a similar level of risk for cardiovascular death as does diabetes (with no airflow limitation).

This observation extends that from the preceding chapter by showing that severe airflow limitation is preferentially associated with greater non-lung cancer deaths from cardiovascular disease and respiratory disease. My results suggest that competing causes of death may underlie the poorer outcomes in the most high risk group. This observation, together with more aggressive cancer and reduced surgery, argues that high risk smokers with GOLD 3-4 disease might gain little from screening. We plan to test this further in the full 18,000 NLST cohort.

The implications from Chapter 5 are that these results raise concern around the issue of comorbidity in screening participants although the study was somewhat underpowered and so this result alone is insufficient to change clinical practice.

Next steps from Chapter 5

Given the finding reported in Chapter 5, next steps would include using lung cancer screening registries that have recorded pre-existing comorbidity and correlate these to surgical rates and survival results following screening. For example, use of single arm observational studies analysis of screening registries, to compare outcomes according to comorbid disease at baseline.

10.6 Airflow limitation, comorbid disease, and risk of lung cancer ($PLCO_{(M2012)}$): effects on lung cancer screening outcomes in the NLST (Chapter 6)

Unlike other cancer types (e.g., breast, prostate), lung cancer screening participants have a high prevalence of comorbid disease, due primarily to a long history of smoking exposure. This study examined the fundamental supposition in lung cancer screening, namely, that the highest risk group according to the $PLCO_{M2012}$ lung cancer risk model, get the most benefit. The results of this study suggest that for individuals in the 5th quintile, (the highest risk group), where COPD prevalence is highest, there was no lung cancer death reduction for those in the CT arm. This may be explained in part by there being little difference in surgical rates between those screened with CT vs Chest X-ray. High rates of pre-existing comorbid disease may reduce the likelihood of surgery. In this study there was a non-linear relationship between risk and benefit in CT screening. This suggests that overtreatment may be an unintended consequence of screening the highest risk (Quintile 5) participants. The most important finding is the identification of a 'sweet-spot' for screening, where the middle 20-80% risk group (Quintiles 2-4), get the most benefit from screening and should be the target group to screen. This distortion has been supported by recent papers showing high comorbid disease and reduced life expectancy undermine the benefits of screening in the highest risk quintile. Therefore, I would contend that the true effectiveness of CT screening is not the numbers diagnosed with lung cancer, but rather, identifying those for whom curative treatment is viable and likely results in prolonged survival.

In summary, this study explored the role of the $PLCO_{m2012}$ lung cancer risk tool in screening outcomes.

The important observation was that increased risk according to the $PLCO_{m2012}$ risk also increases the prevalence of COPD and again seemed to reduce the screening benefit in those at highest risk (Quintile 5). Again, this study is too underpowered to change practice. This observation is unique in showing that risk-based tools, such as the well validated $PLCO_{M2012}$, entirely overlooks this limitation of the model in selecting smokers for lung cancer screening. It also suggests that biomarkers of lung cancer biology may play an important role in better delineating who will do best from screening (see Chapter 9).

Further, this study confirmed that the relative benefit of screening was maximised in the middle group, i.e., PLCO m2012 risk quintiles 2-4. The implication of this in practice may be to review those with very high risk and to carefully weigh-up the benefits vs the harms of screening for these people.

For those referred for lung cancer CT screening the shared decision making visit should discuss the risks and benefits as well as explore the patient's preferences about screening.

Next steps from Chapter 6

The finding reported in Chapter 6, shows that in the highest risk quintile 5 group the improved survival with screening was less than for those in the middle quintiles (2-4) groups when using the $PLCO_{M2012}$ model. Next steps would again include using large screening registries to compare survival in this intermediate risk group where overall life expectancy has been shown to be better.

A next step would also be to collect data on complications of screening according to $PLCO_{M2012}$ risk quintiles.

10.7 Airflow limitation and survival after surgery for NSCLC: results from a Systematic Review and lung cancer screening trial (Chapter 7)

In this chapter I discuss a study looking at outcomes following surgery for early stage non-small cell lung cancer in non-screened lung cancer cases and compared the survival outcomes in a screened population where for both groups pre-existing spirometry defined

airflow limitation (COPD) was available. This is a two-stage study. Stage one of this study was a systematic review and meta-analysis of papers meeting stringent inclusion criteria as described (see Chapter 7, section 7.2). Stage two involved analysing data from the ACRIN-NZ subset of the National Lung Screening Trial where I looked at survival in those undergoing surgery for lung cancer.

The finding from this study is one in which for non-screened lung cancer patients undergoing surgery, those with airflow limitation had similar survival to those with no airflow limitation. A caveat to this though, was found in one study where those with the most severe grade COPD appeared not to do as well. Using actual trial data (ACRIN-NSLT NZ cohort), I have been able to confirm this finding. The corollary of this is, COPD patients with mild to moderate disease severity get value out of lung cancer screening (see also Chapter 4) consistent with our finding that individuals with undiagnosed airflow limitation do well because they get more surgery than do the individuals with pre-existing diagnosed airflow limitation (COPD). The implication of these results is that treatment choice maybe directed in part by the lung function status.

In summary, this study examined survival after lung cancer surgery in those with pre-exiting mild-moderate COPD Gold grade 1-2 severity, in an unscreened vs screened population.

This study was an opportunity to compare 5year survival after surgery for lung cancer in those with normal lungs and those with mild to moderate COPD and showed that they were the same.

The implication of these results has been clearly stated in Chapter 7, which is that treatment choice and screening decision may be directed in part by the lung function status. See section 7.4.3 (Future implications), page 159.

Next steps from Chapter 7, there are no next steps for this study. However, again lung cancer screening registries maybe used to assess screening outcomes in participants with GOLD 3-4 at baseline.

10.8 Are New Zealand Māori more susceptible to smoking related lung cancer? A comparative case-case study (Chapter 8)

Even though the total population of New Zealand is small at just over 5million, ethnically it is very diverse. New Zealand has a National Health System where everyone has access to nationally funded health care. The Ministry of Health has responsibility for overseeing funding and sets disease specific target strategies. Even so, there are differences in the health outcomes within and between ethnicities, with NZ Maori (the indigenous Polynesian group of New Zealand), seemingly more at risk of adverse health outcomes. This is a comparator study undertaken to compare lung cancer characteristics and outcomes between NZ Maori lung cancer cases and NZ Caucasian lung cancer cases, with both case cohorts identified from a single geographic region to mitigate possible biases due to regional differences. The key findings from this chapter are that although NZ Maori with lung cancer have comparable smoking exposure to Caucasian lung cancer cases, the onset of lung cancer occurs on average at a younger age in Maori, is of a more aggressive histological sub-type and their lung function is worse, i.e., they have more airflow limitation (COPD) and lower FEV_1 than their NZ Caucasian counterparts. The relevance of lower lung function and a greater susceptibility to COPD is important as it may have an impact on lung cancer outcomes. The basis of these findings remains uncertain. One implication of these findings is that if disparities are to be addressed then screening spirometry and targeted lung cancer screening starting at a younger age in Maori may improve outcomes in this high risk group.

In summary, this study compared demographics, lung function and lung cancer characteristics with survival between lung cancer cases in two ethnic groups in New Zealand. A significant key difference was at lower smoking (pack years) there was a marked decrease in lung function. Further, Maori had less adeno carcinoma and more small and non-small cell carcinoma which remained consistent across different stratifications. The results and implications for Maori are clearly stated within the chapter (8.4.5 page 185) and restated above. These show that screening spirometry and targeted lung cancer screening should start at a younger age in more vulnerable groups. Such an approach is planned in pilot studies for lung cancer screening in Maori.

Specifically, my recommendations for Maori are –

- to have spirometry as part of wider lung cancer screening to identify COPD with respect to early diagnosis, lung cancer vulnerability, aggressive cessation intervention and treatment of symptoms to improve quality of care for those with mild to moderate airflow limitation;

-to target younger smokers among Maori for lung cancer screening to recognise the younger age of onset;

-to undertake more extensive molecular typing of lung cancer in Maori given the lower rates of adenocarcinoma, such as somatically activated oncogenes for example EGFR (protein epidermal growth factor), KRAS, ALK, BRAF, for targeted chemotherapy.

Next steps from Chapter 8

The finding reported in Chapter 8 that Maori have higher lung cancer mortality than European remains of interest and requires further investigation.

A second audit from the last five years could be done although there are now 2 studies (Stevens ref 281, and mine) showing the younger age effect in Maori. However, in my study I found that clinical stage of cancer was no different between Maori and European which argues strongly against the proposition that Maori do worse because they can't (or don't) access healthcare for their symptoms. My results showing poorer lung function and more aggressive histology may be alternative explanations for these poor outcomes.

These findings further support the routine use of spirometry in all at risk smokers, regardless of symptoms (see earlier), given the widespread under-diagnosis of COPD and increased risk of lung cancer.

Recently the NZ Health Research Council (HRC), has approved funding for a lung cancer screening trial in Maori. Alongside this study funding has also been granted to enable portable spirometry to be done in consenting individuals to identify COPD and recommend appropriate follow up.

I am involved in advising and writing the COPD in lung cancer section of the grant application and will be involved in training the co-investigators in performing the spirometry measurements.

10.9 Genetic Susceptibility (Chr15q25 variant rs16969968) to COPD, lung cancer and smoking (Chapter 9)

This is a validation study of earlier work we have done looking at the association between the SNP variant rs16969968 (A/G), of the nicotinic cholinergic receptor subunit (*CHRNA5*) gene locus on chromosome 15q25 and lung cancer, airflow limitation and smoking.[11] This is a sub-study using a prospective cohort sample of 9,270 non-Hispanic white high risk smokers, followed for an average of 6.4 years. The results of this study revealed that the high risk AA genotype was associated with lower lung function, greater smoking intensity, the presence of COPD, and the development of lung cancer, regardless of COPD status. Specifically, a mediation analysis showed very nicely the independent contribution of the *CHRNA5* rs16969968 AA genotype, to smoking exposure (i.e., pack years), presence of airflow limitation (COPD GOLD2-4 status) and the development of lung cancer. The results of this study have been presented in abstract form.

Indeed, the results of this prospective study confirmed an earlier case-control study finding from our lab, that this variant was associated with airflow limitation (COPD) and lung cancer.[11] Unlike the GWAS studies reporting the association as being with smoking and lung cancer, the availability of baseline spirometry in this prospective cohort enabled me to identify a possible overlapping effect by COPD status, confirming the close relationship between COPD and lung cancer, at a molecular genetic level.

This study confirms an earlier study from my group showing that COPD may be linked to lung cancer through shared pathogenic pathways. In this case, validation of this relationship in the current study suggests that genetic susceptibility to the two smoking related diseases is overlapping.[35,154] This result also directly implicates the CHRNA receptor, and nicotine itself through binding this receptor, in lung cancer. This potentially represents a paradigm shift in current thinking about nicotine's role in the development of COPD and lung cancer beyond that of addiction only. If nicotine or its metabolites are directly pathogenic when inhaled in the lungs, then any delivery of nicotine to the lung such as vaping, may be harmful.

In summary, Chapter 9 was confirmation of an earlier study of ours (ref 11), showing COPD and lung cancer are linked at the genetic level and shows the nicotinic cholinergic receptor to be directly implicated in COPD, lung cancer and smoking. This may have an effect on the regulation and control of vaping and nicotine in the future.

All this has been clearly discussed and stated in Chapter 9 (sections 9.1, 9.3, 9.4).

My results are from post-hoc secondary analysis from the first and largest lung cancer screening RCT to publish their results. They will need verification in other large RCT's and analysis is currently underway to do this analysis in the UK Biobank cohort of over 250,000 smokers.

Next steps from Chapter 9

The finding reported in Chapter 9 confirms our early findings.

Next steps would be to conduct mechanistic studies to better understand the SNP (allele) effect on COPD/lung cancer.

This would include looking at the SNP effect via the "Nicotinic receptor" on alterations in structure to understand the overlap between COPD and lung cancer.

Subsequent to submitting this thesis and publication of this study, a paper addressing the first of my next steps (mechanistic studies) has been published by Zania Diabasana et al, Int J Molecular Sciences, 2021. In their paper they describe a study using immunostaining on lung tissue sections and showed impaired ciliogenesis[1] and altered production of inflammatory mediators. Work is also underway to combine genetic susceptibility SNPs (including that for the nicotinic receptor described here) to better characterise the risk of lung cancer according to genetic signature.

Foot note [1] Ciliogenesis is the process by which cilia – slender outgrowths on a cell's surface in eukaryotes – form; (i.e., the building of the cell's extracellular fluid mediation mechanism). Motile cilia have a beating motion and serve to move fluids, whereas non-motile, or primary, cilia receive signals from other cells or fluids, (https://www.nature.com).

10.10 Final Remarks

This thesis looked at the close relationship between airflow limitation (COPD) and lung cancer. The studies were carefully designed. The major and consistent finding from each study has confirmed the important effect COPD has on lung cancer not only in terms of elevating lung cancer risk but also in terms of the outcomes achieved from lung cancer screening. An aspect of this thesis has been the ability to examine the potential benefits versus potential harms from screening for lung cancer in a high risk population recruited according to age and smoking exposure criteria. In addition, the lung cancer screening cohort is unique in that they had undergone portable spirometry and blood sample collection for DNA extraction. The outcomes reported from each study show the importance of having an objective biomarker such as spirometry measurement to verify the presence of airflow limitation for accurate classification of COPD. The spirometry data enabled me to stratify the participants not only in terms of their self-reported comorbid status but also on the basis of the severity of airflow limitation (GOLD criteria for COPD). The advantage of this was in enabling me to look more closely at the influence COPD may have on meaningful outcomes in lung cancer. For instance, risk calculators based on modelling may tell you who is most at risk but they tell you nothing about who benefits most or for whom the incidence of lung cancer is incidental to their treatment and final cause of death. Further, the inclusion of the genetic SNP analysis again examined the COPD-lung cancer relationship, adding depth to this thesis and made it a very relevant piece of translational research.

Overall, I have unravelled the relationship of airflow limitation and lung cancer at 5 important levels. These are -

1. a molecular level as shown by the evidence presented in Chapter 9
2. an epidemiological level especially shown by the evidence presented in Chapters 3-8
3. a histological level as shown by the evidence presented in Chapters 5--8
4. a treatment level as shown by the evidence presented in Chapters 5-7
5. screening outcome level as shown by the evidence presented in Chapters 5-7

Throughout this thesis the overriding aim guiding the studies was to appreciate how improvements in lung cancer would come about through a greater understanding of the connection between smoking, chronic obstructive pulmonary disease, and lung cancer from

the biological and patient perspectives. In this way I have presented new and novel results which may aid in promoting new ways to approach the devastating consequence from smoking.

Based on my results I would make the following recommendations;

1. measure airflow limitation through routine spirometry in all high risk smokers to confirm the presence and severity of this important co-morbid disease,
2. use an outcomes-based approach to targeting lung cancer screening where this intervention averts the most lung cancer per person screened,
3. acknowledge that very high risk smokers may gain less from screening due to competing causes of death, more aggressive forms of lung cancer and lower surgical rates thereby adopting the "sweet spot" approach to screening, and
4. recognise that the genetic variation in the nicotine receptor plays an important but complex role in susceptibility to developing and dying from lung cancer, and that all inhaled nicotine products are potentially carcinogenic,
5. consider the complex relationship between COPD and lung cancer in better understanding the genetic basis of lung cancer and how the use of genetic profiling could augment existing management strategies.

I have made several novel discoveries that contribute to a greater understanding of the inter-relationship between lung cancer and chronic obstructive pulmonary disease. There are several important clinical implications of this work that merit urgent confirmation or adoption in the management of this large group of current or former smokers.

References

1. Le Courteur P, Burreson J. Napoleon's buttons: 17 molecules that changed history. 2004 Penguin Group USA) Inc., New York, NY 10014.

2. Burrows B, Knudson RJ, Cline MG, et al. Qualitative relationships between cigarette smoking and ventilatory function. *Am Rev Respir Dis* 1977; 115: 195-205.

3. Doll R, Peto R, Boreham J, et al. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ* 2004; 328: 1519.

4. Fletcher C, Peto R. The natural history of chronic airflow obstruction. *BMJ* 1977; 1: 1645-1648.

5. Centers for Disease Control and Prevention: https://www.cdc.gov/tobacco/data_statistics/sgr/index.htm (last accessed November 2019).

6. Lee PN, Forey BA, Coombs KJ. Systematic review with meta-analysis of the epidemiological evidence in the 1900s relating smoking to lung cancer. *BMC* 2012; 12(385): 1-90.

7. Cohen BH, Diamond EL, Graves CG, et al. A common familial component in lung cancer and chronic obstructive pulmonary disease. *Lancet* 1977; 2(8037): 253-526.

8. Adcock IM, Caramori G, Barnes PJ. Chronic Obstructive Pulmonary Disease and lung cancer: new molecular insights. *Respiration* 2011; 81: 265-284.

9. Caramori G, Casolari P, Cavallesco GN, et al. Mechanisms involved in lung cancer development in COPD. *The Int J Biochem & Cell Biology* 2011; 43: 1030-1044.

10. Young RP, Hopkins RJ, Gamble GD, et al. Genetic evidence linking lung cancer and COPD: a new perspective. *The Application of Clinical Genetics* 2011; 4: 99-111.

11. Young RP, Hopkins RJ, Hay BA, et al. Lung cancer gene associated with COPD:triple whammy or possible confounding effect? *Eur Respir J* 2008; 32: 1158-1164.

12. Saccone NL, Culverhouse RC, Schwantes-An T-H, et al. Multiple independent loci at chromosome 15q25.1 affect smoking quantiy: a meta-analysis and comparison with lung cancer and COPD. *PLoS Genet* 2010; 6(8): e1001053.

13. Le Marchland L, Derby KS, Murphy SE, et al. Smokers with the CHRNA lung cancer-associated variants are exposed to higher levels of nicotine equivalents and a carcinogenic tobacco-specific nitrosamine. *Cancer Res* 2008; 68(22): 9137-9140.

14. Wang J, Spitz MR, Amos CI, et al. Mediating effects of smoking and Chronic Obstructive Pulmonary Disease on the relation between the CHRNA5-A3 genetic locus and lung cancer risk. *Cancer* 2010; 116: 3458-62.

15. Hopkins RJ, Duan F, Gamble GD, ...Young RP, et al. Chr15q25 genetic variant (rs16969968) independently confers risk of lung cancer, COPD and smoking intensity in a prospective study of high risk smokers. *Thorax* 2021; 76(3): 272-280.

16. Hole DJ, Watt GCM, Davey-Smith G, et al. Impaired lung function and mortality risk in men and women: findings from the Renfrew and Paisley Prospective Population Study. *BMJ* 1996; 313: 711-715.

17. Sullivan SD, Ramsey SD, Lee TA. Economic burden of COPD. *Chest* 2000; 117(2 Suppl): 5S--9S.

18. Ostbye T, Taylor DH. Effect of smoking on Years of Healthy Life (YHL) lost among middle-aged and older Americans. *Health Serv Res* 2004; 39(3): 531-552.

19. Shavelle RM, Paculdo DR, Kush SJ, et al. Life expectancy and years of life lost in chronic obstructive pulmonary disease: Findings from the NHANES III Follow-up Study. *Int J COPD* 2009; 4: 137-148.

20. Young RP, Hopkins RJ, Eaton TE. Forced expiratory volume in one second: not just a lung function test but a marker of premature death from all causes. *Eur Respir J* 2007; 30: 616-622.

21. Jia H, Lubetkin EI. Dose-response effect of smoking status on quality-adjusted life years among U.S. adults aged 65 years and older. *J Public Health (Oxf)* 2017; 39(4): e194-e201.

22. World Health Organisation: http://www.who.int/mediacentre/factsheets/fs310/en/ WHO Top 10 causes of death paper (last accessed February 2020).

23. Murray CJL and The US Burden of Disease Collaborators. The state of US Health, 1990-2016: Burden of Diseases, Injuries, and Risk Factors Among US States. *JAMA* 2018; 319: 144-1472.

24. Albert RK, Spiro SG, Jett JR. Clinical Respiratory Medicine, Third Edition. ISBN:978-0-323-04825-52008, Mosby, Inc., an affiliate of Elsevier Inc.

25. Ministry of Health, New Zealand: https://www.health.govt.nz/nz-health-statistics < last accessed March 2020)

26. Rennard SI. COPD: overview of definitions, epidemiology, and factors influencing its development. *Chest* 1998; 113(4 Suppl): 235S-41S.

27. Siafakas NM, Vermeire P, Pride NB, et al. Optimal assessment and management of chronic obstructive pulmonary disease (COPD). The European Respiratory Society Task Force. *Eur Respir J* 1995; 8: 1398-1420.

28. Pauwels RA, Buist AS, Calverley PM, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. *Am J Respir Crit Care Med* 2001; 163: 1256-76.

29. Ministry of Health, New Zealand: https://www.health.govt.nz/our-work/preventative-health-wellness/tobacco-control

30. Yao H, Rahman I. Current concepts on the role of inflammation in COPD and lung cancer. *Curr Opin Pharmacol* 2009; 9(4): 375-383.

31. Young RP, Hopkins RJ, Marsland B. The Gut-Liver-Lung Axis: Modulation of the innate immune response and its possible role in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2016; 54(2): 161-169.

32. Young RP, Hopkins RJ. The Mevalonate Pathway and Innate Immune Hyper-Responsiveness in the Pathogenesis of COPD and Lung Cancer: Potential for Chemoprevention. *Curr Molec Pharma* 2017; 10: 46-59.

33. Tockman MS, Anthonisen NR, Wright EC, et al. Airways obstruction and the risk for lung cancer. *Ann Intern Med* 1987; 106:512-518.

34. Young RP, Hopkins RJ, Christmas T, et al. COPD prevalence is increased in lung cancer, independent of age, sex and smoking history. *Eur Respir J* 2009; 34: 380-386.

35. Young RP, Hopkins RJ. How the genetics of lung cancer may overlap with COPD. Invited review *Respirology* 2011; 16: 1047-1055.

36. El-Zein RA, Young RP, Hopkins RJ, et al. Genetic Predisposition to Chronic Obstructive Pulmonary Disease and/or Lung Cancer: Important Considerations When Evaluating Risk. *Cancer Prev Res* 2012; 5: 522-527.

37. Hopkins RJ, Young RP, Duan F, et al. Airflow limitation and lung cancer risk in a large prospective screening study (NLST-ACRIN cohort analysis, N=18,714). *Am J Respir Crit Care Med* 2016; 193: A6178.

38. Hopkins RJ, Duan F, Chilkes C, et al. Airflow limitation confers the same risk of cardiovascular death as diabetes in high risk smokers in the NLST-ACRIN sub-study. *Am J Respir Crit Care Med* 2018; 197: A3543.

39. Mannino DM, Thorn D, Swensen A, et al. Prevalence and outcomes of diabetes, hypertension and cardiovascular disease in COPD. *Eur Respir J* 2008; 32: 962-969.

40. Wannamethee SG, Shaper AG, Papacosta O, et al. Lung function and airway obstruction: associations with circulating markers of cardiac function and incident heart failure in older men – the British Regional Heart Study. *Thorax* 2016; 71: 526-534.

41. Salomaa E_R, Sallinen S, Hiekkanen H et al. Delays in the diagnosis and treatment of lung cancer. *CHEST* 2005; 128(4): 2282-2288.

42. Knight SB, Crosbie PA, Balata H, et al. Progress and prospects of early detection in lung cancer. *Open Biology* 2017; 7: 170070.

43. Malvezzi M, Carioli G, Bertuccio P, et al., European cancer mortality predictions for the year 2018 with focus on colorectal cancer. *Annals of Oncology* 2018; 29(4): 1016-1022.

44. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2019. *CA Cancer J Clin* 2019; 0: 1-28.

45. Barta JA, Powell CA, Wisnivesky JP. Review: Global Edipdemiology of lung cancer. *Annals of Global Health* 2019; 85(1): 8,1-16.

46. Bradley SH, Kennedy MPT, Neal RD. Recognising lung cancer in primary care. *Adv Ther* 2019; 36(1): 19-30.

47. Lundback B, Lindberg A, Lindstrom M, et al. Not 15 but 50% of smokers develop COPD-Report from the Obstructive Lung Disease in Northern Sweden Studies. *Respiratory Medicine* 2003; 97:115-122.

48. Hill K, Goldstein RS, Guyatt GH, et al. Prevalence and underdiagnosis of chronic obstructive pulmonary disease among patients at risk in primary care. *CMAJ* 2010; 182(7): 673-8.

49. Young RP, Hopkins RJ, Gamble GD. Clinical applications of gene-based risk prediction for lung cancer and the central role of chronic obstructive pulmonary disease. *Frontiers in Genetics/Applied Genetic Epidemiology* 2012; 3(Article 10): 1-7.

50. Hopkins RJ, Young RP, Hay BA, et al. Lung cancer risk testing enhances NRT uptake and quit rates in randomly recruited smokers offered a gene-based risk test. *Am J Respir Crit Care Med* 2012; 185: A2590.

51. Samet JM. Smoking kills: Experimental proof from the Lung Health Study. *Ann Int Med* 2005; 142(4): 299-301.

52. Skillrud DM, Offord KP, Miller RD. Higher risk of lung cancer in chronic obstructive pulmonary disease: A prospective, matched, controlled study. *Ann Int Med* 1986; 105 (4): 503-507.

53. Mannino DM, Aguayo SM, Petty TL, et al. Low lung function and incident lung cancer in the United States: Data from the First National Health and Nutrition Examination Survey follow-up. *Arch Intern Med* 2003; 163: 1475-1480.

54. Houghton AM, Mouded M, Shapiro SD. Common origins of lung cancer and COPD. *Nat Med* 2008; 14: 1023-1024.

55. Sekine Y, Katsura H, Koh E, et al. Early detection of COPD is important for lung cancer surveillance. *Eur Respir J* 2012; 39: 1230-1240.

56. Laniado-Laborin R. Smoking and Chronic Obstructive Pulmonary Disease (COPD). Parallel Epidemics of the 21st Century. *Int J Res Public Health* 2009; 6: 209-224.

57. Khan S, Fell P, James P. Smoking-related chronic obstructive pulmonary disease (COPD). *Diversity and Equality in Health and Care* 2017; 11: 267-71.

58. Rennard SI. Pathogenesis of COPD. *Clinical Cornerstone* 2003; 5(1): 11-16.

59. Rahman I, Adcock IM. Oxidative stress and redox regulation of lung inflammation in COPD. *Eur Respir J* 2006; 28: 219-242.

60. Young RP, Hopkins RJ, Eaton TE. Pharmacological actions of statins: potential utility in COPD. *Eur Respir Rev* 2009; 18: 222-232.

61. Mroz RM, Lisowski P, Tycinska A, et al. Anti-inflammatory effects of atorvastatin treatment in chronic obstructive pulmonary disease. A controlled pilot study. *J Physiol Pharmacol* 2015; 66(1): 111-128.

62. Global Initiative for Chronic Obstructive Lung Disease: (www.GOLD.org accessed February 2, 2020).

63. Snider GL. Chronic obstructive Pulmonary Disease: a Definition and Implications of Structural Determinants of Airflow Obstruction for Epidemiology. *Am Rev Respire Dis* 1989; 140: S3-S8.

64. Stenton C. The MRC breathlessness scale. *Occupational Medicine* 2008; 58: 226-227.

65. Celli BR, Cote CG, Marin JM, et al. The body-mass index, airflow obstruction, dyspnoea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med* 2004; 350: 1005-1012.

66. Weatherall M, Marsh S, Shirtcliffe P, et al. Quality of life measured by the St George's Respiratory Questionnaire and spirometry. *Eur Respir J* 2009; 33: 1025-1030.

67. Jones PW, Harding G, Berry P, et al. Development and first validation of the COPD Assesssment Test. *Eur Respir J* 2009; 34: 648-654.

68. Bidan CM, Veldsink AC, Meurs H, et al. Airway and extracellular matirx mechanics in COPD. *Frontiers in Physiology* 2015; 6(346): 1-9.

69. Burgess JK, Mauak T, Tjin G, et al.The extracellular matirx – the under-recognized element in lung disease? *J Pathology* 2016; 240: 397-409.

70. Dunhill MS. Gatter KC. Cellular heterogeneity in lung cancer. *Histopathology* 1986; 10:461-475.

71. Muscat JE, Wynder EL. Lung cancer pathology in smokers, ex-smokers and never smokers. *Cancer Letter* 1995; 1(6): 1-5.

72. Deng P, Hu C, Zhou L, et al. Clinical characteristics and prognostic significance of 92 cases of patients with primary mixed-histology lung cancer. *Molecular and Clinical Oncology* 2013; 1: 863-868.

73. de Bruin EC, McGranahan N, Swanton C. Analysis of intratumor heterogeneity unravels lung cancer evolution. *Molecular & Cellular Oncology* 2015; e985549: 1-3.

74. Zappa C, Mousa SA. Non-small cell lung cancer: current treatment and future advances. *Transl Lung Cancer Res* 2016; 5(3): 288-300.

75. Beckles MA, Spiro SG, Colice GL, et al. The physiologic evaluation of patients with lung cancer being considered for resectional surgery. *CHEST* 2003; 123: 105S-114S.

76. Schiller JH. A new standard of care for advanced lung cancer. *NEJM* 2018; 378: 2135-2137.

77. Shlomi D, Ben-Avi R, Balmor RG, et al. Screening for lung cancer: time for large-scale screening by chest computed tomography. *Eur Respir J* 2014; 44: 217-238.

78. Hopkins RJ, Ko J, Gamble GD, Young RP. Lung cancer surgery for early stage non-small cell lung cancer: Systematic Review of mortality according to the presence of COPD. *Am J Respir Crit Care Med* 2017; 195: A4887.

79. Araujo LH, Horn L, Merritt RE, et al. Chapter 69 - Cancer of the lung: Non-small cell lung cacner and small cell lung cancer. In: Niederhuber JE, Armitage JO, Doroshow JH, Kastan MB, Tepper JE, eds. *Abeloff's Clinical Oncolgy*. 6th Edition, Philadelphia, Pa: Elsevier; 2020.

80. Young RP, Duan F, Greco E, Hopkins RJ, et al. Lung cancer-specific mortality reduction with CT screening: outcomes according to airflow limitation in the ACRIN-NLST study (N=18,475). *Am J Respir Crit Care Med* 2016; 193: A6166.

81. Scotland, lung cancer awareness campaign: http://www.getcheckedearly.org/last accessed Dec 2019.

82. Crosbie PA, Balata H, Evison M, et al. Implementing lung cancer screening: baseline results from a community-based 'Lung Health Check' pilot in deprived areas of Manchester. *Thorax* 2019; 74: 405-409.

83. Bach PB, Katen MW, Thornquist MD, et al. Variation in lung cancer risk among smokers. *J Natl Cancer Inst* 2003; 95: 470-478.

84. Spitz MR, Hong WK, Amos CI, et al. A risk prediction model for prediction of lung cancer. *J Natl Cancer Inst* 2007; 99: 715-726.

85. Young RP, Hopkins RJ, Hay BA, et al. A gene-based risk score for lung cancer susceptibility in smokers and ex-smokers. *Postgrad Med J* 2009; 85: 515-524.

86. Tammemagi MC, Katki HA, Hocking WG, et al. Selection criteria for lung cancer screening. *N Eng J Med* 2013; 368: 728-736.

87. Waters EA, Taber JM, McQueen A, et al. Translating cancer risk prediction models into personalized cancer risk assessment tools: Stumbling blocks and strategies for success. *Cancer Epidemiol Biomarkers Prev* 2020; 29(12): 2389-2394.

88. Young RP, Hopkins RJ, Lam V, et al. Low-dose CT lung cancer Screening in the community: A prosepctive cohort study (REACT) incorporating a gene-based lung cancer risk test. *Am J Respir Crit Care Med* 2015; 191: A3569.

89. Scott RJ, Billings P, Gamble GD, et al. Gene-based risk tests improve engagement in lung cancer screening in 3 community studies. *Am J Respir Crit Care Med* 2020; 201: A5923.

90. National Lung Screening Trial Research Team. Writing committee Aberle D, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011; 365: 395-409.

91. Optican RJ, Chiles C. Implementing lung cancer screening in the real world: opportunity, challenges and solutions. *Transl Lung Cancer Res* 2015; 4(4): 353-364.

92. Ruchalski K, Gutierrez A, Genshaft S, et al. The evidence for low-dose CT screening of lung cancer. *Clinical Imaging* 2016; 40: 288-295.

93. Young RP, Duan F, Chiles C, et al. Airflow limitation and histology shift in the National Lung Screening Trial: The NLST-ACRIN Cohort Study. *Am J Respir Crit Care Med* 2015; 192(9): 1060-1067.

94. Wender R, Fontham ETH, Barrera E, et al. American Cancer Society Lung Cancer Screening Guidelines. *CA Cancer J Clin* 2013; 63: 106-117.

95. Ruano-Ravina A, Heleno B, Fernandez-Villar A. Lung cancer screening with low-dose CT (LDCT), or when a public health intervention is beyond the patient's benefit. *J Epidemiol Community Health* 2015; 69(2): 99-100.

96. Bach PB, Gould MK. When the average applies to no one: personalized decision making about potential benefits of lung cancer screening. *Ann Intern Med* 2012; 157: 571-573.

97. Young RP, Hopkins RJ. Measures of outcome in lung cancer screening: maximising the benefits. *J Thorac Dis* 2016; 8(10): E1317-E1320.

98. Green AK, Bach P. Model-based eligibility for lung cancer screening: where theory meets practice. *Ann Intern Med* 2018; 168(3): 223-224.

99. Young RP, Hopkins RJ. Diagnosing COPD and targeted lung cancer screening. *Eur Respir J* 2012; 40: 1063-1064.

100. Advani S, Braithwaite D. Optimizing selection of candidates for lung cancer screening: role of comorbidity, frailty and life expectancy. *Transl Lung Cancer Res* 2019; 8 (Suppl 4): S454-S459.

101. Caverly TJ, Fagerlin A, Wiener RS, et al. Comparison of Observed Harms and Expected Mortality Benefit for Persons in the Veterans Health Affairs Lung Cancer Screening Demonstration Project. *JAMA Int Med* 2018; 178(3): 426-428.

102. Young RP, Hopkins RJ. Chronic obstructive pulmonary disease (COPD) and lung cancer screening. *Transl Lung Cancer Res* 2018; 7(3): 347-360.

103. Young RP, Hopkins RJ. The potential impact of chronic obstructive pulmonary disease in lung cancer screening: implications for the screening clinic. *Expert Review Of Respiratory Medicine* 2019; 13(8): 699-707.

104. Jaklitsch MT, Jacobson FL, Austin JHM, et al. The Americn Association for Thoracic Surgery guidelines for lung cancer screening using low-dose computed tomography scans for lung cancer surviviors and other high risk groups. *J Thorac Cardiovasc Surg* 2012; 144(1): 33-38.

105. Moyer VA, on behalf of the U.S. Preventive Services Task Force. Screening for Lung cancer: U.S. Preventive Services Task Force recommendation statement. *Annals Internal Medicine* 2014; 160: 330-338.

106. Wood DE, Kazerooni E, Baum SL, et al. Lung cancer screening, version 1.2015; featured updates to the NCCN guidelines. *J Natl Ccompr Canc Netw* 2015; 13(1): 23-34.

107. Maisonneuve P, Bagnardi V, Bellomi M, et al. Lung cancer risk prediction to select smokers for screening CT – a model based on the Italian COSMOS trial. *Cancer Prev Res* 2011; 4: 1778-1789.

108. Harris R, Sawaya GF, Moyer VA, et al. Reconsidering the Criteria for evaluating proposed screening programs: Reflections from 4 current and former members of the U.S.Preventive Services Task Force. *Epidemiologic Reviews* 2011; 33: 20-35.

109. Tammemagi MC, Church TR, Hocking WG, et al. Evaluation of the lung cancer risks at which to screen ever- and never-smokers; screening rules applied to the PLCO and NLST cohorts. *PLoS Med* 2014; 11: e1001764.

110. Mazzone P, Powell CA, Arenberg D, et al. Components necessary for high-quality lung cancer screening: American College of Chest Physicians and American Thoracic Soicety Polciy Statement. *CHEST* 2015; 147(2): 295-303.

111. Callister MEJ, Baldwin DR, Akram AR, et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules. *Thorax* 2015; 70: ii1-ii54.

112. Tanner NT, Dai L, Bade BC, et al. Assessing the Generalizability of the National Lung Screening Trial: Comparison of Patients with Stage1 Disease. *Am J Respir Crit Care Med* 2017; 196(5): 602-608.

113. Rivera MP, Tanner NT, Silvestri GA, et al. Incorporating Coexisting Chronic Illness into Decisions about Patient selection for Lung cancer Screening. An Official American Thoracic Society Research Statement. *Am J Respir Crit Care Med* 2018; 198(2): e3-e13.

114. Iaccarino JM, Steiling KA, Wiener RS, et al. Lung Cancer Screening in a Safety-Net Hospital: Implications of Screening a Real-World Population versus the National Lung Screening Trial. *Annals ATS* 2018; 15(12): 1493-1495.

115. Mazzone PJ, Tenenbaum A, Seeley M, et al. Impact of a Lung Cancer Screening Counseling and Shared Decision-Making visit. *CHEST* 2017; 151(3): 572-578.

116. Grose D, Devereux G, Brown L, et al. Variation in comorbidity and clinical management in patients newly diagnosed with lung cancer in four Scottish centres. *J Thorac Oncol* 2011; 6: 500-509.

117. Gould Mk. Lung cancer screening in individuals with chronic obstructive pulmonary disease: Finding the sweet spot. *Am J Respir Crit Care Med* 2015; 192: 1027-8.

118. Young RP, Hopkins RJ, Duan F, et al. Stratification of NLST-ACRIN screening participants identifies the "sweet spot" of screening by identifying early stage lung cancers most amenable to curative surgery (N=10,054). *Am J Respir Crit Care Med* 2017; 195: A5127.

119. Gould MK, Munoz-Plaza CE, Hahn EE, et al. Comorbidity profiles and their effect on treatment selection and survival among patients with lung cancer. *Ann ATS* 2017; 14(10): 1571-1580.

120. Hopkins RJ, Young RP, Duan F, et al. Lung cancer screening and the effects of competing causes of death in the ACRIN-NSLT. *Am J Respir Crit Care Med* 2017; 195: A5174.

121. Berry SD, Ngo L, Samelson EJ, et al. Competing risk of death: An important consideration in studies of older adults. *J Am Geriatr Soc* 2010; 58(4): 783-787.

122. Kumar V, Cohen JT, van Klaveren D, et al. Risk-targeting lung cancer screening: a cost-effective analysis. *Ann Intern Med* 2018; 168(3): 161-169.

123. Kovalchik SA, Tammemagi M, Berg CD, et al. Targeting of low dose CT screening according to the risk of lung-cancer death. *N Eng J Med* 2013; 369: 245-254.

124. Licker MJ, Widikker I, Robert J, et al. Operative Mortality and Respiratory Complications After Lung Resection for Cancer: Impact of Chronic Obstructive Pulmonary Disease and Time Trends. *Ann Thorac Surg* 2006; 81: 1830-8.

125. Molina JR, Yang P, Cassivi SD, et al. Non-Small Cell Lung Cancer: Epidemiology, Risk Factors, Treatment and Survivorship. *Mayo Clin Proc* 2008; 83(5): 584-94.

126. Silvestri GA, Young RP. Strange bedfellows: The interaction between COPD and lung cancer in the context of lung cacner screening. *Ann ATS* 2020; 17(7): 810-812.

127. de Torres JP, Casanova C, Marin JM, et al. Exploring the impact of screening with low-dose CT on lung cancer mortality in mild to moderate COPD patients: A pilot study. *Respir Med* 2013; 107: 702-707.

128. Young RP, Hopkins RJ. CT screening in COPD: the impact on lung cancer mortality-Letter. *Resp Med* 2014; 108: 813-814.

129. Hoesein FAAM, de Jong PA, Mets OM. Optimizing lung cancer screening: nodule size, volume doubling time, morphology and evaluation of other disease. *Annals Translational Med* 2015; 3: 1-3.

130. King PT. Inflammation in chronic obstructive pulmonary disease and its role in cardiovascular disease and lung cancer. *Clin Trans Med* 2015; 4: 26.

131. Fry JS, Hamling JS, Lee PN. Systematic review with meta-analysis of the epidemiological evidence relating FEV1 decline to lung cancer risk. *BMC Cancer* 2012; 12: 498.

132. Sanchez-Salcedo P, Wilson DO, de Torres JP, et al. Improving selection for lung cancer screening; the potential role of empysema. *Am J Respir Crit Care Med* 2015; 191: 924-931.

133. Cheung LC, Berg CD, Castle PE, et al. Life-Gained-Based Versus Risk-Based Selection of Smokers for Lung Cancer Screening. *Ann Intern Med* 2019; 171: 623-632.

134. Mazzone PJ. Obstacles to and Solutions for a successful Lung Cancer Screening program. *Sem Respir Crit Care Med* 2016; 37: 659-669.

135. Biswas A, Mehta HJ, Folch EE. Chronic obstructive pulmonary disease and lung cancer: inter-relationships. *Curr Opin Pulm Med* 2018; 24: 152-160.

136. Banas E, Halbert CH, Dai L, et al. The effects of lung cancer screening decsion intervention on patient knowledge, intentions to screen and decsional satisfaction. *Am J Respir Crit Care Med* 2017; 195: A5176.

137. Byrne MM, Lillie SE, Studts JL. Lung cancer screening in a community setting: Characteristics, motivations, and attitudes of individuals being screened. *Health Psychology Open* 2019; January-June: 1-9.

138. Young RP, Hopkins RJ. Increasing smokers' risk perception improves CT screening participation. *Thorax* 2012; 67(9): 834-835.

139. Bartlett EC, Kemp SV, Ridge CA, et al. Baseline results of the West London lung cancer screening pilot study – Impact of mobile scanners and dual risk model utilisation. *Lung Cancer* 2020; 148: 12-19.

140. Lal A, Phillips S, Russell C, et al. The novel use of fast track CT to select patients for lung cancer clinics: effect on clinic efficiency, waiting times, and patient satisfaction. *Postgrad Med J* 2011; 87: 264-268.

141. Evison M, Hewitt K, Lyons J, et al. Implementation and outcomes of the RAPID programme: Addressing the front end of the lung cancer pathway in Manchester. *Clinical Medicine* 2020; 20(4): 401-405.

142. Caverly TJ, Cao P, Hayward RA, et al. Identifying Patients for Whom Lung Cancer Screening is Preference-Sensitive. *Ann Int Med* 2018; 169: 1-9.

143. Rajupet S, Doshi D, Wisnivesky JP, et al. Attitudes about lung cancer screening: Primary care providers versus specialists. *Clinical Lung Cancer* 2017; 18(6): e417-23.

144. Henderson LM, Benefield TS, Bearden SC, et al. Changes in Physician knowledge, attitudes, beliefs, and practices regarding lung cancer screening. *Annals ATS* 2019; 16(8): 1065-1069.

145. Cruz CSD, Tanoue LT, Matthay RA. Lung cancer: epidemiology, aetiology, and prevention. *Clin Chest Med* 2011; 32: 605-644.

146. Haiman CA, Stram DO, Wilkens LR, et al. Ethnic and Racial Differences in the Smoking-related Risk of Lung Cancer. *N Eng J Med* 2006; 354: 333-342.

147. Risch N. Dissecting racial and ethnic differences. *N Eng J Med* 2006; 345: 408-411.

148. Hopkins RJ, Duan F, Young RP. Lung cancer susceptibility, ethnicity, and the benefits of computed tomography screening. *Am J Respir Crit Care Med* 2015; 192: 1394-1396.

149. Etzel CJ, Kachroo S, Liu M, et al. Development and validation of a lung cancer risk prediction model for African-Americans. *Cancer Prev Res* 2008; 1: 255-265.

150. Khoury MJ, Beaty TH, Tockman MS, et al. Familial aggregation in chronic obstructive pulmonary diseae: Use of the loglinear model to analyze intermediate environmental and genetic risk factors. *Genetic Epidemiology* 1985; 2: 155-166.

151. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer: analyses of chorts of twins from Sweden, Denmark and Finland. *N Eng J Med* 2000; 343: 78-85.

152. Wyss AB. Multiethnic meta-analysis identifies ancestry-specific and cross-ancestry loci for pulmonary function. *Nature Communications* 2018; 9: 2976.

153. Young RP, Hopkins RJ, Hay BA, Gamble GD. GSTM1 null genotype in COPD and lung cancer: evidence of a modifier or confounding effect? *The Application of Clinical Genetics* 2011; 4: 137-144.

154. Young RP, Hopkins RJ, Whittington CF, et al. Individual and Cumulative Effects of GWAS Susceptibility Loci in Lung Cancer: Associations after Sub-Phenotyping for COPD. *PLoS ONE* 2011; 6(2): e16476: 1-13.

155. Wilk JB, Chen T-h, Gottlieb DJ, et al. A Genome-Wide Association Study of Pulmonary Function Measures in the Framingham Heart Study. *PLoS Genetics* 2009; 5(3): e1000429.

156. Gim J, An J, Sung J, et al. A between ethnicities comparison of chronic obstructive pulmonary disease genetic risk. *Frontiers in Genetics* 2020; 11: article329.

157. Hopkins RJ, Kendall C, Gamble GD, Young RP. A comparison of lung cancer in New Zealand Europeans and Māori: Are Māori more susceptible to smoking? *Am J Respir Crit Care Med* 2016; 193: A6658.

158. Loring B. Litrature review Respiratory Health for Maori. New Zealand: The Asthma and Respiratory Foundation of New Zealand (Inc.); 2009 Jan. Available from:http://asthmafoundation.org.nz

159. Hunt D, Blakely T, Woodward A, et al. The smoking-mortality association varies over time and by ethnicity in New Zealand. *Int J Epidemiol* 2005; 34: 1020-1028.

160. Ministry of Health. Cancer: New Registrations and Deaths 2011. Wellington: Ministry of Health; 2014. Available from: https://www.health.govt.nz

161. Global Lung Cancer Coalition eAtlas: http://www.lungcancercoalition.org/e-atlas/ last accessed 01 February 2019.

162. Hancock DB, Eijgelsheim M, Wilk JB, et al. Meta-analyses of genome-wide association studies identify multiple novel loci associated with pulmonary function. *Nat Genet* 2010; 42(1): 45-52.

163. Repapi E, Sayers I, Wain LV, et al. Genome-wide association study identifies five loci associated with lung function. *Nature Genetics* 2009*;* 42: 36–44.

164. Pillai SG, Ge D, Zhu G, et al. a Genome-Wide Association Study in Chronic Obstructive Pulmonary Disease (COPD): Identification of two major susceptibility loci. *PLoS ONE Genetics* 2009; 5(3): e1000421.

165. Amos CI, Wu X, Broderick P, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nature Genetics* 2008; 40: 616–622.

166. Hung RJ, Mckay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 2008; 452: 633-637.

167. Thorteirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature 2008*; 452(7187): 638-642.

168. Young RP, Hopkins RJ, Hay BA, et al. FAM13A locus in COPD independently associated with lung cancer – evidence of a molecular genetic link between COPD and lung cancer. *App Clin Genet* 2011; 4: 1–10.

169. National Lung Screening Trial Research Team. Writing Committee: Aberle DR, Adams AM, Berg CD et al. Baseline characteristics of participants in the randomized National Lung Screening Trial. *J Natl Cancer Inst* 2010; 102: 1771-1779.

170. Hinshaw LB, Jackson SA, Chen MY. Direct mailing was a successful recruitment strategy for a lung-cancer screening trial. *J Clinical Epidemiology* 2007; 60: 853-857.

171. National Lung Screening Trial Research Team. Writing committee: Aberle DR, Adams AM, Berg CD et al. The National Lung Screening Trial: Overview and study design. *Radiology* 2011; 258(1): 243-253.

172. Gareen IF, Sicks JR, Jain AA, et al. Identifying and collecting pertinent medical records for centralized abstraction in a multi-center randomized clinical trial: The model used by the American College of Radiology arm of the National Lung Screening Trial. *Contemporary Clinical Trials* 2013; 34: 36-44.

173. Patz EF, Caporaso NE, Dubinett SM, et al. NLST ACRIN Biomarker Repository Originating from the Contemporary Screening for the Detection of Lung Cancer Trial (NLST, ACRIN 6654): Design, Intent, and Availability of Specimens for Validation of Lung Cancer Biomarkers. *J Thorac Oncol*. 2010; 5(10): 1502–1506.

174. Wille MMW, Dirksen A, Ashraf H, et al. Results of the randomized Danish Lung Cancer Screening Trial with focus on high risk profiling. *Am J Respir Crit Care Med* 2016; 193(5): 542-551.

175. Paci E, Puliti D, Pegna AL, et al. Mortality, survival, and incidence rates in the ITALUNG randomised lung cancer screening trial. *Thorax* 2017; 72: 825-831.

176. Pastorino U, Silva M, Sestini S, et al. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. *Annals of Oncology* 2019; 30: 1162-1169.

177. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Eng J Med* 2020; 382(6): 503-13.

178. The International Classification of Diseases Ninth Revision, Clinical Modification (ICD-9-CM) system. Available from: https://www.cdc.gov/nchs/icd/icd9cm

179. American Medical Association Current Procedural Terminology code set. Available from: https://www.ama-assn.org/practice-management/cpt

180. Miller MR, Hankinson J, Brusasco V, et al. ATS/ERS Task Force. Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.

181. Moher D, Liberati A, Tetzlaff J, et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 2009; 6: e1000097.

182. Liberati A, Aaltman DG, Tetzlaff J et al. The PRISMA Statement for Reporting Systmatic Reviews and Meta-analyses of Studies that Evaluate Health Care Intervention: Explanation and elablorations. *PloS Medicine* 2009; 6(7): e1000100.

183. Rami-Porta R, Bolejack V, Giroux DJ, et al. The IASLC Lung Cancer Staging Project: The New Database to Inform the Eighth Edition of the TNM Classification of Lung Cancer. *J Thorac Oncol*. 2014; 9 (11): 618–1624.

184. de Torres J, Bastarrika G, Wisnivesky JP, et al. Assessing the relationship between lung cancer risk and emphysema detected on low dose CT of the chest. *Chest* 2007; 132: 1932-1938.

185. Wilson DO, Weissfeld JL, Balkan A, et al. Association of radiographic emphysema and airflow obstruction with lung cancer. *Am J Respir Crit Care Med* 2008; 178: 738-744.

186. Calabra E, Randi G, La Vecchia C, et al. Lung function predicts lung cancer risk in smokers: a tool for targeting screening programmes. *Eur Respir J* 2010; 35: 146-151.

187. de Torres JP, Marin JM, Casanova C, et al. Lung Cancer in Patients with Chronic Obstructive Pulmonary Disease : Incidence and Predicting Factors. *Am J Respir Crit Care Med* 2011; 184: 913-919.

188. Chang KC, Leung CC. Lung Cancer is More Common in Early GOLD Stages of COPD: A Spurious Association? *Am J Respir Crit Care Med* 2012; 185: 1128.

189. de Torres JP, Zulueta JJ, Casanova C, et al. Lung Cancer is More Common in Early GOLD Stages of COPD: A Spurious Association? Response. *Am J Respir Crit Care Med* 2012; 185: 1128-1129.

190. Caramori G, Adcock IM, Casolari P, et al. Unbalanced oxidant-induced DNA damage and repair in COPD; a link towards lung cancer. *Thorax* 2011; 66: 521-527.

191. Crystal RG. Airway basal cells: The "smoking gun" of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2014; 190: 1355-1362.

192. Man SFP, Connett JE, Anthonisen NR, et al. C-reactive protein and mortality in mild to moderate chronic obstructive pulmonary disease. *Thorax* 2006; 61: 849-853.

193. Allin KH, Bojesen SE, Nordestgaard. Baseline C-Reactive Protein Is associated With Incident Cancer and Survival in Patients With Cancer. *J Clin Oncology* 2009; 27 (13): 2217-2224.

194. Sohal SS, Reid D, Soltani A, et al. Reticular basement membrane fragmentation and potential mesenchymal transition is exaggerated in the airways of smokers with chronic obstructive pulmonary disease. *Respirol* 2010; 15: 930-938.

195. Gohy ST, Hupin C, Fregimilicka C, et al. Imprinting of the COPD airway epithelium for dedifferentiation and mesenchymal transition. *Eur Respir J* 2015; 45: 1258-1272.

196. Wan ES, Hokanson JE, Murphy JR, et al. Clinical and radiographic predictors of GOLD-Unclassified smokers in the COPDGene study. *Am J Respir Crit Care Med* 2011; 184: 57-63.

197. Young RP, Duan F, Hopkins RJ, et al. COPD-related risk factors in screen detected lung cancer: preliminary results of a sub analysis of the NLST. *Am J Respir Crit Care Med* 2013; 187: A2345.

198. Tammemagi CM, Pinsky PF, Caporaso NE, et al. Lung Cancer Risk Prediction: Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Models and Validation. *J Natl Cancer Inst* 2011; 103: 1058-1068.

199. Tammemagi MC, Lam SC, McWilliams AM et al. Incremental Value of Pulmonary Function and Sputum DNA Image Cytometry in Lung Cancer Risk Prediction. *Cancer Prev Res* 2011; 4: 552-561.

200. Berry CE, Wise RA. Mortality in COPD: Causes, Risk factors, and Prevention. *COPD* 2010; 7: 375-382.

201. Carlsten C, Sagoo GS, Frodsham AJ, et al. Glutathione S-transferase M1 (GSTM1) polymorphisms and lung cancer: a literature based systematic HuGE review and meta-analysis. *Am J Epidemiol* 2008; 167: 759-774.

202. Kish K, Gurney JW, Schroeder DR, et al. The correlation of emphysema or airway obstruction with the risk of lung cancer: a matched case controlled study. *Eur Respir J 2002*; 19: 1093-1098.

203. Maldonado F, Bartholmai BJ, Swensen SJ, et al. Are airflow obstruction and radiographic evidence of emphysema risk factors for lung cancer? A nested case-control study using quantitative emphysema analysis. *Chest* 2010; 138: 1295-1302.

204. Strange C. COPD and lung cancer: Are they both airways diseases? *Chest* 2010; 138: 1289-1290.

205. Gierada DS, Guniganti P, Newmans BJ, et al. Quantitative CT assessment of emphysema and airways in relation to lung cancer risk. *Radiology* 2011; 261: 950-959.

206. Carrozzi L, Viegi G. Lung cancer and chronic obstructive pulmonary disease; the story goes on. *Radiology* 2011; 261: 688-691.

207. Smith BM, Pinto L, Ezer N, et al. Emphysema detected on computed tomography and risk of lung cancer: A systematic review and meta-analysis. *Lung Cancer* 2012; 77: 58-63.

208. Young RP, Hopkins RJ. Estimating Overdiagnosis of lung cancer. *Ann Int Med* 2013; 158: 635-636.

209. Young RP, Hopkins RJ. Targeted CT image screening and its effect on lung cancer detection rate. *Chest* 2013; 144(4): 1419-1421.

210. Lowry KP, Gazelle GS, Gilmore ME, et al. Personalizing annual lung cancer screening for patients with chronic obstructive pulmonary disease: a decision analysis. *Cancer* 2015; 121: 1556-1562.

211. Bach PB. Computed tomographic screening for lung cancer – response letter. *JAMA* 2012; 308: 1320-1321.

212. Veronesi G, Maisonneuve P, Bellomi M, et al. Estimating overdiagnosis in low-dose computed tomography screening for lung cancer: A cohort study. *Ann Int Med* 2012; 157: 776-784.

213. Saghir Z, Dirksen A, Ashraf H, et al. CT screening for lung cancer brings forward early disease. The randomised Danish Lung Cancer Screening Trial: status after five annual screening rounds with low-dose CT. *Thorax* 2012; 67: 296-301.

214. Horeweg N, van der Aalst CM, Thunnissen E, et al. Characteristics of Lung Cancer Detected by Computer Tomography Screening in the Randomized NELSON Trial. *Am J Respir Crit Care Med* 2013; 187: 848-854.

215. Infante M, Cavuto S, Lutman FR, et al. Longterm follow-up results of the DANTE trial, a randomised study of lung cancer screening with spiral computed tomography. *Am J Respir Crit Care Med* 2015; 191: 1166-1175.

216. Young RP, Hopkins RJ. Mortality reduction, overdiagnosis, and the benefit-to-harm ratio of computed tomography screening. *Am J Respir Crit Care Med* 2015; 192: 398-399.

217. Young RP, Hopkins RJ. A clinical practice guideline update on the diagnosis and management of stable chronic obstructive pulmonary disease. [Letter] *Ann Intern Med* 2012; 156: 68-69.

218. Humphrey LL, Deffebach M, Pappas M, et al. Screening for lung cancer with low-dose computed tomography: a systematic review to update the US Preventive Services Task Force recommendation. *Ann Int Med* 2013; 159(6): 411-20.

219. Hopkins RJ, Duan F, Chiles C, et al. Reduced Expiratory Flow Rates among Heavy Smokers Increases Lung Cancer Risk: Results from the National Lung Screening Trial-American College of Radiology Imaging Network Cohort. *Annals ATS* 2017; 14: 392-402.

220. Colhoun HM, Betteridge DJ, Durrington PN, et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes study (CARDS): multicentre randomised placebo-controlled trial. *Lancet* 2004; 364: 685-696.

221. Tan KS, Eguchi T, Adusumilli PS. Competing risk and cancer-specific mortality: why it matters. *Oncotarget* 2018; 9: 7272-7273.

222. Vestbo J, Lange P. Can GOLD stage 0 provide information of prognostic value in chronic obstructive pulmonary disease? *Am J Respir Crit Care Med* 2002; 166: 329-332.

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

224. Vrbica Z, Labor M, Gudelj I, et al. Early detection of COPD patients in GOLD 0 population: an observational non-interventional cohort study –MARKO study. *BMC Pulmonary Medicine* 2017; 17: 36.

225. Young RP, Hopkins RJ. Interleukin-6 and statin therapy: potential role in the management of COPD. *Resp Res* 2013:14: 74.

226. Ridker PM, MacFayden JG, Thuren T, et al. Effect of interleukin -1β inhibition with canakinumab on lung cancer incident lung cancer in patients with atherosclerosis: exploratory results from a randomised, double-blind, placebo-controlled trial. *The Lancet* 2017; 390: 1833-1842.

227. Young RP, Hopkins RJ. Update on the potential role of statins in chronic obstructive pulmonary disease and its co-morbidities. *Expert Rev Respir Med* 2013; 7(5): 533-544.

228. Mannino DM. GOLD Stage 0 COPD: Is it real? Does it matter? *Chest* 2006; 130: 309-310.

229. Wan ES, Fortis S, Regan EA, et al. Longitudinal phenotypes and mortality in Preserved Ratio Impaired Spirometry in the COPDGene Study. *Am J Respir Crit Care Med* 2018; 198(11): 1397-1405.

230. Lin K, Watkins B, Johnson T, et al. Screening for Chronic Obstructive Pulmonary disease Using Spirometry: Summary of the Evidence for the US Preventive Services Task Force. Agency for Healthcare Research and Quality (US): 2008. March. Report No. 08-05113-EF-1.

231. Qaseem A, Wilt TJ, Weinberger SE, , et al. American College of Physicians. Diagnosis and Management of Stable Chronic Obstructive Pulmonary Disease: A Clinical Practice Guideline Update from the American College of Physicians, American College of Chest Physicians, American Thoracic Society, and European Respiratory Society. *Ann Int Med* 2011; 155: 179-191.

232. Gould MK. Precision Screening for Lung cancer: Risk-based but Not Always Preference-Sensitive. *Ann Int Med* 2018; 169: 52-53.

233. Bhatt JR, Klotz L. Overtreatment in cancer – is it a problem? *Expert Opin Pharmacother* 2016; 17(1): 1-5.

234. Young RP, Hopkins R. Identifying patience for whom lung cancer screening is preference -sensitive. *Ann Int Med* 2018; 169: 822-823.

235. Young RP, Duan F, Greco E, Hopkins RJ, et al. SNP-based risk score out performs a clinical model for dying of lung cancer in the NLST-ACRIN sub-study (N=10,054). *Am J Respir Crit Care Med* 2017; 195: A6154.

236. Katki HA, Kovalchik SA, Berg CD, et al.. Development and Validation of Risk Models to Select Ever-smokers for CT Lung Cancer Screening. *JAMA* 2016; 315: 2300-2311.

237. ten Haaf K, Jeon J, Tammemagi MC, et al. Risk prediction models for selection of lung cancer screening candidates: A retrospective validation study. *PLoS Med* 2017; 14(4): e1002277.

238. Tammemagi MC. Selecting lung cancer screenees using risk prediction models- where do we go from here. *Transl Lung Cancer Res* 2018; 7(3): 243-253.

239. Young RP, Hopkins RJ, Duan F,et al. The relationship between lung cancer risk according to the $PLCO_{2012}$ model and prevalence or presence of COPD in the NLST-ACRIN sub-study (N=10,054). *Am J Respir Crit Care Med* 2018; 197: A4411.

240. Papi A, Casoni G, Caramori G, et al. COPD increases the risk of squamous histological subtype in smokers who develop non-small cell lung carcinoma. *Thorax* 2014; 59: 679-681.

241. Young RP, Hopkins RJ, Duan F, et al. Genetic-based approach to stratifying risk of lung cancer outperforms the $PLCO_{2012}$ model – optimisation of CT screening outcomes in the NLST-ACRIN. Sub-study (N=10,054). *Am J Respir Crit Care Med* 2018; 197: A4422.

242. WHO. World Cancer Report 2014. International Agency for Research on Cancer 2014.

243. Strand TE, Rostad H, Moller B, et al. Survival after resection for primary lung cancer: a population based study of 3211 resected patients. *Thorax* 2006; 62: 710-5.

244. Young RP, Hopkins RJ. Lung function predicts lung cancer. *Eur Respir J* 2010; 35: 1421-1422.

245. Romano PS, Mark DH. Patient and Hospital Characteristics Related to In-Hospital Mortality after Lung Cancer Resection. *Chest* 1992; 101: 1332-7.

246. Iwasaki A, Shirakusa T, Enatsu S, et al. Surgical Treatment for Lung Cancer with COPD Based on the Global Initiative for Chronic Obstructive Lung Disease (GOLD). *Thorac Cardiovasc Surg* 2005; 53: 162-7.

247. Sekine Y, Behnia M, Fujisawa T. Impact of COPD on pulmonary complications and on long-term survival of patients undergoing surgery for NSCLC. *Lung Cancer* 2002; 37: 95-101.

248. López-Encuentra A, Astudillo J, Cerezal J, et al. Prognostic value of chronic obstructive pulmonary disease in 2994 cases of lung cancer. *Eur J Cardiothorac Surg* 2005; 27: 8-13.

249. Sekine Y, Suzuki H, Yamada Y, et al. Severity of chronic obstructive pulmonary disease and its relationship to lung cancer prognosis after surgical resection. *Thorac Cardiovasc Surg* 2013; 61(2): 124-30.

250. Lee SJ, Lee J, Park YS, et al. Impact of Chronic Obstructive Pulmonary Disease on the Mortality of Patients with Non-Small-Cell Lung Cancer. *J Thorac Oncol* 2014; 9: 812-17.

251. Santambrogio L, Nosotti M, Baisi A, et al. Pulmonary lobectomy for lung cancer: a prospective study to compare patients with forced expiratory volume in 1s more or less than 80% of predicted. *Eur J Cardiothorac Surg* 2001; 20: 684-7.

252. Ueda K, Jinbo M, Li TS, et al. Computed Tomography-Diagnosed Emphysema, Not Airway Obstruction, Is Associated with the Prognostic Outcome of Early-Stage Lung Cancer. *Clin Cancer Res* 2006; 12(22): 6730-6.

253. Nakajima T, Sekine Y, Yamada Y, et al. Long-term Surgical Outcome in Patients with Lung Cancer and Coexisting Severe COPD. *J Thorac Cardiovasc Surg* 2009; 57: 339-42.

254. Sawabata N, Miyoshi S, Matsumura A, et al. Prognosis of smokers following resection of pathological stage I non-small-cell lung carcinoma. *Gen Thorac Cardiovasc Surg* 2007; 55: 420-4.

255. Yoshida Y, Kage H, Murakawa T, et al. Worse Prognosis for Stage IA Lung Cancer Patients with Smoking History and More Severe Chronic Obstructive Pulmonary Disease. *Ann Thorac Cardiovasc Surg* 2015; 21: 194-200.

256. Global Initiative for Chronic Obstructive Lung disease: www.GOLD.org (accessed May 2017).

257. Kuo CH, Wu CY, Lee KY, et al. Chronic Obstructive Pulmonary Disease in Stage I Non-small Cell Lung Cancer That Underwent Anatomic Resection: The Role of a Recurrence Promoter. *COPD* 2014; 11: 407-13.

258. Wilson DO, Ryan A, Fuhrman C, et al. Doubling times and CT screen-detected lung cancers in the Pittsburgh Lung Screening Study. *Am J Respir Crit Care Med* 2012; 185: 85-89.

259. Janssen-Heijnen MLG, van Erning FN, De Ruysscher DK, et al. Variation in causes of death in patients with non-small cell lung cancer according to stage and time since diagnosis. *Annals of Oncology* 2015; 26:902-907.

260. Mazzone P. Lung cancer screening: examining the issues. *Cleveland Clinic J Med* 2012; 79: e-S1-e-S6.

261. Zhai R, Yu X, Shafer A, et al. The Impact of Coexisting COPD on Survival of Patients With Early-Stage Non-Small Cell Lung Cancer Undergoing Surgical Resection. *Chest* 2014; 145(2): 346-53.

262. Subotic DR, Mandaric DV, Eminovic TM, et al. Influence of chronic obstructive pulmonary disease on postoperative lung function and complications in patients undergoing operations for primary non-small cell lung cancer. *J Thorac Cardiovasc Surg* 2007; 134(5): 1292-9.

263. Korst RJ, Ginsberg RJ, Ailawadi M, et al. Lobectomy Improves Ventilatory Function in Selected Patients with Severe COPD. *Ann Thorac Surg* 1988; 66: 898-902.

264. Brenner DJ. Radiation Risks Potentially Associated with Low-Dose CT Screening of Adult Smokers for Lung Cancer. *Radiology* 2004; 231: 440-5.

265. Kroenke K, Lawrence VA, Theroux JF, et al. Postoperative Complications After Thoracic and Major Abdominal Surgery in Patients With and Without Obstructive Lung Disease. *Chest* 1993; 104: 1445-51.

266. Hu Y, McMurry TL, Isbell JM, et al. Readmission following lung cancer resection is associated with a six-fold increase in 90-day postoperative mortality. *J Thorac Cardiovasc Surg* 2014; 148(5): 2261-67.

267. Seok Y, Lee E, Cho S. Respiratory Complications during Mid- and Long-Term Follow-Up Periods in Patients Who Underwent Pneumonectomy for Non-Small Cell Lung Cancer. *Ann Thorac Cardiovasc Surg* 2013; 19: 335-40.

268. Senbaklavaci O. Lobar Lung Resection in Elderly Patients With Non-Small Cell Lung Carcinoma: Impact of Chronic Obstructive Pulmonary Disease on Surgical Outcome. *Int Surg* 2014; 99: 319-24.

269. Stolz A, Pafko P, Harustiak T, et al. Risk factor analysis for early mortality and morbidity following pneumonectomy for non-small cell lung cancer. *Barisl Lek Listy* 2011; 112(4): 166-9.

270. Sekine Y, Kesler, KA, Behnia M, et al. COPD May Increase the Incidence of Refractory Supraventricular Arrhythmias Following Pulmonary Resection for Non-small Cell Lung Cancer. *Chest* 2001; 120: 1783-90.

271. Curkendall SM, DeLuise C, Jones JK, et al. Cardiovascular Disease in Patients with Chronic Obstructive Pulmonary Disease, Saskatchewan Canada. *Ann Epidemiol* 2006; 16: 63-70.

272. Volpino P, Cangemi R, Fiori E, et al. Risk of mortality from cardiovascular and respiratory causes in patients with chronic obstructive pulmonary disease submitted to follow-up after lung resection for non-small cell lung cancer. *J Cardiovasc Surg* 2007; 48: 375-83.

273. Young RP, Hopkins RJ, Midthun DE. Computed tomographic screening for lung cancer. *JAMA* 2012; 308: 1320.

274. Mattson ME, Pollack ES, J W Cullen. What are the odds that smoking will kill you? *Am J Pub Health* 1987; 77: 425-431.

275. Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev* 1988; 10: 87-121.

276. Gala G. Northern cancer network: Cancer in the Northern Region of New Zealand Health Needs Assessment -2009.

277. Friedlander JS, Friedlander FR, Reed FA, et al. The genetic structure of Pacific Islanders. *PLoS Genet* 2008; 4: 173-190.

278. Harwood M, Aldington S, Beasley R. Lung cancer in Māori: A neglected priority. *NZ Med J* (Editorial). 2005; (Apr): 118: 1213).

279. Blakely T, Fawcett J, Hunt D, et al. What is the contribution of smoking and socioeconomic position to ethnic inequalities in mortality in New Zealand? *Lancet* 2006; 368: 44-52.

280. Sutherland T, Aitken D. Ethnic and socioeconomic inequalities in lung cancer in a New Zealand population. *Respirology*. 2008; 13: 590-593.

281. Stevens W, Stevens G, Kolbe J, et al. Ethnic differences in the Management of Lung Cancer in New Zealand. *J Thorac Oncol*. 2008 ; 3(3): 237-44.

282. Ministry of Health. Tobacco Use 2012/13: New Zealand Health Survey. Wellington: Ministry of Health; 2014.

283. Milivojevic-Poleksic L, Wells AU, Moody A, et al. Spirometric lung volumes in the adult Pacific Islander population: comparison with predicted values in a European population. *Respirology* 2001; 6: 247-253.

284. Dransfield MT, Davis JJ, Gerald LB, et al. Racial and gender differences in susceptibility to tobacco smoke among patients with chronic obstructive pulmonary disease. *Respir Med* 2006; 100: 1110-1116.

285. Jedrychowski W, Becher H, Wahrendorf J, et al. Effect of tobacco smoking on various histological types of lung cancer. *J Cancer Res Clin Oncol* 1992; 118(4): 276-282.

286. Glass WI. Ventilatory function differences between Polynesian and European rope workers. *NZ Med J* 1962; 61: 433-444.

287. De Hamal FA, Glass WJ. Observations on Māori-European lung function differences. *Aust NZ J Med* 1975; 5: 44-48.

288. Derby KS, Cuthrell K, Caberto C, et al. Nicotine metabolism in three ethnic/racial groups with different risks of lung cancer. *Cancer Epidemiol Biomarkers Prev* 2008 Dec; 17(12): 3526-35.

289. Park SL, Carmella SG, Ming X, et al. Variation in the levels of the lung carcinogen NNAL and its glucuronides in the urine of cigarette smokers from five ethnic groups with differing risks of lung cancer. *Cancer Epidemiol Biomarkers Prev* 2015; 24: 561-569.

290. Correa P, Fontham E, Pickle LW, et al. Passive smoking and lung cancer. *Lancet* 1983; 2: 595-597.

291. Aldington S, Harwood M, Cox B, et al. Cannabis use and risk of lung cancer: a case-control study. *Eur Respir J* 2008; 31(2): 280-286.

292. Wallace TA, Martin DN, Ambs S. Interaction among genes, tumor biology and the environment in cancer health disparities: examining the evidence on a national and global scale. *Carcinogenesis* 2011; 32: 1107-1121.

293. Odonez-Mena JM, Schottker B, Mons U, et al. Quantification of the smoking-associated cancer risk with rate advancement periods: meta-analysis of individual participant data from cohorts of the CHANCES consortium. *BMC Medicine* 2016; 14: 62: 1-15.

294. Liu DMKI, Kwee SA. Demographic, treatment, and survival patterns for Native Hawaiians with lung cancer treated at a community medical centre from 1995-2001. *Pacific Health Dialogue* 2004; 11: 139-145.

295. Miquel JF, Covarrubias C, Villaroel L, et al. Genetic epidemiology of cholesterol cholelithiasis among Chilean Hispanics, Amerindians, and Maoris. *Gastroenterology* 1998; 115: 937-946.

296. Liu T, Xie CB, Ma WJ, et al. Association between CYP2A6 genetic polymorphisms and lung cancer: a meta-analysis of case-control studies. *Environ Mol Mutagen* 2013; 54(2): 133-40.

297. Miyazaki M, Yamazaki H, Saoo K, et al. Mechanisms of chemopreventive effects of 8-methoxypsoralen against 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone-induced mouse lung adenomas. *Carcinogenesis* 2005; 26(11): 1947-55.

298. Fagan P, Moolchan ET, Pokhrel P, et al. Biomarkers of tobacco smoke exposure in racial/ethnic groups at high risk for lung cancer. *Am J Pub Health* 2015; 105: 1237-1245.

299. Herzog TA, Pokhrel P. Ethnic differences in smoking rates, nicotine dependence and cessation related variables among adult smokers in Hawaii. *J Comm Health* 2012; 37: 1226-1233.

300. Schabath MB, Hernandez LM, Wu X, et al. Dietary Phytoestrogens and Lung Cancer Risk. *JAMA* 2005; 294(12): 1493-1504.

301. Serjeantson SW. The Colonization of the Pacific - A Genetic Trail. Oxford University Press; 1989.

302. Saccone NL, Wang JC, Breslau N, et al. The *CHRNA5-CHRNA3-CHRNB4* nicotinic receptor subunit gene cluster affects risk for nicotine dependence in African-Americans and in European-Americans. *Cancer Res* 2009; 69: 6848-6856.

303. Chen L-S, Baker T, Hung RJ, et al. Genetic risk can be decreased: Quitting smoking decreases and delays lung cancer for smokers with high and low *CHRNA5* risk genotypes - A meta-analysis. *EBioMedicine* 2016; e00734: 1-8.

304. Gabrielsen ME, Romundstat P, Langhammer A, et al. Association between a 15q25 gene variant, nicotine-related habits, lung cancer and COPD among 53 307 individuals from the HUNT study in Norway. *Eur J Hum Genet* 2013; 21: 1293-1299.

305. Wang Y, Broderick P, Matakidou A, et al. Chromosome 15q25 (*CHRNA3-CHRNA5*) variation impacts indirectly on lung cancer risk. *PLoS ONE* 2011; 6(4): e19085.

306. Wang JC, Cruchaga C, Saccone NL, et al. Risk for nictotine dependence and lung cancer is conferred by mRNA expression levels and amino acid change in *CHRNA5*. *Human Molecular Genetics* 2009; 18(16): 3125-3135.

307. Smith RM, Alachkar H, Papp AC, et al. Nicotinic α5 receptor subunit mRNA expression is associated with distant 5' upstream polymorphisms. *Eur J Hum genet* 2011; 19: 76-83.

308. Young RP, Hopkins RJ. Etzel C, El-Zein R. Genetics of lung cancer susceptibility and COPD. *Lancet Oncol* 2011; 12(7): 622-623.

309. Mohamed Hoesein FA, de Jong PA. Personalized lung cancer screening: the value of spirometry and emphysema as risk modifiers. *Ann Transl Med* 2016; 4(15): 293.

310. Young RP, Hopkins RJ, Hay BA, et al. Lung cancer susceptibility model based on age, family history and genetic variants. *PLoS ONE* 2009; 4(4): e5302.

311. McKay JD, Hung RJ, Han Y, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nature Genetics* 2017; 49: 1126-1132.

312. Lassi G, Taylor AE, Timpson NJ, et al. The *CHRNA5-A3-B4* gene cluster and smoking: from discovery to therapeutics. *Trends in Neuroscience* 2016; 39(12): 851-861.

313. Scott A, Lugg ST, Aldridge K, et al. Pro-inflammatory effects of e-cigarette vapour condensate on human alveolar macrophages. *Thorax* 2018; 73: 1161-1169.

314. Tang MS, Wu XR, Lee HW, et al. Electronic-cigarette smoke induces lung adenocarcinoma and bladder urothelial hyperplasia in mice. *PNAS* 2019; 116 (43): 21727-21731.

315. Young RP, Hopkins RJ. Incorporating genomic data into multivariate risk models for lung cancer. *Genet Med* 2013; 15(8): 667-668.

316. VanderWeele TJ, Asomaning K, Tchetgen EJ, et al Genetic variation on 15q25.1, smoking, and lung cancer: An assessment of mediation and interaction. *Am J Epidemiol* 2012; 175(10): 1013-1020.

317. Improgo MRD, Schofield MD, Tapper AR, et al. From smoking to lung cancer: The *CHRNA5/A3/B4* connection*. Oncogene* 2010; 29(35): 4874-4884.

318. Lam DC-L, Luo SY, Fu K-H, et al. Nicotinic acetylcholine receptor expression in human airway correlates with lung function. *Am J Physiol Lung Cell Mol Physiol* 2016; 310: L232-L239.

319. Young RP, Hopkins RJ, Duan F, et al. Genetic risk score from 12 SNP panel predicts lung cancer lethality in the National Lung Screening Trial (NLST) – A validation study in the NLST_ACRIN (N=10,054). *Am J Respir Crit Care Med* 2018; 197: A2475.

320. DeMeo DL, Mariani T, Bhattacharya S, et al. Intergration of genomic and genetic approaches implicates *IREB2* as a COPD susceptibility gene. *Am J* Human *Genetics* 2009; 85: 493-502*.*

321. Mazzone PJ, Sears CR, Arenberg DA, et al. Evaluating molecular biomarkers for the early detection of lung cancer: When is a biomarker ready for clinical use? *Am J Respire Crit Care Med* 2017; 1296: e15-29.