QUANTIFYING IMPORTANT RISK FACTORS AND SURVIVAL FOLLOWING TREATMENT IN PEOPLE WITH LUNG CANCER USING ROUTINELY COLLECTED NATIONAL DATA

Dr Helen Powell BMedSci BMBS MRCP

Thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

December 2013

TABLE OF CONTENTS

ABSTRACT

Background

Survival for people with lung cancer is poor both in comparison with other cancers and for the United Kingdom (UK) compared with other developed countries. Inequalities in access to care for people with lung cancer have been demonstrated using large, routinely collected, datasets. One especially useful resource in this context is The National Lung Cancer Audit (NLCA) which was set up in 2002 with the aim of improving outcomes for people with lung cancer. It has collected data on people with primary lung cancer from hospital trusts in England and Wales since 2004. This is now the largest database of people with lung cancer in Europe containing over 150,000 cases and with close to 100% case ascertainment; approximately 35,000 new cases are added each year.

Methods

In addition to the NLCA, routinely collected primary care data from The Health Improvement Network, and the database which results from the clinical coding of all inpatient hospital admissions in England (the Hospital Episodes Statistics (HES) database) were used to investigate several clinical questions in lung cancer. Records for people in the NLCA were linked with their HES records by the Health and Social Care Information Centre (HSCIC). Death registration is mandatory in the UK and these records were obtained from the Office for National Statistics (ONS), and linked with HES data. The ONS death data were used to provide accurate and complete follow-up for mortality and survival analyses.

The observational studies in this thesis used matched case-control methodology and multivariate logistic regression to investigate the association between sex, smoking quantity, chronic obstructive lung disease (COPD), and lung cancer.

Case control and cohort studies were performed to investigate early mortality after lung cancer surgery and treatment decisions in small cell lung cancer (SCLC). Multivariate logistic regression was used to generate a score to predict the risk of early mortality after lung cancer surgery. Survival analyses including Kaplan Meier curves and Cox regression were used to determine the most accurate definition of surgery and chemotherapy from the NLCA and HES databases and to provide information on outcomes after chemotherapy.

Results

Sex significantly modified the effect of smoking on lung cancer (multiplicative test for interaction likelihood ratio $p<0.0001$) with women at higher risk for the same quantity smoked. Chronic obstructive pulmonary disease was strongly associated with lung cancer in univariate analysis (odds ratio 11.47, 95% confidence interval 9.38-14.02 for people with recently diagnosed COPD compared with those without COPD) however this was heavily confounded by smoking and strongly related to the timing of diagnosis.

For people with non-small cell lung cancer the 90-day mortality after lung cancer surgery was 5.9%. Factors which were significantly associated with this outcome (and therefore make up the predictive score) included age, co-morbidity index, performance status, procedure type, stage. Seventy per cent of people with histologically confirmed SCLC were treated with chemotherapy however this varied according to several factors including the referral method and socioeconomic status. Survival after chemotherapy for people with SCLC in the NLCA was similar to that reported in clinical trials.

Conclusions

I have used routinely collected clinical data to address important questions surrounding the aetiology and treatment of lung cancer. The work in thesis provides evidence to support the growing body of work suggesting that women

are at higher risk of lung cancer per quantity of cigarettes smoked, and challenges the commonly held belief that COPD is a strong independent risk factor for lung cancer.

I have used the NLCA-HES linked data to add to our knowledge of the validity of treatment records in the NLCA, to produce a new predictive score for early mortality following lung cancer surgery which is now being validated in more than one independent dataset and to provide the oncology community with information on real-life treatment decisions and associated outcomes for small cell lung cancer.

Qualitative analyses of patient and clinicians attitudes, new data linkages, and information on organisational level variables are highly important in the next stages of research into inequalities in lung cancer care, and several studies are ongoing as a result of the research in this thesis.

ACKNOWLEDGEMENTS

I am grateful to Dr Barbara Iyen-Omofoman for her help in getting started with the THIN data and to Dr Tricia McKeever for her patience in answering all my basic Stata questions in the first few months. I would also like to thank Dr Anna Rich whose MD thesis and Royal College of Physicians fellowship were the foundations of this work and whose experience of her Royal College of Physicians fellowship attracted me to apply for the post in the first place.

My clinical research fellowship was funded by the Nottingham Respiratory Research Unit (NRRU) through the National Institute of Health Research and I am grateful to Professor Alan Knox (Director, NRRU) for this opportunity.

I am also grateful to Dr Vanessa Potter for her oncology expertise, and the members of our PhD steering group (Professor Anne Tattersfield, Professor Mick Peake, Dr Paul Beckett, Dr Anna Rich, Dr Catherine Free, Dr Roz Stanley, Dr Laura Jones, Jesme Fox, Aamir Khakwani, Mr John Duffy, John Southall, Dr Manpreet Bains, Dr Sadia Anwar, Dr Mat Callister and Dr Ian Woolhouse) who have provided ideas for new lines of research and opportunities to present my results to an ever expanding group of highly respected lung cancer clinicians and researchers. In particular I would like to acknowledge the contribution of Dr Roz Stanley formerly of the Health and Social Care Information Centre without whose hard work and patience I would not have had any NLCA data to work with.

I would like to mention my husband Tom and thank him for generally being great and continuing to work nights and on calls while I adopted more of a student lifestyle! Finally, I am extremely grateful to my PhD supervisors Professors Richard Hubbard and David Baldwin and Dr Laila Tata for their help and advice, encouragement and support, throughout my time as a clinical research fellow. I genuinely couldn't have asked for a better team of supervisors and am sorry to be leaving you a little sooner than planned!

PUBLICATIONS ARISING

Abstracts of work in this thesis which I have presented at conferences can be found in Appendix A. Papers published in peer reviewed journals are listed here:

Powell HA, Iyen-Omofoman B, Hubbard RB, Baldwin DR, Tata LJ. The Association Between Smoking Quantity and Lung Cancer in Men and Women. **Chest** 2013;143(1):123-9.

Powell HA, Iyen-Omofoman B, Baldwin DR, Hubbard RB, Tata LJ. Chronic Obstructive Pulmonary Disease and Risk of Lung Cancer: The Importance of Smoking and Timing of Diagnosis. **Journal of Thoracic Oncology**. 2013;8(1):6-11

Powell HA, Tata LJ, Baldwin DR, Stanley RA, Khakwani A, Hubbard RB. Early mortality after surgical resection for lung cancer: an analysis of the English National Lung cancer audit. **Thorax.** 2013;68(9):826-34.

Powell HA, Tata LJ, Baldwin DR, Potter VA, Stanley RA, Khakwani A, Hubbard RB. Treatment decisions and outcomes in small cell lung cancer. **British Journal of Cancer**. In press December 2013.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFO	Airflow Obstruction
AHD	Additional Health Data
AIDS	Acquired Immune Deficiency Syndrome
ASA	American Society of Anaesthesiologists
A1AT	Alpha-1 Antitrypsin (deficiency)
CHART	Continuous Hyper-fractionated Accelerated Radiotherapy
COAD	Chronic Obstructive Airways Disease
COPD	Chronic obstructive pulmonary disease
CCI	Charlson co-morbidity index
CI	Confidence interval
CT	Computerised Tomography
CVS	Cardiovascular system
DLCR	Danish Lung Cancer Registry
ECG	Electrocardiogram
EGFR	Epidermal Growth Factor Receptor
EPASS	Estimation of Physiologic Ability and Surgical Stress
EPIC	Epidemiology and Pharmacology Information Core
ERS	European Respiratory Society
ESOS	European Society Objective Score
ESSS	European Society Subjective Score

FEV1	Forced expiratory volume in 1 second
GP	General Practitioner
Gy	Gray
HCP	Healthcare professional
HR	Hazard ratio
HES	Hospital Episodes Statistics
HIV	Human Immunodeficiency Virus
HQIP	Healthcare Quality and Improvement Partnership
HSCIC	Health and Social Care Information Centre
IASLC	International Association for the Study of Lung Cancer
ICD-10	International Classification of Diseases - Revision 10
ILCOP	Improving Lung Cancer Outcomes Project
InPS	In Practice Systems
IQR	Interquartile range
LSOA	Lower Super Output Area
m	Metres
MDT	Multidisciplinary team
MRC	Medical Research Council
NAEDI	National Awareness and Early Diagnosis Initiative
NHS	National Health Service
NICE	National Institute for health and Clinical (or Care) Excellence
NLCA	National Lung Cancer Audit

NSCLC	Non-small cell lung cancer
NUH	Nottingham University Hospitals
ONS	Office of National Statistics
OPCS	Office of Population Censuses and Surveys
OR	Odds ratio
PET	Positron Emission Tomography
POSSM	Physiological and operative severity score for the enumeration of mortality and morbidity
PS	Performance status
QOF	Quality and Outcomes Framework
SACT	Systemic Anti-Cancer Therapy
SBRT	Stereotactic Body Radiotherapy
SCLC	Small cell lung cancer
SD	Standard deviation
SEER	Surveillance, Epidemiology and End Results
SNoMed	Systematised Nomenclature of Medicine
THIN	The Health Improvement Network
TNM	Tumour Nodes Metastases
TKI	Tyrosine kinase inhibitor
UICC	Union Internationale Contre Le Cancer
UK	United Kingdom
US	United States

VALSG Veterans' Administration Lung Study Group

VATS Video assisted thoracic surgery

CHAPTER 1: INTRODUCTION

This chapter covers the evolution of our knowledge about risk factors for lung cancer, some definitions for the medical and organisational terminology which will be used in this thesis, and a brief overview of existing treatments and how these affect survival. I will also discuss overall survival from lung cancer, the inequalities which are known to exist in lung cancer care, and current strategies to reduce the mortality burden and improve survival for people with lung cancer in the United Kingdom.

This is followed by the thesis justification, objectives and an outline of subsequent thesis chapters.

1.1 Incidence and risk factors

Lung cancer is the second commonest cancer in the UK, after breast cancer, with an estimated 42,000 new cases diagnosed in 2010.(1) Worldwide, lung cancer is the most common cancer with approximately 1.61 million new cases diagnosed in 2008. Incidence rates are highest in Europe and Northern America and lowest in parts of Africa.(2)

1.1.1 Cigarette smoking

When deaths from lung cancer started to increase in the UK at the start of the 20th century, scientists and doctors initially attributed this to increases in air pollution due to industrialisation and increasing road traffic. Exposure to gas during the war and even a recent influenza pandemic were also suggested as potential causes for the rising incidence of this disease. In the late 19th century, however, a machine had been developed which rolled cigarettes, making them widely available and much more affordable. The prevalence of smoking had risen sharply, initially in men, and particularly in the armed forces, and this was followed about 20 years later by a dramatic increase in the incidence of lung cancer.

Epidemiological studies first made the link between smoking and lung cancer in the 1930s, but it was not widely accepted, even by clinicians (many of whom enjoyed smoking cigarettes), until at least the 1950s. A landmark UK study in this respect was the British Doctors Study by Doll and Hill who collected information on smoking habits and cause of death for a cohort of British doctors in the 1930 -40s. They published their findings in 1956, providing evidence of the marked increase in incidence of lung cancer since the increase in cigarette consumption and of the increased mortality from lung cancer in smokers compared with non-smokers, and in heavy smokers compared with lighter smokers. In addition they reported that upper respiratory tract and upper

gastrointestinal malignancies as well as coronary thrombosis appeared to be more common in smokers than non-smokers although numbers with this outcome at that time were small; smoking is now a well-established risk factor for all of these diseases. (3)

Overall trend in smoking prevalence

Figure 1-1 shows the trend in smoking prevalence in the United States from 1900 to 1998 and some of the measures introduced to encourage smoking cessation.(4) Smoking prevalence in other developed countries followed a similar pattern and trends in Great Britain since 1950 are shown in Figure 1-2.

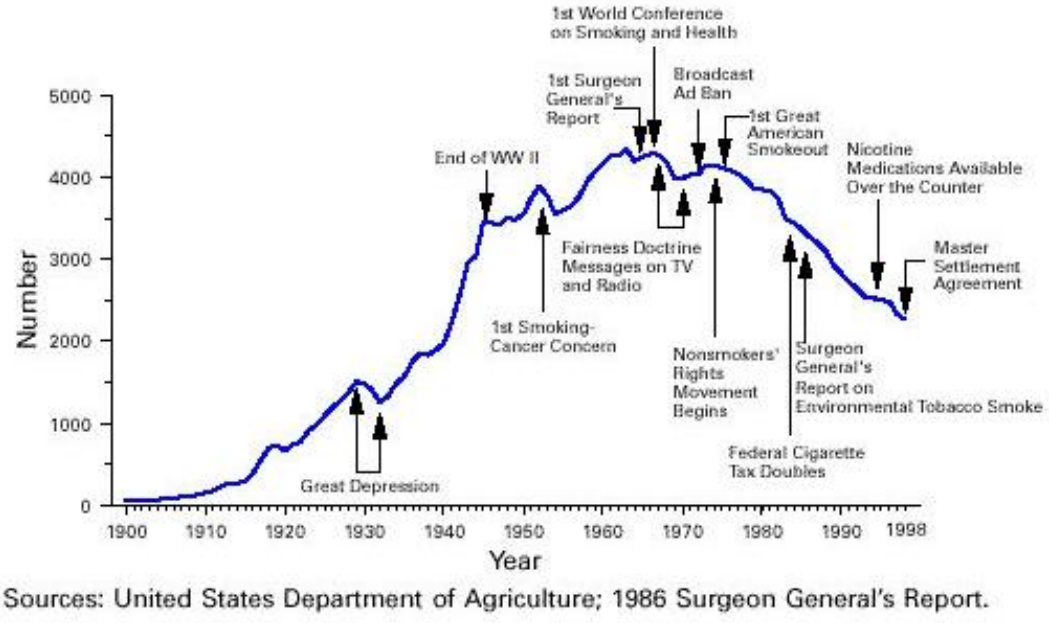


Figure 1-1: Annual adult per capita cigarette consumption and major smoking and health events United States 1900-1998 (4)

Smoking in women

By the time Doll and Hill, and other clinicians, began to report the strong association between cigarette smoking and lung cancer, tobacco advertising had taken off and the prevalence of cigarette smoking in UK males peaked at approximately 65% in the early 1940s.(5) When smoking prevalence in men

eventually started to decline, tobacco companies targeted women with advertising campaigns such as those for *Virginia Slims*.(6) The prevalence of smoking in women in the UK increased to a peak of approximately 45% in the 1960s (Figure 1-2). In some European countries and particularly in developing countries the prevalence of smoking in women is now higher than that of men.(7)

In the United Kingdom (UK) both smoking and lung cancer are still more common in men, however the ratio is falling and compared with 39:10 in 1975 the ratio of lung cancer in men compared with women is now 12:10.(5)

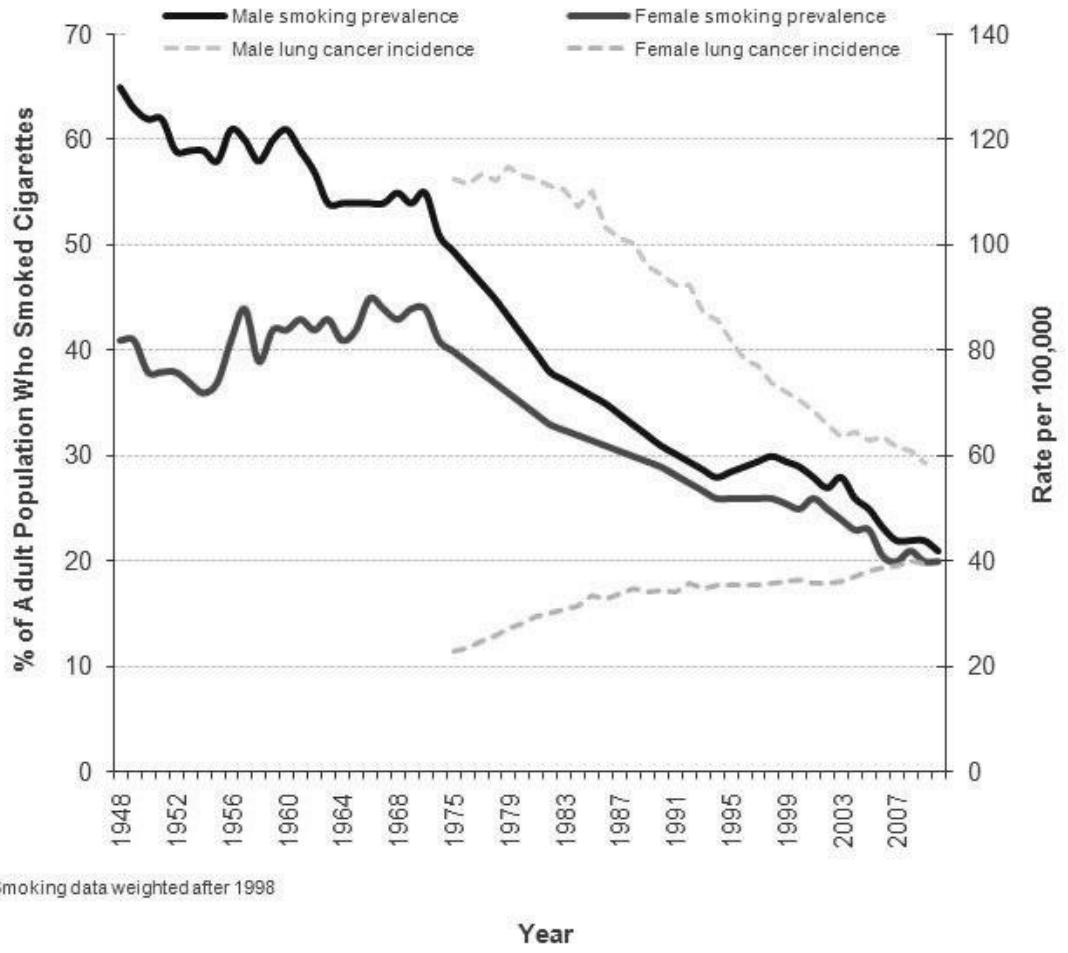


Figure 1-2: Smoking prevalence and lung cancer incidence, by sex, Great Britain, 1948-2010 (5)

Tobacco control and changes in incidence

The risk of lung cancer reduces significantly in people who stop smoking before middle age, (8) and tobacco control has been the single biggest factor to date in reducing the number of deaths due to lung cancer. Figure 1-2 demonstrates the peaks in incidence of lung cancer in men and women and how this relates to changes in smoking prevalence. The incidence of lung cancer in men peaked in 1980 but in women it continues to rise.

Financial burden

Lung cancer remains the second most common cancer in the UK and the financial burden is considerable with the estimated cost to the UK economy of £2.4 billion each year, £9,071 per patient annually, which is far higher than the cost of any other cancer despite survival rates being among the lowest. (9)

1.1.2 Other risk factors

There are many reported risk factors for lung cancer. Radon gas and occupational exposure to substances such as asbestos are well established as causes of lung cancer, particularly in smokers. It has also been suggested that lung cancer is more common in people with other chronic lung diseases such as pulmonary fibrosis and chronic obstructive pulmonary disease (COPD), even after accounting for smoking. (10, 11)

A history of lung cancer in a first-degree relative is associated with a two-fold increased risk of lung cancer regardless of smoking status and suggests the possibility of a hereditary predisposition to lung cancer or shared environmental risk factor exposure by members of the same family. The increased risk in individuals less than 60 years of age who have a first degree relative diagnosed with lung cancer at less than 60 years has been found to be five-fold. (12, 13) There is, however, the possibility of ascertainment bias here in that people may be more aware of the symptoms of lung cancer and potentially more likely to

present to their doctor if they have seen a relative with the disease, but also because people may not know what their relatives died of.

1.2 Classification of lung cancer

1.2.1 Histology

Lung cancer can broadly be divided into small-cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Malignant mesothelioma is another tumour which affects the thoracic cavity and is strongly associated with asbestos exposure. Whilst the pathological features, treatment and prognosis for NSCLC and SCLC are discussed in this chapter, mesothelioma is not within the scope of this research and will not be covered. Data for patients with a known diagnosis of mesothelioma are excluded from the studies in this thesis and where the term 'lung cancer' is used this does not include mesothelioma.

Non small-cell lung cancer (NSCLC)

The majority (more than 80%) of lung cancers are NSCLC, and most are of the adenocarcinoma, squamous cell carcinoma or large cell subtypes, with adenocarcinomas recently having overtaken squamous cell as the most common subtype. Adenocarcinoma is much more common than squamous cell carcinoma in non-smokers, (14) but both are still more common in smokers than in non-smokers.

Small-cell lung cancer (SCLC)

Small-cell lung cancer accounts for between 10% and 18% of all lung cancer, and almost always occurs in smokers. The incidence is declining as a result of decreasing prevalence of cigarette smoking.

Small cell lung cancer is so termed because of the microscopic appearance of the tumour cells which, in comparison to NSCLC cells, are small. These tend to be rapidly dividing tumours which frequently results in metastases being present at the time of diagnosis. (15)

1.2.2 Stage

The extent of disease for any tumour is described as the stage; until recently staging systems in lung cancer have differed according to the tumour type.

The extent of disease in patients with NSCLC is described using the Union Internationale Contre Le Cancer (UICC) tumour, node, metastasis (TNM) staging system which assigns a stage between I and IV depending on the size of tumour and any invasion into other structures within the chest (tumour or 'T' stage), the location of any lymph nodes which are affected by the cancer (nodal or 'N' stage), and the presence or absence of spread to distant structures (metastatic or 'M' stage). (16)

The staging system used for SCLC was, until recently, that described by the Veterans' Administration Lung Study Group (VALSG) as 'limited' or 'extensive' depending on whether the full extent of the disease is confined to one side of the chest and could be captured in a single radiotherapy field (limited stage if this would be technically possible, extensive stage if not).(17)

Research has suggested that further classifying SCLC by the UICC TNM staging system used for NSCLC may improve the accuracy with which outcomes and treatment response (particularly from radiotherapy) can be predicted, and therefore current recommendations are that TNM staging is also used for SCLC. (14, 18) If conversion is necessary, limited disease broadly includes T1-4, N0-3, M0 and extensive disease includes T1-4, N0-3, M1a/b in the updated TNM staging classification.

1.3 Treatment options and their effect on survival

The most effective intervention in reducing the number of deaths due to lung cancer has been the promotion of smoking cessation, (8) but for those who have already developed the disease treatment options depend on how far the disease has spread (the stage) and the general condition of the patient. Common treatments for NSCLC are described below; small-cell lung cancer behaves, and is therefore treated, differently and the management is described under a separate sub-heading.

1.3.1 Non-small cell lung cancer

Surgery

Patients with NSCLC who present at an early stage (stage I-IIIA) may be suitable for surgical resection. The most common type of surgical procedure is a lobectomy (Figure 1-3), which removes the affected lobe of the lung along with the blood vessels and the lymphatics. The aim is to entirely remove the tumour so that the patient is cured and surgery therefore offers the biggest improvement in survival of all treatments for NSCLC. For patients with early stage disease, 5-year survival has been reported as up to 73% following tumour resection. (19) If the tumour is too big or mediastinal lymph nodes are affected a bi-lobectomy or pneumonectomy may be performed (Figure 1-3). If the tumour has spread outside the lung, or to lymph nodes on the opposite side of the chest then it is not possible to completely resect the tumour; these cases are sometimes described as not being 'resectable'.

Due to advances in preoperative staging it is quite unusual in current practice to find at the time of operation that the tumour has spread to the extent that it is not resectable. This still occurs occasionally and these procedures are termed 'open and close' as usually nothing is removed. Even with sophisticated pre-operative staging techniques some patients will develop a recurrence of the

cancer very soon after surgery at the site of the resection or as metastases in the lymph nodes or other organs. In these patients it is highly probable that micro metastases were present at the time of the operation and unfortunately the surgery was not successful at removing the entire tumour. These patients will have a considerably poorer prognosis than those for whom surgery is successful.

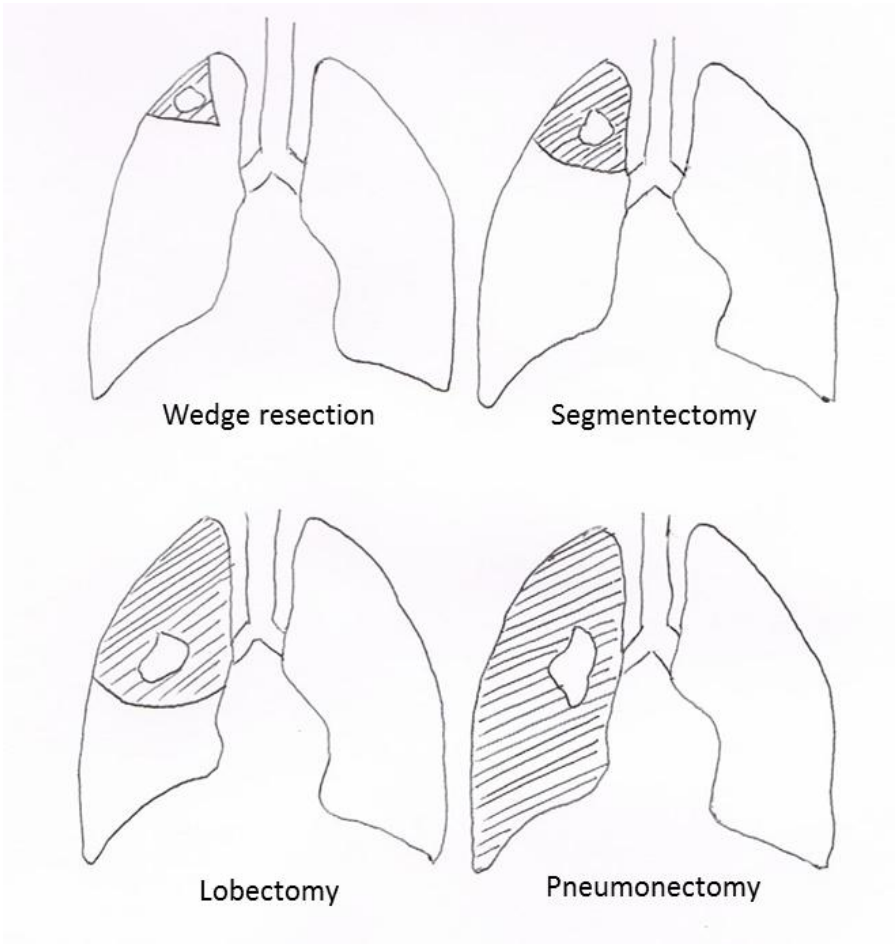


Figure 1-3: Types of surgery for lung cancer

Post-operative complications may occur immediately or later on after surgery. Immediate complications include bleeding, leakage of air from the area where the lung was removed preventing the lung from re-inflating, infection, pulmonary emboli and cardiac complications. In some cases these complications can be fatal and some patients are considered too high risk for major thoracic surgery due to

co-existing health problems. This is sometimes described as whether or not the patient is 'operable', and is the subject of Chapter 6 of this thesis. In some of these cases a less complicated procedure such as a wedge resection or segmental resection (segmentectomy - Figure 1-3) may be performed in an attempt to cure patients who are not considered fit enough for a lobectomy. Wedge resections are sometimes performed for both diagnostic and treatment purposes in patients who have small lung lesions (termed nodules) confined to a small area of lung, however there is concern that a wedge resection is more likely to leave residual tumour compared with a segmentectomy or lobectomy.

Radiotherapy

If surgery is not an option, or the patient's preference is for a less invasive treatment, 'radical' radiotherapy may be given with curative intent. The unit for measuring dose of ionising radiation is Gray (Gy) and the total dose is usually spread out over time (fractionated) to allow normal cells to recover. Common side effects of radiotherapy are due to damage to normal tissue and when radiotherapy is given to the chest they include fatigue, breathlessness, cough and inflammation of the oesophagus. There is also a risk of damage to the spinal cord if the tumour is close to the vertebrae.

One form of radical radiotherapy is continuous hyper-fractionated accelerated radiotherapy (CHART). A randomised controlled trial which compared CHART (36 fractions of 1.5 Gy radiotherapy 3 times per day to give 54 Gy in 12 consecutive days) with conventional radiotherapy (30 fractions of 2 Gy to a total dose of 60 Gy in 6 weeks) reported a 2-year survival for people with locally advanced NSCLC treated with CHART or radiotherapy of 29% and 20% respectively. (20)

Stereotactic body radiation therapy (SBRT) has recently been introduced to UK practice. This involves a very high dose radiation delivered to a well-defined area of lung. It is only suitable for certain small tumours which are an appropriate distance from a major airway. At present, SBRT is only used for patients with

NSCLC who decline, or who are not felt to be fit enough for, thoracic surgery. Local control rates as high as 90% have been reported in some studies however it difficult to compare outcomes with surgery when there is such a difference in fitness between the patients who undergo each treatment.(21)

Chemotherapy

In NSCLC, chemotherapy is not usually given curative intent; it would usually be considered a palliative treatment to improve symptoms and quality of life, unless given as an adjunct to surgery or induction prior to radiotherapy (see below). It is used for people who have advanced disease which is not resectable and cannot be radically treated with radiotherapy, including people who have previously been treated and the tumour has recurred. Patients must be fit enough to withstand the potential side effects and toxicities.

Common chemotherapy regimens for NSCLC involve a platinum agent (Cisplatin or Carboplatin) combined with another drug such as Paclitaxel, Etoposide, Vinorelbine, Docetaxel, Gemcitabine or Pemetrexed. Recent evidence suggests that poorer outcomes are seen if Pemetrexed is used for squamous cell lung cancer and therefore the National Institute for Health and Clinical Excellence (NICE) recommends that Pemetrexed is only used for non-squamous tumours. (22) This is also referred to as targeted chemotherapy. The side effects of these drugs vary but may include gastro-intestinal disturbance, fatigue, hair loss, mouth ulcers and hearing loss. There is also a risk of life-threatening bone marrow toxicity which may cause a severe deficiency of white blood cells (neutropenia) or platelets which can in turn lead to life-threatening infections or bleeding.

Whilst treatment with chemotherapy alone will not cure patients or even have much of an effect on medium to long-term survival, it may prolong survival in the short term. Estimates for absolute increase in median survival compared with best supportive care suggest that this is in the region of 1.5 months (from

4.5 months to six months). (23) Second and third line chemotherapy may be given with improvements in survival for some patients.

Adjuvant and induction chemotherapy

Adjuvant chemotherapy is given following surgery. Studies have shown an absolute improvement in 5-year survival of 4% with adjuvant chemotherapy, (24) and current guidelines recommend that it is considered in patients with post-operative pathological stage II or higher and tumours greater than 4cm maximum diameter. (14, 25)

Chemotherapy given with the intention of down-staging the tumour prior to conventional radical radiotherapy is sometimes termed induction chemotherapy. Meta-analyses suggest that this improves 2-year survival by 4–7%. (26)

Biological therapy

Another form of targeted therapy is biological therapy, a relatively recent development in systemic therapy for NSCLC, although this has been a part of treatment of other solid and haematological tumours for some time. The first pathway to be exploited was a mutation on the Epidermal Growth Factor Receptor (EGFR) gene; the growth of tumours with this mutation can be slowed by tyrosine kinase inhibitors (TKIs). Erlotinib and Gefitinib are the TKIs currently recommended for use in the UK. (27, 28)

Studies of Erlotinib suggest a 2 month overall survival benefit when compared with placebo in people with advanced NSCLC previously treated with chemotherapy, (29) and in the UK it is now recommended for first line use in patients with EGFR mutations.(27) Treatment of advanced pulmonary adenocarcinoma in non-smokers or former light smokers with Gefitinib resulted in 12-month progression-free survival of 24.9% compared with 6.7% in those treated with carboplatin–paclitaxel. (30) Gefitinib is now approved by NICE as an alternative first line treatment for EGFR mutation positive tumours.(28)

The side effects of the TKIs are considerably less severe than those of standard chemotherapy regimens with the most common being a skin rash and diarrhoea. They also have the advantage of being orally administered in contrast to most other chemotherapy drugs which usually require the patient to attend a day-case unit for intravenous therapy.

1.3.2 Small-cell lung cancer

Small cell lung cancer is usually a very aggressive tumour and often presents at an advanced stage; the majority of patients die due to systemic disease. Surgery is not usually considered to be an option as the tumour has spread too far by the time of diagnosis. The rapidly dividing nature of the tumour cells does, however, mean that the sensitivity to radiotherapy and chemotherapy is often better than in NSCLC.

Chemotherapy

Initial studies of chemotherapy in SCLC in the late 1960s using cyclophosphamide found a modest improvement in survival compared with placebo.(31) It was soon discovered that a combination of a platinum based agent and another active agent produced a much greater survival benefit, (32) and that administering these drugs at the same time rather than sequentially produced the best results.(33) Current practice is therefore to use combination chemotherapy, administered simultaneously. The management of small cell lung cancer has not really changed over the past decade and there are few clinical trials in progress.

Given there is strong evidence that platinum based chemotherapy improves survival compared with non-platinum based regimes,(34) the most commonly used chemotherapy regimens for SCLC are Etoposide with either Carboplatin or Cisplatin (two dose options for Cisplatin are 60 or 80 mg/m^2 depending on patient fitness and co-morbidities). Carboplatin and Cisplatin must be

administered intravenously and each dose takes approximately 1 hour to administer. The side effects and toxicities are similar to those described above for chemotherapy in NSCLC. Dose reductions may be necessary for patients with liver or renal impairment. The platinum agent is given on day 1 and Etoposide on days 1, 2 and 3. Etoposide can be given orally as well as intravenously however bio-availability is better if it is given by the intravenous route; often the first dose of each cycle is intravenous and the two subsequent doses oral so the patient doesn't have to attend hospital on so many occasions.

It would be usual to give a 'cycle' (a full dose of both drugs on day 1 and of Etoposide on days 2 and 3) every 3 weeks. The number of cycles given depends on whether the patient experiences intolerable side effects or toxicity, but the aim would be to complete 3-4 cycles before assessing whether there has been a response to treatment. If there is a good response radiologically and clinically, a total of 6 cycles are given. If there is no improvement, or if there is an increase, in the burden of disease, the risks of further chemotherapy are felt to outweigh the benefits and no further cycles will be given. Once 6 cycles of chemotherapy have been given the patient is usually followed up every 2-3 months for evidence of progression of disease. In some cases further cycles of chemotherapy are given if the disease progresses.

This sort of chemotherapy treatment is sometimes referred to as 'palliative' because it is not given with the aim of curing the patient. The aim is to reduce the disease burden and therefore the patient's symptoms, and in addition increase life expectancy. Both limited and extensive stage disease can be treated with chemotherapy however patients with extensive stage disease are often more frail than those with limited disease and frequently may not tolerate aggressive chemotherapy due to poor performance status and multiple co-morbidities.

The results of clinical trials show that median survival can improve to 8-12 months for people with extensive stage disease, (35-37) and up to 2 years for those with limited stage disease, particularly when combined with radiotherapy (see below).(35, 38, 39)

Radiotherapy

For patients with limited stage SCLC (disease confined to one half of the chest without any distant metastases), there is good evidence that radiotherapy combined with chemotherapy improves survival when compared with chemotherapy alone with estimated median survival between 18 and 24 months. (38, 40) In a few cases chemo-radiotherapy can lead to long term survival akin to cure, however the risks of toxic side effects (pneumonitis, oesophagitis and neutropenic sepsis) are increased when both treatment modalities are used.(41, 42)

Radiotherapy can either be given at the same time as chemotherapy (usually with the first or second cycle) or after 4-6 cycles of chemotherapy have been completed; these are termed concurrent and sequential chemo-radiotherapy respectively. Radiotherapy would usually be given every weekday for 3 weeks to a total of about 45Gy but the optimal dose and dose per treatment (termed fractionation) is still unclear.(43, 44) There is some evidence to suggest that concurrent chemo-radiotherapy confers a small survival advantage over sequential chemo-radiotherapy, (38, 45) although the former is associated with an increased risk of toxic side effects. Some studies, including one in the UK, (46) have failed to show a benefit of early over late radiotherapy, however this may be due to inadequate or incomplete chemotherapy doses in some patients in these studies. In patients who have good performance status and limited stage disease current recommendations are for concurrent chemo-radiotherapy. (14, 43, 44)

Most of the above evidence is based on studies of patients with limited stage disease. Some patients with extensive disease respond very well to chemotherapy and in these cases thoracic radiotherapy is considered if there is complete response at distant sites and a good response in the thorax.(14)

Prophylactic cranial irradiation

Small cell lung cancer grows and spreads rapidly and distant metastases are often evident at the time of presentation. Micro-metastases may also be present outside the lung in patients who appear to have limited stage disease on imaging. For this reason patients whose disease burden is reduced, or does not progress, after first-line treatment are offered prophylactic cranial irradiation with the aim of preventing or delaying the growth of brain metastases.(14) This has been shown to confer an overall survival advantage in patients with extensive stage disease, (47) and a reduction in the incidence of brain metastases in limited stage SCLC. (48)

Disease recurrence

In a few cases of, usually limited stage, SCLC the disease becomes undetectable both clinically and radiologically after chemo-radiotherapy, however patients are followed up as it usually recurs. There are insufficient data from studies of second-line chemotherapy to determine the most effective second-line chemotherapy agents, but response is dependent on the response to first-line therapy, time interval since finishing first-line treatment, residual toxicity and performance status.(43) Symptomatic patients who are unlikely to benefit from second-line chemotherapy are considered for palliative radiotherapy.

1.3.3 Palliative care

Supportive, symptom based, or palliative care is an important part of the management of patients with any cancer, and is particularly important in lung cancer given the large number of patients with advanced disease at presentation

for which there is no possibility of long term survival. Palliative care involves the management of symptoms such as cough, breathlessness, pain and haemoptysis and covers a range of interventions such as psychological support for patients and their families, local radiotherapy for haemoptysis or bone pain, and morphine infusions to alleviate the symptoms of pain or breathlessness at the end of life.

A randomised controlled study of early palliative care in patients with advanced lung cancer, published in 2010, showed improvements in quality of life, reduction in symptoms and even a survival benefit (11.6 months vs. 8.9 months median survival) in patients who received standard oncology care plus early palliative care compared with those who only received standard oncology care; this was despite fewer patients in the early palliative care group receiving aggressive end-of-life care. (49)

1.4 Structure of lung cancer care in the UK

Before discussing inequalities in lung cancer care and survival, and current strategies to improve these, I will briefly discuss the organisation of care for people with lung cancer in the UK and introduce some terminology which will be used later in this thesis when assessing the effects of organisational level factors on treatment and outcomes.

1.4.1 Primary care

General practice

All UK residents should be registered with a general practitioner and are entitled to free consultations and treatment (with the exception of fees for some vaccinations and prescriptions) paid for by the National Health Service (NHS). General practitioners (GPs) work in primary care practices; each practice is run by one or more GPs who are responsible for a proportion of people in their local area.

Presentation and referral

General practitioners manage chronic disease and should also be the first point of call for non-emergency new presentations. People with symptoms suggestive of cancer (of any site) should be referred urgently by their GP to the appropriate secondary care service; in the case of suspected lung cancer this is the lung cancer multi-disciplinary team (MDT, described below). Since the NHS Cancer Plan was published in 2000,(50) there has been a 2-week-wait system whereby patients referred with suspected cancer must be seen in secondary care within 2 weeks of referral.

This system relies on the ability of the GP to recognise the signs of lung cancer, organise the appropriate tests, and make the referral to secondary care. The National Institute for Health and Care Excellence (NICE) has produced guidelines to assist GPs with recognising and managing these patients and advises either a

chest radiograph followed by referral or immediate referral to the lung cancer team for certain high risk patients.(14)

1.4.2 Secondary care

Trusts

Secondary care in the UK is provided predominantly by NHS hospital trusts (sometimes termed 'NHS trusts' or 'hospital trusts') and is also free to patients at the point of access (although a few people pay or have medical insurance which pays for them to be seen and/or treated in the private sector). Whereas a hospital is generally a single secondary care facility on one physical site, an NHS trust is made up of one or more (although usually less than four) hospitals which are under the same management and usually in fairly close geographical proximity.

Cancer networks

From 2000 until the NHS reforms of 2012, NHS trusts were further grouped geographically in terms of cancer care into cancer networks. Networks consisted of between three and twelve NHS trusts, with the exception of the Welsh cancer network which comprised all 17 Welsh trusts. Clinicians representing the trusts in each network worked together with local primary care representatives and other NHS services to improve performance, facilitate communication and engagement around cancer issues, and deliver high quality, integrated cancer services for their populations.

Diagnostic pathway

Patients are usually referred to the lung cancer team by their GP after presenting in primary care with relevant symptoms and often having had a chest radiograph. Some patients, however, do not present to their GP and are identified from acute hospital admissions, emergency department attendances or consultations with consultants in other specialities. These cases are referred

directly to the lung cancer MDT without the need for the patient to consult their GP.

The first consultation in secondary care is almost always with a respiratory physician. In many trusts a computerised tomography (CT) scan is arranged prior to this consultation to expedite the diagnostic pathway. The respiratory physician usually arranges diagnostic tests and then the patient is discussed at the trust's lung cancer MDT meeting where a management plan is agreed.

The multi-disciplinary team

The lung cancer MDT should include one or more respiratory physicians with a special interest in lung cancer, a radiologist, histo-pathologist, lung cancer clinical nurse specialist, an oncologist who can either offer both radiotherapy and chemotherapy, a palliative care physician, a thoracic surgeon, and an MDT administrator. The team discusses all cases of lung cancer and agrees management plans which are then communicated to patients and implemented by the appropriate member of the team. Some patients require repeated discussions after additional information or investigations are obtained.

Availability of services

Most trusts have respiratory physicians and radiologists at one of their hospitals (although not all individual hospitals will have them on-site), however services such as thoracic surgery are not available at all trusts. Approximately 32 of the 162 NHS trusts in England have a thoracic surgery service and therefore the thoracic surgical representative on the MDT for the majority of trusts will be employed and operate at another trust. The same is true for chemotherapy with some trusts referring patients elsewhere for treatment; this means that patients may have to travel quite long distances for pre-operative assessments and / or treatment.

1.5 Survival and inequalities in lung cancer

1.5.1 Lung cancer survival in the UK

Over 70% of patients with lung cancer in England present at an advanced stage when cure is not possible, and thus the prognosis is poor.(51) English data from the ONS show that for people who were diagnosed with lung cancer from 2005 to 2009 who were followed up until 2010, survival at one-year was 29% for men and 33% for women, and just 8% and 9% respectively at five years.(52) Five year survival from lung cancer in the UK has not changed substantially over the last 20 years.

Five-year survival from lung cancer is extremely poor compared with other common cancers. In the UK, 5-year survival from breast and bowel cancer are currently estimated at 85% and 54% respectively which means that although lung cancer is only the second most common tumour type overall (excluding non-melanoma skin cancer), more people die each year from lung cancer than from breast, bowel and prostate cancer put together.

1.5.2 International differences in survival

People with lung cancer in other countries in Europe, and North America seem to have considerably better 5-year survival than the UK. Coleman and colleagues reported age standardised 5-year survival based on UK cancer registries to be 8.8% for people diagnosed in 2005 to 2007, compared with 18.4% and 17.0% from Canadian and Australian registries respectively; the figures for Denmark and Norway were more similar to the UK but still higher at 10.9% and 14.4%. One-year survival for people diagnosed in the same time period followed a similar pattern with 29.7% of the UK lung cancer population surviving a year after diagnosis compared with 42.8% in Australia, 43.1% in Canada, 34.9% in Denmark and 39.2% in Norway.(53)

Reasons for these considerable differences are unclear but possible explanations include differences in the way that lung cancer cases and survival statistics are reported (i.e. who is included in the case definition), or true differences in survival due to earlier diagnosis or more aggressive investigation and treatment (for example higher histological verification and resection rates) in these countries.(54)

1.5.3 Inequalities in treatment rates

Histological confirmation

Radical treatment is not usually completed without histological confirmation, although surgery is sometimes offered without prior histological confirmation where there is a high probability of cancer. The few patients who have radical treatment without histological confirmation are those treated with radical radiotherapy where biopsy is considered to be high risk. A study using the large EUROCARE database showed that England had the lowest rate of histological verification of lung cancer cases of all the European countries included. (54) It has also been shown that histological confirmation rates vary between National Health Service (NHS) trusts, even after accounting for patients' age and performance status. (55)

Whilst some of the difference, particularly for international comparisons, may be due to differences in data collection and whether or not the population is nationally representative, it is also possible that individual clinician and patient attitudes in different institutions as well as potential differences in stage at presentation affect how aggressively a tissue diagnosis is pursued.

Surgical resection rates

It is estimated that 13.7% of all patients with lung cancer in the UK had surgery in 2010.(56) Substantially higher resection rates were reported in other

countries including the US (27%) and a region of the Netherlands (20%). A study using the EUROCARE database confirmed that the proportion of patients receiving surgery was higher in Switzerland, France and The Netherlands, than in the UK and Spain. (57-59)

Once more, this variation could partly be explained by the denominator used by each country but resection rates also vary between NHS trusts in the UK (Figure 1-4). Using the National Lung Cancer Audit (NLCA), Rich and colleagues found that people were 51% more likely to have surgery if they were first seen at a trust where there was a thoracic surgery service on site, (60) and Brown et al reported that people with lung cancer identified from the Southend Lung Cancer Registry were more likely to have had treatment (surgery or chemotherapy) if they were seen by an accredited chest physician at some point in their diagnostic pathway. (61)

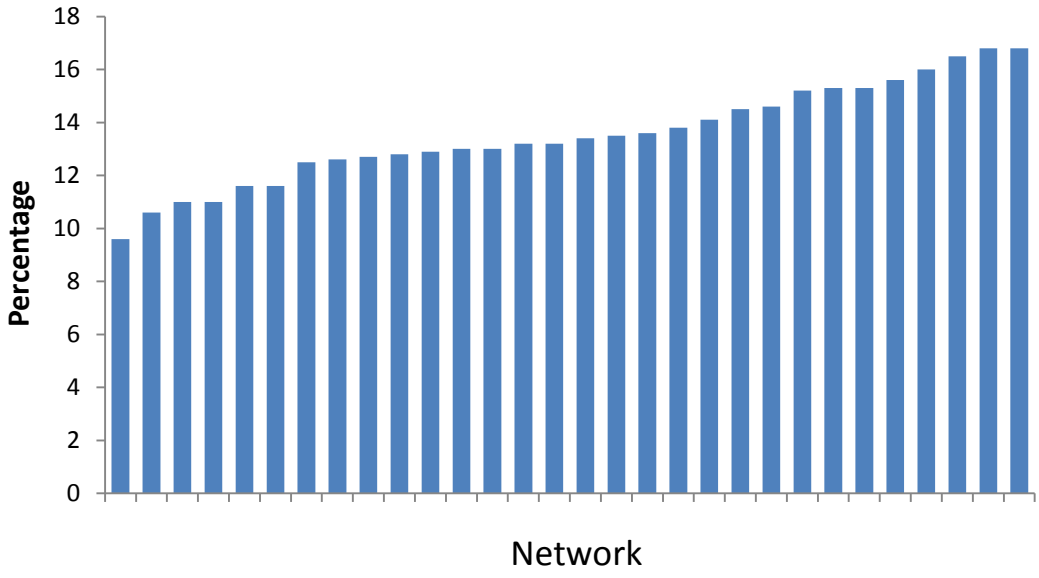


Figure 1-4: Proportion of cases of lung cancer resected in 2010 by English cancer network (Source: NLCA)

Given the dramatic improvements in survival after surgery for those who are eligible (section 1.3.1), an understanding of the reasons for the variation in

resection rates and whether any of these are modifiable, is important in the drive to improve survival.

1.6 Current strategies to reduce lung cancer mortality and improve survival

The promotion of smoking cessation remains the most important factor in reducing the incidence of lung cancer, and thus the annual mortality burden. Strategies have included legislation to control the sales and marketing of tobacco products, restrictions on smoking in public areas, and mass media campaigns to educate the public on the health risks of smoking. (62)

Several important initiatives have been developed with the aim of improving survival for people with lung cancer. These centre on earlier diagnosis so that more people can be treated with curative intent.

1.6.1 Screening

The potential benefits of screening for lung cancer are that tumours will be detected at an earlier stage when curative treatment is more likely to be possible, and when patients are potentially more able to withstand these treatments. The disease is a huge public health burden, computerised tomography (CT) scanning is a highly sensitive test, and there are several well established treatments meaning that lung cancer meets the major criteria for screening. (63) The target population would consist of smokers and ex-smokers between the ages of approximately 40 and 75. Older patients have not been included in screening trials because they are often not suitable for curative treatment and thus earlier detection would not confer a benefit; the upper age limit is, however, debatable.

The drawbacks of screening programs are that they are expensive to set up and run, and that the necessary radiation from a screening CT scan poses a risk of future malignancy for the patient. A CT scan is not specific for lung cancer and

benign nodules, infections and other inflammatory processes may be detected in the process of screening, which require further investigation and treatment; this would add to the costs of a screening program. Some tumours may be very slow growing and resection or radiotherapy, particularly in patients with other co-morbidities, may not actually improve survival. There is also concern that the patients who would be likely to attend for screening are not those at highest risk of lung cancer. This problem is partly related to the higher incidence of lung cancer in more deprived areas, and also to beliefs that it is a self-inflicted disease not worthy of treatment, or that treatments are futile.(64, 65)

A large randomised controlled trial in the US recently reported that CT screening was associated with a 20% relative reduction in mortality from lung cancer, and a 6.7% reduction in all-cause mortality, when compared with chest x-ray screening. (66, 67) A European trial and a pilot study in the UK have been performed and are currently in the follow-up stages. (68, 69)

1.6.2 Early diagnosis initiatives

Further strategies to increase the proportion of patients eligible for radical treatment include increasing public and general practitioner awareness of lung cancer. The National Awareness and Early Diagnosis Initiative (NAEDI) was set up in 2007 by the Department of Health and Cancer Research UK.(70) The project covers interventions across several tumour sites but for lung cancer it has included a mass media 'three-week-cough' campaign encouraging people to see their general practitioner (GP) if they have had a cough for more than 3 weeks. Following the intervention there was evidence of an increase in unprompted awareness of cough and hoarseness (41% to 50%), and persistent cough (12% to 15%), as symptoms of lung cancer in the target audience (people over 50 years old from deprived areas). (71) There was also an increase of approximately 30% in two week wait referrals for suspected lung cancer in the campaign months. The campaign has now been implemented nationwide and

changes in stage at presentation before and after the campaign are being assessed.

A large proportion of patients with lung cancer are diagnosed as the result of an emergency admission rather than outpatient referral to the lung cancer team, and these patients have poorer outcomes. (56) There is therefore pressure on GPs to identify people with lung cancer earlier, particularly as survival is better in countries who do not have the primary care system where patients have direct access to secondary care. (72) Epidemiological analyses have been performed to describe the features and presenting symptoms in primary care of patients who go on to be diagnosed with lung cancer. (73, 74)

1.6.3 Identifying and addressing inequalities

Data collection

The NLCA collects data on people with lung cancer in the UK and produces an annual report so that NHS trusts can compare their rates of investigations and treatments with national averages. (56) The audit was set up with the aim of improving lung cancer survival and one way in which it aims to do this is by reducing inequalities in care through publication of these reports. The NLCA is described in detail in Chapter 2.

The European Initiative for Quality Management in Lung Cancer Care was established in 2009 by the European Respiratory Society (ERS) with the aim of sustainably improving the quality of care for people with lung cancer in Europe. The ERS taskforce have piloted the European Lung Cancer Audit which is an international program of data collection to evaluate the provision of lung cancer care across Europe and survey the resources available so that inequalities can be identified without concern over differences in methods of data collection. (75)

Resection rates

There has been a drive to increase resection rates for NSCLC with particular focus on older patients and those with poorer performance status or multiple co-morbidities who anecdotally had not been offered surgery in the past due to concerns that they may be at high risk of perioperative mortality. It is important that patients are provided with the best possible estimate of their level of surgical risk, and the likely benefits, before a treatment plan is decided. There are tools available to assist clinicians with estimating perioperative mortality risks and these are discussed in detail in chapter 6; they do, however have recognised limitations.

In the drive to increase resection rates it is also important to understand what patients and clinicians think is an acceptable level of risk so that people are not denied this potentially life-saving treatment because of concerns that their risk is too high. Little work has been done on this to date and this will be discussed further in chapter 9.

Resection rates are slowly but steadily increasing, particularly in older patients, and it is predicted that this will save lives in the long term. (76) It is too soon, however, to see any clinically significant improvement in lung cancer survival in the UK. (77)

Multi-disciplinary Team performance

The lung cancer MDT was described in section 1.4. Almost all patients with lung cancer in the UK are discussed by a lung cancer MDT before their management plan is decided.(51) This in itself is expected to have reduced inequalities in lung cancer care compared with the previous system where a single clinician was responsible for deciding on a management plan, however variation still exists in the quality of MDT management of lung cancer,(51); the way that an MDT is run and the opinions of individual decision-makers may contribute to ongoing inequalities in investigation and treatment rates and consequently survival.

Acknowledging these differences, the National Cancer Peer Review programme was set up in 2004 to monitor the quality of cancer services in the UK. The programme involves self-assessments by individual MDTs and external reviews of teams by professional peers against nationally agreed quality measures (78).

A further initiative, the Improving Lung Cancer Outcomes Project (ILCOP) was set up in 2010 and involved members of MDTs from different institutions visiting neighbouring MDTs and providing feedback on how improvements could be made (79). This resulted in improvements in quality and efficiency of MDT working at the local level by providing individual recommendations for each institution (80).

1.7 Chapter summary

In this chapter I have described the main risk factors for lung cancer and how changes in the prevalence of smoking have affected lung cancer incidence and mortality. I have described the classification, staging and main treatment options for lung cancer and in so doing have introduced some medical and organisational terminology relevant to lung cancer care specifically in the UK. I have also described the effect of common treatments on survival and how survival in the UK is poor when compared with other developed countries. I have introduced the subject of inequalities in lung cancer care and some current strategies to reduce these and improve lung cancer survival in the UK.

This is now followed by the thesis justification, aims and objectives and an outline of subsequent thesis chapters.

1.8 Justification of thesis

The initial studies in this thesis use primary care data to study the start of the lung cancer patient's journey by investigating risk factors, specifically focusing on smoking. Knowledge and quantification of the effects of these factors on a person's risk of developing lung cancer are important for identifying people with the disease in a timely manner.

The majority of the work in this thesis follows on from work carried out by Dr Anna Rich and published in her MD thesis. (81) Dr Rich validated lung cancer cases in the National Lung Cancer Audit (NLCA) database, (82) developed a method to measure co-morbidity using linked data from the Hospital Episodes Statistics (HES) database, and then used records cases first seen up to the end of 2008 to provide evidence of inequalities in access to surgery and chemotherapy for people with NSCLC and SCLC respectively. (60, 83)

I will use the link between the NLCA and HES to examine and validate records of surgical and chemotherapy treatment in each database and determine the most accurate definition of each to be used in future studies. I will build on the work on surgery in NSCLC by investigating the factors associated with early mortality after potentially curative resection with the aim of producing a predictive score to aid patient selection in clinical practice. I will also build on the work on SCLC by using a more recent (and therefore larger) dataset to look for inequalities in chemotherapy treatment between trusts, and take the work on survival further by including the number of chemotherapy cycles received in the analysis.

Further rationale for each study is given at the start of each chapter.

1.9 Thesis objectives

Using data from The Health Improvement Network (THIN), Hospital Episodes Statistics (HES), the National Lung Cancer Audit (NLCA) and the Office for National Statistics (ONS) which are anonymous patient databases described in detail in Chapter 2, I set out to achieve the following objectives:

1. Investigate whether sex modifies the effect of quantity of cigarettes smoked on lung cancer risk.

2. Establish whether COPD is an independent risk factor for lung cancer.

3. Identify risk factors for early death after surgery for NSCLC

4. Identify factors associated with chemotherapy use in people with SCLC and how this affects survival

As a speciality registrar in respiratory medicine I also aimed to maintain and further develop the skills required to manage patients with lung cancer and to lead a successful lung cancer service. This includes an in depth knowledge and understanding of previous and current research and political influences in the

field, and how these influence current policies. This objective was met through a series of clinical tutorials, observation of clinical practice, and practical experience (Appendix B).

1.10 Outline of thesis sections

A short description of the work described in each chapter is given below:

Chapter 2: Description of databases - A description of the sources of routinely collected data used for this thesis, the populations studied, definitions of common variables, and some generic strengths and limitations of the data.

Chapter 3: Smoking quantity and lung cancer in men and women – A case-control study using primary care data to establish whether the effect of smoking is the same in men and women.

Chapter 4: Is chronic obstructive pulmonary disease an independent risk factor for lung cancer? – A case-control study using primary care data investigating in detail the association between chronic obstructive pulmonary disease and lung cancer, accounting for smoking and timing of diagnoses.

Chapter 5: Validation of records of surgical procedures – A comparison of records of thoracic surgical procedures in two different databases with the aim of determining the most appropriate definition of surgery for future studies.

Chapter 6: Risk factors for early death following surgery for lung cancer - A description of risk models in thoracic surgery followed by a study investigating factors associated with early death after lung cancer surgery resulting in a new predictive model.

Chapter 7: Validation of records of chemotherapy and radiotherapy - A comparison of records of chemotherapy and radiotherapy in two databases with the aim of determining the most appropriate definitions for future studies.

Chapter 8: Treatment decisions and outcomes in small cell lung cancer – A study investigating factors associated with receipt of chemotherapy and completion of a course, and how these factors affected survival.

Chapter 9: Ongoing research – A description of ongoing studies which have resulted from the work in this thesis including validation of the predictive model and a qualitative study exploring attitudes to risk in lung cancer surgery.

Chapter 10: Summary and suggestions for further research – A summary of all of the studies described in this thesis and some proposals for future research.

1.11 Data organisation and statistical methods

All data organisation and statistical analyses were performed using Stata MP Version 11 or 12 (StataCorp, Texas).

In order to acquire skills of data organisation and analysis I completed the following modules which are part of the University of Nottingham Masters in Public Health degree course:

- Research methods in epidemiology and basic statistics – *Self-taught from lecture notes August- September 2011*
- Data Organisation and Management in Epidemiology (DOME) – *October 2011- January 2012*
- Advanced statistical methods – *February- May 2012*

Acknowledgements

The studies described in chapters 3 and 4 used a case-control dataset from a large primary care database. As will be described in chapter 2, section 2.1.3, the case-control dataset was extracted from the database prior to the start of my research. A smoking variable had been defined by Dr Barbara Iyen-Omofoman

during the course of her PhD thesis which I adapted for use in this thesis.(74) I used Read code lists for chronic obstructive pulmonary disease, asthma and pneumonia which had been compiled previously in the Division of Epidemiology and Public Health. I performed all other data organisation and all statistical analyses, with assistance from Dr Laila Tata (Associate Professor and PhD supervisor - University of Nottingham) and Dr Tricia McKeever (Associate Professor – University of Nottingham) for some of the more complex data management in the early stages of my research.

The studies described in chapters 5 to 8 used a linked dataset which I acquired from the Health and Social Care Information Centre by completing a data sharing agreement. Dr Anna Rich used an earlier extract of this linked dataset for her MD thesis and developed a method of calculating a Charlson co-morbidity index and a surrogate start date for people with missing date of diagnosis (sections 2.2.5 and 2.3.5). The Charlson index and start date variables used for this thesis are based on the code lists and methods used by Dr Rich.(81) Dr Laila Tata analyses the NLCA data for the annual reports, (56) and had therefore defined stage and histological subtype variables in this database; I adapted these for the work in this thesis. I performed all other data organisation and statistical analysis.

CHAPTER 2: DESCRIPTION OF DATABASES

In this chapter I describe the sources of routinely collected data which were used for this thesis:

- The Health Improvement Network (THIN)
- The National Lung Cancer Audit (NLCA)
- Hospital Episodes Statistics (HES)
- The Office for National Statistics (ONS)

This is followed by a description of the populations studied and some of the common variables used. I will also discuss some of the generic strengths and weaknesses of studies using these data.

2.1 The Health Improvement Network (THIN)

2.1.1 Background

The system of primary care or general practice in the UK was discussed in section 1.4. All general practices in the UK keep computerised patient records to facilitate consultations, ensure timely and appropriate follow-up, and for evidence of activity for audit and financial purposes. In Practice Systems (InPS) provide *Vision* software which is the interface used by about 2000 general practices in the UK to record these data. Doctors, nurses and administrative staff record data during their day-to-day interactions with patients, and can also upload retrospective data into the patients' records.

The Health Improvement Network is a research database which was set up through collaboration between the Epidemiology and Pharmacology Information Core (EPIC) and InPS in 2002. The data held in primary care patient records using *Vision* software are downloaded by EPIC on a monthly basis, and added to existing files to create a database which is available to researchers. The data are contained in four separate files: patient, medical, therapy and additional health data (AHD) as described in Table 2-1.(84) Each patient has a unique identifier to allow linkage of patient, medical, additional health, and therapy data.

Table 2-1: Description of data files in THIN

Data file	Description
Patient data	Demographic information (including date of birth, sex, practice registration date and date of death)
Medical	Read codes for diagnoses, symptoms, investigations, procedures and hospital admissions.
Additional Health Data (AHD)	Information on lifestyle such as smoking, weight and height, and preventative healthcare such as screening.
Therapy	Drug prescriptions

2.1.2 Ethical approval

The work in this thesis which used the THIN database received ethical approval from the Cegedim Strategic Data Medical Research scientific review committee in 2009. Individual identifiable information was not available to myself or any of the researchers involved in the studies.

2.1.3 Data extract for this thesis

The THIN database used to identify cases and controls for the studies in chapters 3 and 4 was extracted in October 2009 and therefore contains data entered between 2002 and 2009. In October 2009, 446 UK general practices contributed data to THIN and the database contained records for over 8.2 million people.(84) More than 3.2 million of these patients were actively registered and could be prospectively followed.

A lung cancer case-control dataset was created from the October 2009 extract of THIN by Mr Chris Smith and Dr Barbara Iyen-Omofoman (Department of Epidemiology and Public Health, University of Nottingham) in the course of Dr Iyen-Omofoman's PhD project.(74) Cases were patients who had a diagnosis of lung cancer first recorded between 1st January 2000 and 28th July 2009 and at least 12 months of prospectively computerised data prior to this cancer diagnosis date (i.e. they were actively registered with a GP for at least 12 months before diagnosis – this helped to ensure that they were incident cases). Controls were patients with no evidence of current or past lung cancer and were excluded if they had less than 12 months of data before their index date, which was defined as the date of lung cancer diagnosis in their matched case.

The methods for creating this dataset are described in detail in Dr Iyen-Omofoman's PhD thesis, (74) but the stages prior to my use of the data are briefly outlined below:

1. Barbara Iyen-Omofoman (BIO) compiled a list of Read codes (Appendix C) with which people with lung cancer in THIN could be identified, and confirmed these with her PhD supervisors.

2. Chris Smith (CS) extracted all data on all patients with a read code for lung cancer from the entire THIN population.

3. BIO identified the incident cases of lung cancer and performed several data cleaning tasks to ensure that these were true incident cases. This included the exclusion of any cases where the date of death or final contribution of data was >31 days before the lung cancer diagnosis date.

4. BIO and CS worked together to identify and extract matched controls. Up to four controls were matched to each case on sex, year of birth and the general practice with which they were registered. They assigned an index date to each control which was equal to the date of diagnosis in the matched case.

This case-control dataset contained 12,121 incident cases with a first record of lung cancer between January 2000 and July 2009. A total of 48,216 controls were identified: 11,960 cases were matched with 4 controls, 84 with 3 controls, 47 with 2 controls and 30 with 1 control. The full dataset contained data on 60,337 people.

2.1.4 THIN variables used for this thesis

The following variables were used for the studies in chapters 3 and 4. Further details of specific variables for each of these studies are given with the individual methods sections.

Patient data

Townsend score is a measure of deprivation and disadvantage (commonly termed 'socioeconomic status') derived from the 1991 census data and based on levels of unemployment, non-car ownership, non-home ownership and home overcrowding.(85) Residential areas are divided in to Lower Super Output Areas

(LSOAs) by postcode with approximately 1500 homes in each area; each area is linked with a Townsend score for deprivation. For the purpose of these analyses Townsend scores were divided into quintiles; quintile 5 is the most deprived and quintile 1 represents the least deprived (or most advantaged) quintile of society.

Age at diagnosis was calculated using the first recorded diagnosis of lung cancer and the patient's date of birth; for controls the age at index date was calculated.

Medical data

Read codes are a standard classification system used by general practitioners in the UK to record patients' medical information. Diagnoses made both in general practice and in secondary care are recorded and lists of read codes can be used to identify people with a particular diagnosis, symptom or procedure in the THIN database. Lists of the Read codes used in this thesis can be found in Appendix C.

Additional health data

Smoking status: Read code lists for smoking status, including quantity smoked, had previously been developed, and validated within the Division of Epidemiology and Public Health at the University of Nottingham. These code lists (Appendix C) were used to identify all records of smoking status within each patient's record. Cases and controls were categorised as current, ex, or never-smokers according to the codes recorded prior to their lung cancer diagnosis or index date.

Smoking quantity: The record of number of cigarettes smoked per day was identified using the AHD codes in Appendix C. Quantity was defined as light (1-9 cigarettes per day), moderate (10-19 cigarettes per day) or heavy (20 or more cigarettes per day). Current or ex-smokers who had no record of their daily cigarette consumption were recorded as smokers with unknown quantity and those with no recorded smoking information were included as 'missing smoking status'. The highest quantity ever recorded was used, but the most recent

quantity (excluding records in the six months prior to lung cancer diagnosis or index date so as not to capture potential reductions in quantity smoked due to suspicion of lung cancer) was used in sensitivity analyses.

2.1.5 Strengths & weaknesses

There are some important strengths and weakness which are common to both of the studies based on THIN in this thesis. These are briefly introduced here and discussed further in the relevant chapters.

Sample size

There were 3.2 million actively registered patients in the October 2009 extract of THIN. Studies using THIN to investigate diseases such as lung cancer, which have a high incidence in the UK, therefore have the advantage of being able to include several thousand cases.

Unselected population

All sections of the population are represented in THIN due to the number and spread of practices which contribute data. A validation study has been performed to assess lung cancer cases in THIN, comparing the incidence and survival with data from national cancer registries and the National Lung Cancer Audit. Incidence and survival both overall and by sex, age at diagnosis and at death, geographical area, and level of socioeconomic deprivation in THIN are comparable to data from these other sources, although THIN does appear to capture a higher proportion of lung cancer incidence in more recent years (after 2004). (65)

Prospective data entry

Exposure data are recorded prospectively which minimises recall bias.

Data completeness

A patient's GP should co-ordinate all of their medical care and should be informed of, and record, everything that affects that patient. A primary care database should, therefore, contain codes for all diagnoses for every patient. There may, however be times when this does not occur if communication from secondary care is not clear, or if administrative staff do not consider information to be important. In addition, THIN relies on patients consulting their GP which does not always happen and therefore some data are incomplete.

Following the introduction of the new GP contract and Quality and Outcome Framework (QOF) there are financial incentives for GPs to ensure that their patients' electronic records are complete and accurate. This has meant that, for example, the vast majority of patients now have their smoking status recorded, which is clearly important for studies of lung cancer. There are, however, still missing data in areas such as smoking quantity.

2.2 The National Lung Cancer Audit (NLCA)

2.2.1 Background

The NLCA collects data on people with primary lung cancer from trusts in England and Wales (the UK system of NHS hospital trusts was described in section 1.4). It was set up by the Royal College of Physicians in 2002 with the aim of improving outcomes for people with lung cancer and is currently commissioned by the Healthcare Quality Improvement Partnership (HQIP). Data are collected and held by the Health and Social Care Information Centre (HSCIC) in Leeds.

Data collection began in 2004 and all NHS trusts in England and Wales, plus a few private healthcare trusts, now contribute. Input of data is non-mandatory and therefore case ascertainment, particularly in the early years, was not complete. In 2004 and 2005 case ascertainment was estimated at 19% and 42% respectively, in 2006, 60% of the expected cases were entered and from 2009 onwards at least 97%.(51) Staff members at each trust have secure access to all of their own data through the NLCA website and an annual report is published providing trusts with overall figures, with which they can compare and evaluate their performance.(51) This encourages clinicians to ensure that their data are both accurate and complete.

2.2.2 Data entry

Data are usually entered as a result of the first discussion at the lung cancer multi-disciplinary team (MDT) meeting (as described in section 1.4) and consequently patients who are never seen in secondary care are not captured by this database. There is no consistent method by which trusts enter or upload data to the NLCA database and it may be done by respiratory physicians, lung cancer specialist nurses, lung cancer co-ordinators, specialist audit data managers, or administrators depending on the facilities available at each trust.

Patient data are entered into an online form which is shown in Appendix D. The intention of the audit team was that Trusts uploaded data during the course of the patient's journey through the lung cancer MDT process, and that a final check was performed shortly after treatment was completed, thereby ensuring all data were accurate and that any changes to treatment plans were updated. Due to time constraints in the NHS many trusts actually enter data in chunks, often immediately before the closing date which, for patients first seen in each calendar year, is the following June.

Data entry is never closed so information on new patients can be entered, or records of existing patients changed, at any point in time. Information about any patient can be entered by more than one NHS Trusts, with the latest entry overwriting preceding entries. Whilst changes made after June each year would not contribute to the NLCA annual report, they replace previous versions of the database and therefore are included in the database used for the studies in this thesis.

2.2.3 Ethical approval

This is discussed in section 2.3.3.

2.2.4 Data extracts for this thesis

The HSCIC does not release NLCA data to researchers until the annual report which covers the most recent year of data collection has been published.(51) Two extracts of the NLCA database were used for this thesis: The first, which was used for the study of surgery in NSCLC (chapters 5 and 6), was extracted in August 2011. This raw data file contained records for a total of 156,325 people with primary lung cancer or mesothelioma first seen in England between 1st January 2004 and 31st December 2010.

The study of chemotherapy in SCLC (chapters 7 and 8) was undertaken approximately one year after the work on surgery by which time a more recent

extract of data was available. This was extracted in July 2013 and the raw data file contained records for a total of 178,428 people with primary lung cancer or mesothelioma first seen in England between 1st January 2004 and 31st December 2011.

2.2.5 NLCA Variables used for this thesis

The NLCA database contains demographic data, dates and modalities of referral, diagnosis and treatments, and data on stage, performance status and lung function. Limited data on co-morbidity are requested but these fields are not mandatory and at present are not reliable. Further details of a few key variables are given below with further details in Chapters 5 - 8.

Patient data

Townsend score: Lower Super Output Area (LSOA) data are included in the NLCA (as in THIN – see section 2.1.4) and can be used to calculate a Townsend score. For the purpose of these analyses Townsend scores were divided into quintiles where 5 was the most deprived and 1 the least deprived (or most advantaged) quintile of society.

Lung function is recorded in the NLCA as absolute volume and as percentage of predicted forced expiratory volume in 1-second (FEV1). During data cleaning any measurements of >150% or <10% predicted were deleted as these were felt likely to reflect errors in data entry.

Performance status is a subjective measure of assess how a disease affects the daily living abilities of the patient. It is recorded in the NLCA on a scale of 0 to 4 as defined by the Eastern Cooperative Oncology Group (Table 2-2), at a single point in time; this is usually the day of MDT discussion and therefore usually reflects the performance status recorded by the respiratory physician at the initial consultation. (86)

Table 2-2: Description of Eastern Co-operative Group performance status

Performance status	Description
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair
5	Dead

Tumour data

Stage: The standard NLCA definition of tumour stage considers that post-treatment (usually post-surgical) records of stage are more accurate than pre-treatment records and therefore prioritises these records, only using pre-treatment records when post-treatment stage is missing. For the studies described in this thesis, however, pre-operative stage was prioritised over post-operative stage because the focus here was on the pre-operative plan (i.e. surgery with curative intent) and pre-operative estimation of operative risks. If pre-treatment stage was missing the post-treatment stage was used to reduce the amount of missing data.

NSCLC stage is recorded in the NLCA according to the Union for International Cancer Control (UICC) Tumour Node Metastases (TNM) staging system.(16) In 2009, after work by the International Association for the Study of Lung Cancer (IASLC),(19) the UICC staging system changed slightly and Revision 7 was introduced.(16) During 2010 participating trusts were given the option to enter stage using revision 6 or revision 7, with a field indicating which system was

used. It is not possible to convert stages recorded using version 6 to version 7 using the data provided in the NLCA, and therefore for this work the recorded stage was used without taking into account which version was used.

SCLC stage is now recorded using the TNM system as recommended by the UICC. In the extracts of the NLCA used for this thesis, however, it was almost always recorded using the Veterans' Administration Lung Study Group system (limited or extensive). For any cases where the more recent TNM staging system was used this was converted to limited (T1-4, N0-3, M0) or extensive (M1a/b) as appropriate.(16)

Histological subtype: The options for entering histological subtypes (using Systematised Nomenclature of Medicine (SNOMed) coding), and the NLCA classification of NSCLC, SCLC, carcinoid and mesothelioma are shown in Table 2-3. The NLCA assumes that patients who do not have a record of pre- or post-treatment histology have NSCLC unless an ICD-10 code for mesothelioma is recorded elsewhere (it would usually be clear clinically and radiologically that the diagnosis was mesothelioma, even without histology). This is because the vast majority of cases of lung cancer are NSCLC. It would be unusual not to obtain tissue in carcinoid given the good prognosis with treatment, but some cases of SCLC which did not have a histological diagnosis may be misclassified as NSCLC using this definition.

Table 2-3: Systematised Nomenclature for Medicine (SNoMed) codes and classification of histology in the NLCA

SNoMed code	Description	NLCA histology category
M8010/2	Carcinoma in situ	**NSCLC**
M8041/3	Small cell carcinoma	**SCLC**
M8046/3	Non-small cell carcinoma (includes adenosquamous carcinoma)	**NSCLC**
M8070/3	Squamous cell carcinoma NOS	**NSCLC**
M8140/3	Adenocarcinoma NOS (without alveolar cell features)	**NSCLC**
M8250/3	Bronchio-alveolar cell carcinoma	**NSCLC**
M8012/3	Large cell carcinoma NOS	**NSCLC**
M8020/3	Large cell – undifferentiated	**NSCLC**
M8013/3	Large cell neuroendocrine	**NSCLC**
M8240/3	Carcinoid tumour NOS (includes atypical carcinoid)	**Carcinoid**
M8980/3	Carcinosarcoma NOS	**NSCLC**
M9050/3	Malignant mesothelioma NOS	**Mesothelioma**
M9052/3	Mesothelioma (epitheliod)	**Mesothelioma**
M9051/3	Mesothelioma (sarcomatoid)	**Mesothelioma**
M8940/3	Mixed tumour (malignant)	**NSCLC**
M9999/3	Other	**NSCLC**

For patients who had surgery there is often a record of post- as well as pre-treatment histology. If pre- and post- treatment histological types differ the standard NLCA definition considers that post-treatment histological type is the most accurate because this is usually based on a larger sample of tissue. In the studies described in this thesis, however, the pre-operative stage was prioritised over post-operative, for the same reasons described above for stage. If pre-treatment histological subtype was missing the post-treatment histology was used to reduce the amount of missing data.

Dates of diagnosis, investigations and treatment

Date of diagnosis (start date): For survival analyses the date of diagnosis was important, however in some cases this information was missing. Dr Anna Rich used the NCLA database for her thesis entitled 'Validation of the National Lung Cancer Audit database and analysis of the information it contains',(81) and in the course of this research developed a method of determining a surrogate date of diagnosis or start-date for each patient as follows:

The date of diagnosis was used unless this field was missing. Where date of diagnosis was missing it was estimated using one the following dates (in order of priority as listed) and the median difference between that date and the date of diagnosis in the rest of the initial cohort (shown in brackets): Date of first NHS Trust appointment (+5 days), date of referral to the lung cancer team (+18 days) or date of multi-disciplinary team meeting (-9 days).

Although the data extracts were different to those used by Dr Rich, the median differences between each of the dates listed above were the same. It was not possible to calculate a start date for a small proportion of patients; in most of the analyses these patients were excluded or a sensitivity analysis excluding these records was conducted.

Investigation and treatments are recorded in various sections of the database. For most fields this is just the date of the investigation or the date of first treatment, but for some treatments more information is given, e.g. procedure type for those who had surgery. This is discussed in much greater detail in chapters 5-8.

Hospital trusts (as described in section 1.4) are identified using standard coding. The NLCA identifies the trust where the patient was first seen for lung cancer, where the diagnosis was made, and also where any treatment was given. The cancer **network** where a patient was first seen or treated is derived

from the trust code. Trusts are grouped into networks as listed in the NLCA Annual Report 2012. (51)

2.2.6 Strengths and weaknesses

There are some important strengths and weakness which are common to all of the studies in this thesis which used the NLCA. These are briefly introduced here and discussed further in the relevant chapters.

Sample size

The NLCA is the largest, unselected, non-Registry lung cancer database in Europe.

Population

The NLCA collects data from all of the NHS Trusts in England that provide care to people with lung cancer, and is therefore much more comprehensive than other similar databases such as the registry linked EUROCARE-4 or the North American Surveillance, Epidemiology and End Results (SEER) programme. EUROCARE is a European collaboration involving 47 cancer registries from 21 countries (including 4 regional registries for England), and SEER represents 26% of the overall North American population; both are large datasets which contain information on incidence, treatment and survival for all major cancers, including but not specific to lung cancer. (87, 88)

Work from the University of Nottingham published in 2011 found that cases in the NLCA database first seen between 2004 and 2008 were representative of patients with lung cancer in England when compared with national cancer registry data.(82)

Data fields

One of the main advantages of the NLCA over cancer registry data is that, although registries capture marginally more cases, the NLCA includes far more detail in data fields such as performance status, lung function and treatment.

The data on co-morbidity collected by the NLCA are, at present, insufficient for studies such as those in this thesis. Fortunately the NLCA database has been linked with HES from which a measure of co-morbidity has been derived (see below).

Errors in data entry

All data in the NLCA database are entered manually, usually by administrative staff following an MDT discussion, and it is likely that there will be some errors in data entry. Provided these errors are random they are unlikely to affect the results of studies such as those reported in this thesis, however large proportions of missing data may lead to concerns over the validity of results.

Changes in patient condition

As described in section 2.2.1, for each patient data are usually entered into the NLCA at a single point in time. Whilst it is possible to change the record an unlimited number of times, it is only possible to record a single performance status, lung function measurement and stage. The NLCA guidelines state that this should reflect the patient's condition at the time they patient are first seen, but means that changes in patient fitness are not captured.

2.3 Hospital Episodes Statistics (HES)

2.3.1 Background

The HES database contains data generated through inpatient admissions to NHS hospitals in England. Data collection started in 1989, however data completeness was not considered to be of a standard suitable for use in research until about 1997. Data are grouped by financial year and every admission is divided into episodes (one episode is a single period of care under one consultant). There is one row of data for each individual episode and therefore one patient may have hundreds of rows of data in just one year.

Data are collected as a part of the process by which NHS Trusts charge for their services but are also used for a range of health services research in the NHS, and by the UK government and other organisations.(89). The database consists of demographic information, dates of admission, discharge, and procedures, diagnosis codes (coded using the International Classification of Diseases revision 10, ICD-10) and procedure codes (coded using the Office of Population Censuses and Surveys Classification of Interventions and Procedures, OPCS-4).

There are separate databases for outpatient and A&E attendances which record diagnoses and procedures made in these settings. These data were not available at the time this work took place therefore only inpatient HES data were used.

2.3.2 Data extracts for this thesis

Data from HES are collected in financial years. At the close of each financial year certain checks and data cleaning procedures are carried out. The data are not, therefore, available to researchers until several months after the end of the financial year.

For the studies used in this thesis, the first HES dataset was extracted in August 2011 and data were available from 1st April 1997 to 31st March 2010 (inclusive). The second HES dataset was extracted in July 2013 and contained HES data for

financial years 1997/98 to 2011/12. Data were obtained for all patients who had a record in the NLCA database at the time of data extraction, linked as described below.

2.3.3 Linkage with the NLCA database

Whilst the NLCA does collects some co-morbidity data, the fields are not mandatory and at present not felt to be reliable. Prior to the start of this thesis, researchers at the University of Nottingham arranged for the HSCIC to link the NLCA database with HES data so that co-morbidity data from HES were available to supplement the NLCA database for lung cancer research.(81, 83)

Linkage was performed by the HSCIC using NHS numbers, with additional checks using date of birth, sex, postal code. The data are pseudonomised meaning that researchers have access to a database identifier which links records in the NLCA and HES, but not to individual NHS numbers with which they could identify people.

2.3.4 Ethical approval

The work in this thesis using the NLCA-HES linked data was approved by the NLCA regulatory body, HQIP. I completed data-sharing agreements with the Health and Social Care Information Centre (HSCIC) allowing me to use the NLCA - HES linked data for the purpose of this thesis and the publications arising. For the latest extract, local ethics committee approval was also required due to changes in procedure at the HSCIC. I therefore obtained approval from the University of Nottingham Medical School Ethics Committee to use NLCA-HES linked data for the work described in chapters 7 and 8.

2.3.5 HES variables used for this thesis

A single row of data in this extract of the HES database included a unique patient identifier (used to link to the NLCA database), sex, ethnicity, date of admission to and discharge from hospital, episode start and end dates, spell start and end dates (a spell is made up of several episodes), up to 20 diagnosis codes per episode, and

up to 24 procedure codes. The patient identifier and admission date may be the same for several rows of data reflecting multiple episodes in a single spell or admission. Additional data are available in HES but were not used for these studies.

Demographics

Ethnicity as recorded in HES was used for the studies in chapters 5-8, and was categorised as White, Black, Asian, other or missing ethnicity.

Surgical procedures and chemotherapy

Since procedures, including thoracic surgery and administration of chemotherapy, are recorded in both HES and the NLCA, it was possible examine differences between treatment records in order to determine which was likely to be the most accurate. This work is the subject of chapters 5 and 7.

Co-morbidity

When entering data, coding staff are expected to code the primary diagnosis or diagnoses for each episode as well as all significant co-morbidities. These ICD-10 codes (listed in Appendix E) were used to determine whether or not each patient had certain diagnoses so that a Charlson co-morbidity index could be calculated.

The Charlson index, developed in 1987, gives each of 19 medical conditions a weighted score, based on their relative mortality risk, and combines them to give the Charlson co-morbidity index (CCI). (90) The 19 medical conditions are listed in table 2-4. The CCI has been widely used in research as a marker of co-morbidity as it provides a means of taking into account several co-morbidities without assigning a different estimate of risk to each individual disease. Modern treatments for HIV and AIDS are such that this is longer considered a rapidly fatal illness but for the purpose of this work the original index was used as the number of patients with lung cancer who also have HIV is negligible. The only modification which was necessary was the exclusion of lung cancer from 'any tumour' because all patients had lung cancer.

Table 2-4: Charlson co-morbidity index, (90)

Assigned weights for diseases	Conditions
1	Myocardial infarct Congestive heart failure Peripheral vascular disease Cerebrovascular disease Dementia Chronic pulmonary disease Connective tissue disease Ulcer disease Mild liver disease Diabetes
2	Hemiplegia Moderate or severe renal disease Diabetes with end organ damage Any tumour Leukaemia Lymphoma
3	Moderate or severe liver disease.
6	Metastatic solid tumour Acquired Immune Deficiency Syndrome

Assigned weights for each condition that a patient has. The total equals the score. Example: Chronic pulmonary (1) and lymphoma (2) = total score (3)

2.3.6 Strengths and weaknesses

There are some important strengths and weakness which are common to all of the studies in this thesis which used HES data. These are briefly introduced here and discussed further in the relevant chapters.

Data completeness

The HES data are predominantly used by secondary care trusts to charge for their services. There is therefore an incentive for managers to ensure that the data are complete, however it is important to remember that the reason for data collection was not the same as that for which it is used in research.

Data entry

All data in the HES database are entered manually, by administrative staff, usually from medical notes at the end of a hospital episode and as such is subject to errors in data entry. The data are, however, audited annually for accuracy.(91)

Inpatient data only

If a patient had never had an admission to hospital prior to the diagnosis of lung cancer their CCI would be 0 even though they may have had one or more of the diagnoses made by their GP. The original Charlson index was also based on secondary care records of disease however this is a potential weakness of using inpatient HES data to derive a measure of co-morbidity.

2.4 Office for National Statistics

The Office for National Statistics (ONS) is the UK's largest independent producer of official statistics and the recognised national statistical institute of the UK. Part of their remit is to collect information on date and cause of death from civil registration records.

2.4.1 Ethical approval

In order to use ONS death data for the work in this thesis I completed a data sharing agreement and became an ONS approved researcher.

2.4.2 Data extraction and linkage

The HSCIC obtained ONS death data for all patients in the NLCA database and provided it merged with the NLCA dataset for the August 2011 extract, and as a separate file with HES identifiers for the July 2013 extract. The linkage was based on NHS number as described above for the NLCA-HES linkage. This was done on 8th August 2011 for the first NLCA extract and on 31st March 2013 for the second extract. Any patient with missing death date could therefore be assumed to be alive, and for survival analyses censored, on this date.

2.4.3 ONS variables used for this thesis

The only ONS variable used for this thesis was date of death.

2.4.4 Strengths and weaknesses

<u>Follow-up time</u>

The ONS continually collect death data. Cross reference and linkage to the NLCA database using NHS number allows almost continuous follow-up and analysis of survival without individual trusts having to collect these data for their own patients.

<u>Accuracy and completeness</u>

Death registration in the UK is mandatory and therefore these data are highly complete and accurate. For a minority of patients there is a delay between death

and death certification, usually due to the need for a coroner's inquiry. This is not a common occurrence however it may lead to a few patients being classified as alive on the censor date because their death has not yet been registered.

2.5 Chapter summary

In this chapter I have described four sources of data which are routinely collected for non-research purposes but which can be used for observational studies of the aetiology, treatment and outcomes of lung cancer in the UK. The following two chapters describe two studies of risk factors for lung cancer using THIN. Chapters 5-8 describe the use of NLCA-HES-ONS linked data to study treatments and outcomes for people with lung cancer.

CHAPTER 3: SMOKING QUANTITY AND LUNG CANCER IN MEN AND WOMEN

This chapter describes the use of a matched lung cancer case-control dataset, derived from THIN, to establish whether the association between smoking quantity and lung cancer is the same in men and women.

3.1 Introduction

3.1.1 Background

Lung cancer kills more women than any other cancer and deaths have exceeded those from breast cancer for the past 20 years. (92) Whilst lung cancer does occur in non-smokers, smoking is by far the most important risk factor, with over 80% of all lung cancer attributable to smoking cigarettes. (93, 94) As described in chapter 1, smoking prevalence in women increased following the end of the Second World War to a peak prevalence of about 40% in Northern Europe in the 1980s. Worldwide, at least 250 million women smoke and, although in high income countries the prevalence is generally decreasing, in some European countries it now exceeds that in men. (7)

3.1.2 Rationale for this study

Most studies quantifying smoking-related cancer risks are in men and these have been extrapolated to female populations, (3, 95) yet evidence from a recent systematic review showed that women who smoke have a 25% greater risk of coronary heart disease than male smokers. (96) This relationship has also been examined in lung cancer but with conflicting results, (97-102) which may in part be due to variation in smoking patterns and prevalence between countries. No previous study has assessed the effect in a UK population.

3.1.3 Aim of this chapter

The aim of this analysis of THIN was to investigate whether the risk of lung cancer differs between men and women with the same recorded quantity of cigarettes smoked, and to test the hypothesis that if women are at higher risk of the effects of cigarette smoke this may be because they have smaller lung volumes than men, and hence a higher dose per lung volume for the same number of cigarettes smoked.

3.2 Methods

3.2.1 Dataset & Study Population

The matched lung cancer case-control dataset, extracted from THIN and described in section 2.1.3, was used for this study.

3.2.2 Definition of Exposures

Sex was obtained from the patient data file.

Smoking status and smoking quantity were defined using the additional health data file as described in section 2.1.4. Smoking was categorised as: Never, light (1-9 cigarettes per day), moderate (10-19 cigarettes per day), heavy (20 or more cigarettes per day), smoker with missing quantity, or missing smoking status.

Quantity smoked was defined as the highest quantity ever recorded prior to the lung cancer diagnosis or index date. The most recent quantity (but excluding records in the six months prior to lung cancer diagnosis or index date so as not to capture potential reductions in quantity smoked due to suspicion of lung cancer) was used in a sensitivity analysis.

3.2.3 Covariate definitions

Townsend quintile, as described in section 2.1.4, was used to define socio-economic status. Quintile 1 represents the most affluent and quintile 5 the most deprived people.

Height: The tallest height ever recorded for each individual was used as a surrogate for lung volume (because height is the predominant determinant of total lung capacity, (103)). Height was categorised according to quintile, 1 being the shortest and 5 the tallest category, for the population overall (i.e. men and women combined). Units are often not recorded in THIN and therefore during data cleaning values of height in the range 1.2-2.0 were presumed to be in

metres and those in the range 120-200 were presumed to be in centimetres. Any values outside this range were assumed to be errors and not used.

3.2.4 Statistical methods

For the population overall, and for males and females separately, proportions of cases and controls across 10-year age bands, Townsend quintiles, smoking quantity, and height quintiles were compared. A conditional logistic regression model was used to calculate odds ratios for lung cancer by smoking quantity in the dataset overall. A multiplicative test for interaction was then used to assess whether the effect of smoking quantity on lung cancer differed between men and women; a p value <0.01 was considered statistically significant.

To test the hypothesis that lung volume (represented by height) explains any difference in effect of smoking quantity on lung cancer risk the conditional logistic regression model was used to estimate odds ratios for lung cancer by height quintile and the interaction was re-assessed in the model which adjusted for height.

3.3 Results

The dataset contained information on a total of 60,337 people: 12,121 incident cases of lung cancer between January 2000 and July 2009, and 48,216 matched controls. Fifty-nine patients (of whom 49% were female) had an age at diagnosis of less than 40 years and were excluded, as were their 236 matched controls, leaving a total of 60,042 patients for analysis.

Forty-one per cent of cases were female. Overall, patients had a median of 9.6 years of data available. A larger proportion of women had never smoked (41% compared to 26% of men) and the proportion of heavy smokers was higher in men (19% compared to 15% of women) (Table 3-1). Fifty-eight per cent of lung cancer cases were in moderate or heavy smokers; this was similar for men and women, but a higher proportion of females who developed lung cancer were recorded as never smokers (13% compared to 8% of males). The height distribution was as expected with the majority of females in quintiles 1-3 and the majority of males in quintiles 3-5. There were no differences in age at diagnosis between men and women and the distribution of socioeconomic deprivation was also very similar.

The odds of lung cancer were much higher in people who smoked compared with those who had never smoked, the odds increasing with quantity of cigarettes smoked (for the heaviest smokers the odds ratio (OR) overall was 15.13) (Table 3-2). The multiplicative test for interaction showed strong evidence of a difference in the effect of quantity smoked on lung cancer between men and women (likelihood ratio test $p<0.0001$).

When compared to men within strata of smoking quantity, the odds ratios for lung cancer in women were 1.02 (95%CI 0.91-1.15) for never smokers, 1.06 (95% CI 0.92-1.23) for light smokers, 1.32 (95%CI 1.20-1.46) for moderate smokers, 1.42 (95% CI 1.31-1.54) for heavy smokers, 0.92 (95%CI 0.84-1.02)

for smokers with unknown quantity and 1.25 (95%CI 1.04-1.50) for those with missing smoking status.

3.3.1 Sensitivity analysis

To investigate the relationship further the same analyses were performed using the latest smoking status recorded up to 6 months prior to the index date. The results were very similar: in current moderate, current heavy, and ex-heavy smokers women had significantly higher odds of lung cancer. For ex-moderate smokers the odds ratios were higher for women but this was not statistically significant (Table 2).

3.3.2 Height

The mean height for the study population was 1.68 metres (m) (standard deviation 0.1m); this was the same for cases and controls. In the overall population, and also when stratified by sex, the odds ratios for lung cancer were not significantly different between the first (shortest) and any other height quintile (Table 2). There were no differences in smoking quantity according to height quintile for men or women. The interaction for the effect of smoking quantity in men and women remained after adjusting for height ($p<0.0001$).

Table 3-1: Quantity smoked, height, Townsend score and age at lung cancer diagnosis for cases and controls overall and by sex

		OVERALL N=60,042					MALES N=35,481					FEMALES N=24,561				
		Cases (n=12,062)		Controls (n=47,980)		TOTAL	Cases (n=7,143)		Controls (n=28,338)		TOTAL	Cases (n=4,919)		Controls (n=19,642)		TOTAL
Length of data	**Median (years)**	9.54		9.56		**9.56**	9.60		9.64		**9.63**	9.48		9.47		**9.47**
		n	**%‡**	**n**	**%‡**	**%‡**	**n**	**%***	**n**	**%***	**%***	**n**	**%†**	**n**	**%†**	**%†**
Smoking	**Never**	1,213	*10.1*	17,976	*37.5*	***32.0***	577	*8.1*	8,655	*30.5*	***26.0***	636	*12.9*	9,321	*47.5*	***40.5***
quantity	**Trivial/light**	1,030	*8.5*	3,113	*6.5*	***6.9***	640	*9.0*	1,978	*7.0*	***7.4***	390	*7.9*	1,135	*5.8*	***6.2***
	Moderate	2,465	*20.4*	4,859	*10.1*	***12.2***	1,303	*18.2*	2,904	*10.2*	***11.9***	1,162	*23.6*	1,955	*10.0*	***12.7***
	Heavy / very heavy	4,547	*37.7*	6,051	*12.6*	***17.7***	2,712	*38.0*	4,102	*14.5*	***19.2***	1,835	*37.3*	1,949	*9.9*	***15.4***
	Smoker but unknown quantity	2,285	*18.9*	10,732	*22.4*	***21.7***	1,607	*22.5*	7,364	*26.0*	***25.3***	678	*13.8*	3,368	*17.1*	***16.5***
	Missing smoking status	522	*4.3*	5,249	*10.9*	***9.6***	304	*4.3*	3,335	*11.8*	***10.3***	218	*4.4*	1,914	*9.7*	***8.7***
Height	**≤1.60**	2,360	*19.6*	9,387	*19.6*	***19.6***	179	*2.5*	760	*2.7*	***2.6***	2,181	*44.3*	8,627	*43.9*	***44.0***
(quintiles)	**>1.62,≤1.66**	1,826	*15.1*	7,108	*14.8*	***14.9***	563	*7.9*	2,319	*8.2*	***8.1***	1,263	*25.7*	4,789	*24.4*	***24.6***
Metres	**>1.66,≤1.72**	2,059	*17.1*	8,150	*17.0*	***17.0***	1,454	*20.4*	5,571	*19.7*	***19.8***	605	*12.3*	2,579	*13.1*	***13.0***
	>1.72,≤1.78	2,485	*20.6*	9,623	*20.1*	***20.2***	2,288	*32.0*	8,827	*31.1*	***31.3***	197	*4.0*	796	*4.1*	***4.0***
	>1.78	1,614	*13.4*	6,584	*13.7*	***13.7***	1,590	*22.3*	6,473	*22.8*	***22.7***	24	*0.5*	111	*0.6*	***0.5***
	Missing	1,718	*14.2*	7,128	*14.9*	***14.7***	1,069	*15.0*	4,388	*15.5*	***15.4***	649	*13.2*	2,740	*13.9*	***13.8***
Townsend	**(least deprived) 1**	2,064	*17.1*	10,779	*22.5*	***21.4***	1,289	*18.0*	6,615	*23.3*	***22.3***	775	*15.8*	4,164	*21.2*	***20.1***
score	**2**	2,233	*18.5*	10,262	*21.4*	***20.8***	1,366	*19.1*	6,060	*21.4*	***20.9***	867	*17.6*	4,202	*21.4*	***20.6***
	3	2,420	*20.1*	9,482	*19.8*	***19.8***	1,452	*20.3*	5,579	*19.7*	***19.8***	968	*19.7*	3,903	*19.9*	***19.8***
	4	2,638	*21.9*	8,755	*18.2*	***19.0***	1,530	*21.4*	5,061	*17.9*	***18.6***	1,108	*22.5*	3,694	*18.8*	***19.6***
	(most deprived) 5	2,232	*18.5*	6,748	*14.1*	***15.0***	1,237	*17.3*	3,915	*13.8*	***14.5***	995	*20.2*	2,833	*14.4*	***15.6***
	Missing	475	*3.9*	1,954	*4.1*	***4.0***	269	*3.8*	1,108	*3.9*	***3.9***	206	*4.2*	846	*4.3*	***4.3***
Age at diagnosis	**40-49**	315	*2.6*	1,260	*2.6*	***2.6***	168	*2.4*	672	*2.4*	***2.4***	*147*	*3.0*	588	*3.0*	***3.0***
index date	**50-59**	1,367	*11.3*	5,467	*11.4*	***11.4***	793	*11.1*	3,172	*11.2*	***11.2***	*574*	*11.7*	2,295	*11.7*	***11.7***
(matched)	**60-69**	3,236	*26.8*	12,934	*27.0*	***26.9***	1,951	*27.3*	7,797	*27.5*	***27.5***	*1,285*	*26.1*	5,137	*26.2*	***26.1***
	70-79	4,520	*37.5*	18,011	*37.5*	***37.5***	2,738	*38.3*	10,896	*38.5*	***38.4***	*1,782*	*36.2*	7,115	*36.2*	***36.2***
	≥80	2,624	*21.8*	10,308	*21.5*	***21.5***	1,493	*20.9*	5,801	*20.5*	***20.6***	*1,131*	*23.0*	4,507	*22.9*	***23.0***

‡ Proportion of population overall; * Proportion of males overall; † Proportion of women overall

Table 3-2: Odds ratios for lung cancer by quantity smoked (highest and latest recorded) for men and women

		OVERALL (N=60,042)				MALE (n=35,481)				FEMALE (n=24,561)			
		Odds ratio (OR)		Adjusted OR*		Odds ratio (OR)		Adjusted OR*		Odds ratio (OR)		Adjusted OR*	
			95% CI		*95% CI*		*95% CI*		*95% CI*		*95% CI*		*95% CI*
Smoking	**Never**	1.00		1.00		1.00		1.00		1.00		1.00	
quantity	**Trivial / light**	5.79	*5.26-6.38*	5.83	*5.30-6.42*	5.61	*4.94-6.38*	5.67	*4.99-6.44*	5.75	*4.95-6.68*	5.78	*4.97-6.71*
Highest	**Moderate**	9.37	*8.63-10.17*	9.43	*8.69-10.24*	8.24	*7.36-9.24*	8.34	*7.44-9.34*	10.78	*9.57-12.15*	10.82	*9.61-12.19*
reported	**Heavy / very**	15.13	*14.00-16.35*	15.30	*14.15-16.54*	12.81	*11.52-14.24*	13.04	*11.73-14.50*	19.10	*16.98-21.49*	19.19	*17.06-21.29*
	Smoker unknown quantity	3.61	*3.34-3.91*	3.65	*3.37-3.95*	3.60	*3.23-3.99*	3.64	*3.28-4.05*	3.32	*2.93-3.75*	3.34	*2.95-3.78*
	Missing smoking status	1.29	*1.15-1.45*	0.99	*0.88-1.12*	1.21	*1.04-1.41*	0.90	*0.76-1.05*	1.34	*1.13-1.60*	1.12	*0.92-1.35*
Smoking	**Never**	1.00		1.00		1.00		1.00		1.00		1.00	
quantity	**Ex light**	7.01	*6.29-7.82*	7.19	*6.45-8.02*	6.58	*5.68-7.61*	6.79	*5.86-7.85*	7.43	*6.30-8.77*	7.57	*6.42-8.93*
Latest	**Ex moderate**	8.76	*7.92-9.70*	9.00	*8.12-9.96*	7.77	*6.78-8.91*	8.06	*7.02-9.24*	10.07	*8.62-1.76*	10.22	*8.75-11.93*
reported †	**Ex heavy**	10.77	*9.73-11.92*	11.03	*9.96-12.21*	9.27	*8.14-10.56*	9.58	*8.41-10.92*	13.84	*11.68-16.41*	13.99	*11.80-16.58*
	Current light	9.32	*8.48-10.25*	9.38	*8.53-10.31*	8.35	*7.35-9.48*	8.42	*7.41-9.56*	10.60	*9.18-12.24*	10.64	*9.21-12.30*
	Current	11.78	*10.79-12.87*	11.77	*10.77-12.86*	10.41	*9.21-11.77*	10.44	*9.23-11.81*	13.49	*11.87-15.34*	13.47	*11.85-15.32*
	Current heavy	15.02	*13.69-16.48*	14.97	*13.64-16.43*	12.73	*11.24-14.41*	12.72	*11.23-14.41*	18.74	*16.26-21.60*	18.67	*16.20-21.52*
	Ex / current unknown quantity	3.91	*3.62-4.23*	3.94	*3.65-4.26*	3.90	*3.52-4.33*	3.95	*3.56-4.38*	3.59	*3.18-4.01*	3.60	*3.19-4.06*
	Missing smoking status	1.30	*1.16-1.46*	1.01	*0.90-1.15*	1.22	*1.05-1.43*	0.92	*0.78-1.08*	1.36	*1.14-1.62*	1.14	*0.94-1.37*
Height	**≤1.60**	1.00		1.00		1.00		1.00		1.00		1.00	
quintile	**>1.62, ≤1.66**	1.02	*0.95-1.09*	1.00	*0.92-1.08*	1.03	*0.86-1.24*	0.99	*0.81-1.21*	1.05	*0.97-1.13*	1.04	*0.95-1.14*
Metres	**>1.66, ≤1.72**	1.00	*0.93-1.08*	0.96	*0.88-1.04*	1.11	*0.94-1.32*	1.07	*0.89-1.29*	0.93	*0.84-1.03*	0.91	*0.81-1.02*
	>1.72, ≤1.78	1.02	*0.94-1.10*	1.00	*0.92-1.09*	1.11	*0.94-1.32*	1.10	*0.91-1.32*	0.98	*0.83-1.16*	1.03	*0.85-1.24*
	>1.78	0.97	*0.89-1.06*	0.97	*0.88-1.07*	1.05	*0.89-1.52*	1.07	*0.88-1.29*	0.85	*0.56-1.33*	0.72	*0.44-4.18*
	Missing	0.94	*0.87-1.02*	1.55	*1.42-1.70*	1.02	*0.86-1.23*	1.80	*1.48-2.18*	0.93	*0.84-1.03*	1.38	*1.22-1.57*

OR Odds ratio; CI Confidence interval. *ORs by smoking quantity adjusted for height, ORs for height adjusted for smoking quantity; † Excludes records within 6 months of cancer or index date

3.4 Discussion

3.4.1 Main findings

The results show a highly significant difference between sexes in the effect of quantity of cigarettes smoked on the odds of developing lung cancer: as a result of ever having smoked heavily women's odds of developing lung cancer are 42% higher than that of men. There was no evidence of an association between height and risk of lung cancer to support the original hypothesis that this difference is due to women having smaller lungs and hence a higher dose of carcinogen per unit lung volume.

3.4.2 Strengths & weaknesses

This study has considerable power with over 12,000 incident cases of lung cancer; it is therefore the largest study to address this issue in an unselected population and also the first UK study on the subject. The generic strengths of studies which use THIN including prospective data collection, data completeness and the validity of lung cancer diagnoses were discussed in section 2.1.5. The sex distribution of lung cancer cases in THIN has been compared and found to be similar to that of the National Lung Cancer Audit. (82)

<u>Smoking</u>

Quantity of cigarettes smoked was defined as the highest quantity ever reported by the patient before their lung cancer diagnosis or matched index date for controls. This method of measuring smoking quantity will not comprehensively represent the variation in patients' lifetime smoking patterns; however it enabled a smoking history to be obtained for over 90% of the population.

There have been changes in patterns of smoking over time and more recently it appears that women are less likely to smoke heavily but also less likely to stop smoking. (8) This was accounted for within the limits of the data by estimating odds ratios for lung cancer based on last reported smoking quantity and status

prior to index date, and significant differences in risk of lung cancer between male and female moderate and heavy smokers were still found.

Histological subtype

The use of a general practice database allowed the identification of a large number of cases and controls and ensured that smoking data were prospectively recorded. However, a weakness of THIN is that there is insufficient information to assess whether the interaction demonstrated applies to all histological types of lung cancer (coding only identifies lung cancer and not the subtype or stage). It is well known that adenocarcinomas are more common than any other histological type in non-smokers and therefore the relationship between sex and smoking for adenocarcinomas may differ from other tumours. This has been examined in some of the previous literature (briefly described in table 3-2) but without any firm conclusions.

3.4.3 Previous research

A summary of previous research on this subject is given in table 3-3 and some of the key studies are summarised below.

Table 3-3: Summary of previous research examining the association between smoking and lung cancer in men and women

Author *(Population)*	**Date published**	**Study design**	**Main findings**	
McDuffie (104) *(Canada)*	1987	Cross-sectional study of 927 lung cancer cases (21% women), diagnosed 1979-83. Retrospectively collected smoking data by postal questionnaire.	Women developed lung cancer at an earlier age while smoking fewer cigarettes and for fewer years than men.	*
Brownson (101) *(United States)*	1992	Registry-based case-control study; 14,596 lung cancer cases diagnosed 1984 - 1990. Controls had other, non-smoking related, cancers. Prospectively recorded smoking data were extracted from hospital records.	Relative risk of lung cancer associated with ever (vs. never) smoking and level of smoking was higher in females than in males overall. For the subgroup of adenocarcinomas there was no significant difference in smoking related risk between men and women.	*
Osann (105) *(United States)*	1993	Registry based case-control study; 1,986 cases (42% female) diagnosed 1984-86. Controls had other, non-smoking related, cancers. Prospectively recorded smoking data were extracted from medical records.	Overall, no statistically significant differences in odds of lung cancer for men and women comparing former or current smokers with never smokers. Odds of small-cell lung cancer relative to never smokers >2-fold higher in women than men but wide confidence intervals and not statistically significant.	
Risch (98) *(Canada)*	1993	Case-control study; 845 cases (52% female) diagnosed 1981-85 identified through monthly examination of medical records in Toronto hospitals. Controls were randomly selected from population listings. Smoking history collected by retrospective questionnaire.	Significantly stronger association between cigarette consumption and lung cancer for females than for males (p=0.01). In 40-pack year smokers relative to non-smokers odds ratio for lung cancer in women was 27.9 (14.9-52) and for men 9.6 (5.64-16.3). Higher odds ratios for females also seen for all histological subgroups.	*
Harris (106) *(United States)*	1993	Case-control study, 4,423 lung cancer cases (34% women). Controls were patients with non-tobacco related diseases from the same hospitals as the cases.	In both black and white populations women were at higher risk of lung cancer than men for each level of smoking compared to baseline never smokers or zero tar consumption.	*

		Retrospective smoking data collected by interview using structured questionnaire.		
Morabia (107) *(United States)*	1991	Hospital based case-control study; 1,358 lung cancer cases (37% women). All interviewed between 1985 and 1990. Retrospective participant reported smoking data collected by interview.	No significant difference in relative risk of lung cancer (assessed in subgroups of four histological types), between men and women when compared with the lightest smokers of the same sex.	
Zang (97) *(United States)*	1996	Hospital based case-control study, 1,887 lung cancer cases (41% female). Controls had non-smoking related diseases all admitted 1981 -1984. Retrospective smoking histories by questionnaire and interviews. Ex-smokers were excluded.	Odds ratios for lung cancer were 1.2 – 1.7-fold higher in women than in men for each level of smoking exposure (pack years, and most recent quantity) compared with never smokers of the same sex. Adjustments for height and weight did not alter the results.	*
Engeland (108) *(Norway)*	1996	Examined trends in the risk of smoking-associated cancers based on registry data comparing 1954-58 with 1989-93; 1,427 lung cancer cases (23% female). Smoking data were obtained from clinical records.	Lung-cancer incidence increased more rapidly, relative to the other smoking-associated cancers, in females than in males.	*
Prescott (109) *(Denmark)*	1998	Cohort study, 867 cases of lung cancer (23% female) diagnosed between 1967 and 1994. Participant-reported smoking data collected on entry to the study and at several points during follow-up.	Incidence rate ratios for lung cancer compared with never smokers of the same sex, did not differ significantly between men and women in for any quantity of cigarettes smoked.	-
Tulinius (110) *(Iceland)*	1997	Cohort study. Lung cancer was a subgroup analysis as the study investigated all cancers; 472 cases identified (42% female) between 1968 and 1995. Prospective smoking data collected on entry.	Relative risk of developing lung cancer for all smoking quantities, compared to never smokers, was increased approximately 2-fold in women compared with men but differences were not statistically significant.	-
Kreuzer (102) *(Germany & Italy)*	2000	Case control study, 4623 lung cancer cases (19% female) diagnosed 1988 - 94. Controls were randomly recruited from the community.	Lung cancer risk in ever compared with never smokers higher in men than women. Further analysis restricted to smokers and adjusted for	-

		Smoking history from retrospective interviews and questionnaires.	quantity and duration found no difference in risk between sexes.	
Bain (100) *(United States)*	2004	Cohort study using two established cohorts one male and one female; 1,266 lung cancer cases identified between 1986 and 2000 (25% female). Prospective smoking data collected on entry to studies.	No significant difference in hazard ratio for lung cancer between men and women within groups of current smokers or former smokers.	-
Henschke (111) *(United States)*	2006	Analysis of data from a screening trial: 269 screen – detected cases of lung cancer, (58% women) between 1993 and 2005. Prospective data on smoking collected on entry to trial.	After adjusting for age and quantity smoked there was a significantly higher risk of lung cancer in women compared with men (odds ratio for lung cancer in women compared with men 1.7 (95% CI 1.3-2.3)	*
Freedman (99) *(United States)*	2008	Cohort study identified 6,324 cases of lung cancer (35% women) between 1996 and 2003. Prospective smoking data collected by postal questionnaire on entry to study.	In current smokers, the hazard ratio for lung cancer in women compared with men, adjusted for smoking quantity was 0.9 (0.8-0.9). In former smokers the same hazard ratio was 0.9 (0.9-1.0).	-
Frueh (112) *(Switzerland)*	2009	Observational cross-sectional study of 683 lung cancer cases (30% women), diagnosed 2004-05. Smoking history from retrospective analysis of medical notes. (Abstract)	Women with lung cancer were more likely to have smoked significantly less than men with lung cancer.	*
Ryu (113) *(Korea)*	2011	Comparison of smoking history between 1,490 men and 104 women diagnosed with lung cancer between 2001 and 2009. Prospective smoking data collected as part of clinical care.	Women with lung cancer had a lower number of pack years compared to men. Women had a significantly higher lung cancer susceptibility index compared with men.	*

Studies marked * found a statistically significant increased risk of lung cancer in females compared with males per quantity smoked; those marked – did not

Studies which support the present findings

Previous research into the difference in effect of smoking quantity on development of lung cancer in men and women includes a study by Zang *et al*, (97) of 1889 cases (781 of whom were female) in which female smokers had a 1.2 – 1.7 fold increase in odds of lung cancer compared with males for all histological subtypes. In a larger case-control study using cancer registry data, 14,596 cases of lung cancer were compared with 36,438 controls who were people diagnosed with other types of cancer during the same time period. (101) The risk of lung cancer was found to be higher in females at each level of smoking, however the authors themselves remark that the validity of their results are limited by a finding of differential misclassification for smoking status among the smoking associated cancers.

Studies which contradict the present results

Bain *et al*. (100) studied the effect of smoking quantity on lung cancer by comparing two separate large previously established cohorts. One of these studies was the Nurses' Health Study which began in 1976 and used postal questionnaires to obtain exposure information on 60,296 women. The second was the 1986 Health Professionals Follow-up Study which also used a postal questionnaire sent to male health professionals aged 40-75 years; there were 25,397 men in the study. Both of these studies were updated by follow-up questionnaires every two years providing updated smoking data and information on newly diagnosed diseases. On comparing these two cohorts Bain *et al* found 1266 incident cases of lung cancer in smokers, having excluded never smokers from the study. No significant differences in hazard ratios for lung cancer between women and men were found when smokers of <25 cigarettes per day were compared with those smoking >25 cigarettes per day who started smoking before and after 20 years of age. This contradicts our findings to some extent, although we compared odds ratios for a different range of smoking habits,

included never smokers as the baseline, and studied a population which was not restricted to health professionals; it is difficult to compare rates of smoking in these cohorts with those of our population due to the exclusion of never smokers.

A more recent prospective cohort study conducted in the United States, (99) also failed to demonstrate a difference in quantity of cigarettes smoked and subsequent risk of lung cancer between the sexes. This study used smoking data collected by questionnaire in the National Institutes of Health – American Association of Retired Persons Diet and Health Study. Smoking status, history and quantity were self-reported on entry to the study by postal questionnaire. Of the 3.5 million people to whom the questionnaire was originally posted, only 17.6% responded, however the questionnaire did include a detailed smoking history allowing the authors to account for the effects of changes in smoking patterns over time. Again, one of the main differences between the population we studied and that of this US study is in selection of participants. The respondents who were included by Freedman *et al* (those who provided complete smoking histories) had an approximately 5% higher incidence of never smokers in both men and women compared to our population which may reflect selection bias in their study or a difference in smoking patterns between people in the US and those in the UK.

The importance of a UK population study is supported by the results of an analysis of 31 studies (more than 480,000 individuals) which suggested that effects of smoking on risk of lung cancer may differ according to sex and country of residence. Hazard ratios for fatal lung cancer in smokers compared to never smokers were 2.48 for Asian men and 2.35 for Asian women, but in the Australia and New Zealand populations 9.87 (95% CI 6.04-16.12) for men and 19.33 (95% CI 10-37.3) for women.(114)

3.4.4 Explaining the difference

There are several possible explanations for the difference in lung cancer risk observed in men and women.

Misclassification of smoking status

It is known that women visit the GP more often than men and from this one could extrapolate that men would be more likely to have missing data. This was true in this study, with 25% of men and 16% of women being known smokers with unknown quantity; however all odds ratios for lung cancer in this category were very similar, strongly suggesting that a difference in missing data does not explain the results.

The prevalence of smoking in men and women in these data is comparable to UK national surveys, (115) however it is still possible that women under-report their smoking quantity during consultations in general practice and that they do this to a greater extent than men. This would partly explain the results but it seems unlikely that the large and significant differences that we have shown can be entirely explained by a difference in reporting between sexes. A higher baseline risk of lung cancer in men could also explain the differences observed in this study, however at least two previous studies have reported similar baseline risk for lung cancer in men and women, (98, 109) and this is supported by the finding of the same incidence of lung cancer in never smokers of both sexes (6%).

It was not possible to account for passive smoking as this is often not recorded in general practice and it is possible that some of the effect demonstrated reflects this.(116) Biomarkers, particularly cotinine, have been used to attempt to validate self-reported smoking status. This has revealed misclassification but mostly confined to trivial smokers and so does not account for the differences seen in moderate and heavy smokers.(117)

Lung volume

Height was used as a surrogate for lung volume, as this is the predominant determinant of total lung capacity, (103) and no difference in lung cancer risk was found across the height quintiles overall or when stratified by sex. The Million Women Study, (118) assessed the effect of height on risk of all cancers in a cohort of middle-aged women between 1996 and 2001. Lung cancer was analysed as a subgroup (approximately 6000 female cases) and whilst overall they found taller stature to be associated with increased relative risk of cancer, for lung cancer they did not find a statistically significant effect. In 2001 a systematic review was published (119) which also found taller people to be more at risk of cancer overall; only one of the ten prospective studies of the effect on lung cancer found a statistically significant increase in relative risk with taller stature (and it is important to note that this study used self-reported height and included just 80 cases). (120) This study addressed this question in a much larger population than any of these previous studies and found no evidence that height has an important effect on the odds of developing lung cancer in males nor in females. Both Bain *et al* and Zang *et al* (discussed above) reported that height, and for Zang *et al* body mass index and weight, had no effect on risk of lung cancer in their populations.

Breathing patterns

An alternative explanation for the study findings is related to evidence that men and women have different breathing, and thus smoking, patterns. Perhaps carcinogens are deposited in female lungs at a higher concentration because of the way they smoke. Ragnarsdottir *et al*. demonstrated a difference in breathing pattern in their study of 100 healthy subjects; whilst during quiet breathing the pattern appears similar, during deep breathing women tend to use their abdominal muscles less than men and their ratio of inspiration to expiration differed. (121) Studies specifically investigating differences in the way people

smoke cigarettes,(122, 123) have found that women tend to have a smaller puff volumes and take longer between puffs, resulting in a lower nicotine exposure, however this ought to make women less, rather than more, susceptible to lung cancer and therefore would not explain most of the study findings.

Metabolic differences

Hormonal effects, for example the effects of oestrogens on development lung tumours, should be considered as a potential explanation for the results. There may be differences in the metabolism of carcinogens and the overall susceptibility to lung cancer between men and women, or oestrogens may affect tumour growth and type. Women with lung cancer have been shown to have better overall survival, (124) this could also be due to a difference in the way tumours exposed to oestrogens behave biologically.

3.4.5 Conclusion

Women are more likely to develop lung cancer than men who smoke similar quantities. This is not explained by the fact that women have smaller lungs than men as there was no effect of height on the odds of having lung cancer. That women have a greater risk of developing lung cancer due to smoking should be taken into account when estimating risk of lung cancer in the context of early referral of symptomatic patients and possibly in lung cancer screening. It is not clear why women are more susceptible to smoking and this merits further research.

Smoking prevalence in women increased from about 1945 and although it never reached the same peak as in men, did not start to decline until the mid-1970s. Lung cancer incidence in women continues to climb gradually and, unlike in men, we have not yet seen the peak (Chapter 1, figure 1-2). (5) These results raise concern that the trajectory of the disease in women may not follow that of men. Whilst we are expecting to see a smoking cessation related decline in lung cancer incidence and mortality in women, if women are more susceptible to

cigarette smoke we may not see this to the same extent that we have witnessed in the male population.

The findings from this study should draw attention to the previously unrecognised direction that the lung cancer epidemic is taking and strongly support the call for smoking cessation programs specifically aimed at women. (125) Reasons for the increased susceptibility of women to cigarette smoke remain unclear but we now have robust evidence for an increased effect of quantity of smoking in women in heart disease (96) and, from this study, in lung cancer. The impact of smoking in women has, until now, been underestimated.

3.5 Chapter summary

In this chapter I have described the use of THIN data to provide further evidence that, for each smoking quantity, women are at higher risk of lung cancer than men. These findings are important because projected lung cancer incidence figures are based entirely on the male population and may therefore be underestimates, and also because in some countries smoking prevalence in women continues to increase. This work was published in CHEST in January 2013.(126)

In undertaking this study I have developed skills in manipulating and interrogating a large database, and acquired knowledge of statistical analyses including logistic regression and analyses of interaction. In the next chapter I will use the same case-control database to investigate the controversial issue of chronic obstructive pulmonary disease as an independent risk factor for lung cancer.

CHAPTER 4: IS CHRONIC OBSTRUCTIVE PULMONARY DISEASE AN INDEPENDENT RISK FACTOR FOR LUNG CANCER?

This chapter describes a second study using the lung cancer case-control dataset from THIN. In this chapter I investigate in detail the association between chronic obstructive pulmonary disease and lung cancer, specifically focusing on adequate adjustment for smoking and the timing of diagnoses of COPD in relation to lung cancer.

4.1 Introduction

4.1.1 Background

Chronic obstructive pulmonary disease (COPD) and lung cancer are two of the most important smoking related diseases worldwide, with a huge combined mortality burden.(127, 128) Many consider COPD to be an independent risk factor for lung cancer, (11, 129) but others argue that they are just manifestations of the same exposure.

4.1.2 Rationale for this study

Brenner *et al* recently published a meta-analysis of the relationship between lung cancer and prior lung diseases.(11) In all of the 39 studies assessing COPD, efforts were made to adjust for smoking. The majority of these studies reported that COPD was associated with an increase in risk of lung cancer, the highest reported increase being 9-fold,(129) although a few, including one study in never smokers, (130) showed reduced risks. The combined relative risk of lung cancer in people with a diagnosis of COPD, chronic bronchitis or emphysema, compared with people without these diagnoses, was 1.83 (95% confidence interval (CI) 1.6-2.11) but the authors reported significant heterogeneity across studies particularly in the populations studied and in definitions of COPD.(11)

When patients with lung cancer first present to a clinician their symptoms may be consistent with a new diagnosis of COPD, and may be recorded as such before the diagnosis of lung cancer is made. Patients referred to secondary care for suspected lung cancer may additionally be investigated for COPD. For this reason in studies of COPD and lung cancer there is likely to be strong ascertainment bias. Brenner *et al* were not able to account for this in their meta-analysis as many studies did not have data on when diagnoses were made.

The most likely explanation for the alleged link between COPD and lung cancer is airway inflammation, (131) and therefore the effect of asthma on lung cancer has also been investigated, with conflicting results.(132-134)

The identification of new factors which contribute to the aetiology of lung cancer are important in identifying patients who will benefit the most from screening, smoking cessation and perhaps chemoprevention (some studies have suggested that the use of statins may reduce the risk of lung cancer, (135)). If there is good evidence of an increased risk of lung cancer using a general practice diagnosis of COPD this would be of great benefit to those working in primary care who decide whether to refer patients for further investigations.

4.1.3 Aim of this chapter

This study used the prospectively collected GP data in THIN to quantify the association between COPD and lung cancer in the UK population, whilst accounting for smoking and the impact of timing of diagnoses. To assess the specificity of any association between COPD and lung cancer other common pulmonary diseases (asthma and pneumonia) were also considered.

4.2 Methods

4.2.1 Study population

This study was based on the lung cancer case control dataset described in section 2.1.3; however for this study patients had to contribute data to THIN for at least the year leading up to their index date to be included. In the previous study (Chapter 3) it was not absolutely necessary for all controls to contribute data in the few months prior to their index date, provided they contributed at least one year of data in total. In this study, which assessed the risk of lung cancer according to history of other respiratory conditions, this was necessary because people with each diagnosis were divided into categories according to when they were first diagnosed.

4.2.2 Definition of Exposures

<u>COPD, pneumonia and asthma</u>

The main exposures of interest were a history of COPD, asthma, or pneumonia. Read codes for COPD, asthma and pneumonia (Appendix C) were used to identify patients with these diagnoses before the lung cancer diagnosis or index date. Dates of first diagnosis were grouped into diagnostic latency categories: within 6 months, between 6 months and 1 year, between 1 year and 5 years, between 5 years and 10 years and 10 years or more before the lung cancer diagnosis or index date.

4.2.3 Covariate definitions

Smoking

Smoking status and smoking quantity were defined using the additional health data file as described in section 2.1.4. Smoking was categorised as: Never, light (1-9 cigarettes per day), moderate (10-19 cigarettes per day), heavy (20 or more cigarettes per day), smoker with missing quantity, or missing smoking status. Where there was more than one smoking record the highest smoking quantity recorded before the lung cancer or index date was used.

Socio-economic status

Townsend quintile, as described in section 2.1.4, was used to define socio-economic status. Quintile 1 represents the most affluent and quintile 5 the most deprived people.

COPD severity

Severity of COPD was represented by the percentage of predicted forced expiratory volume in 1 second (FEV1) and grouped as recommended in the 2010 National Institute for Health and Clinical Excellence (NICE) guidelines for the management of COPD in adults. (136) FEV1 >80% of predicted with a read code for COPD was classed as mild, 50-80% moderate, 30-50% severe and <30% very severe.

Records of FEV1 were extracted from the database, excluding any measurements not recorded in litres or millilitres. The most recent measurement was used, excluding measurements taken after, or within 6 months of, the lung cancer or index date as these values may represent cancer-induced changes in lung function. The patient's age at the time of the measurement and height record closest to that date were used to calculate predicted FEV1 using the equations published by Crapo et al. (137)

4.2.4 Statistical methods

Conditional logistic regression was used to estimate odds ratios of lung cancer associated with a prior diagnosis of COPD (overall and by severity), asthma and pneumonia, according to the timing of diagnoses. Any changes in effect after adjusting for smoking and Townsend quintile were explored.

To account for any diagnostic uncertainty or overlap between asthma and COPD, patients who had records of both were identified and analyses repeated in those exclusively with COPD or asthma.

4.3 Results

The initial lung cancer case-control dataset contained information on a total of 60,337 people: 12,121 incident cases of lung cancer between January 2000 and July 2009, and 48,216 matched controls. Following the exclusion of 10,442 controls for which data recording ended before the cancer diagnosis or index date, it also was necessary to exclude 174 cases for which no controls remained. A further 228 patients with age at diagnosis <40 years were excluded and this left a total of 11,888 cases and 37,605 controls in the analysis (overall N=49,493) with 5,256 cases matched with 4 controls, 4,008 cases with 3 controls, 1,933 cases with 2 controls and 691 cases with one control.

Patients had a median of 9.6 years (interquartile range 5.7 - 13.5 years) of prospectively recorded general practice data before their index date; cases had a median of 9.5 years and controls 9.4 years.

Fifty-nine per cent of cases were male and the majority (59%) were over 70 years old at diagnosis. Cases had greater socioeconomic deprivation than controls, with 19% of cases being in the highest Townsend quintile compared with 14% of controls (Table 4-1).

People with cancer were more likely to smoke (90% of cases compared with 61% of controls had ever smoked), and were more likely to have smoked heavily (38% and 13% respectively). Smoking status was missing for 4% of cases and 9% of controls.

Table 4-1: Description of cases and controls

		Cases (11,888)		Controls (37,605)	
		n	%	n	%
Sex (matched)	Female	4,863	*40.9*	15,639	*41.6*
	Male	7,025	*59.1*	21,966	*58.4*
Age at diagnosis	40-49	315	*2.6*	1,056	*2.8*
(years, matched)	50-59	1,366	*11.5*	4,732	*12.6*
	60-69	3,229	*27.2*	10,946	*29.1*
	70-79	4,501	*37.9*	14,499	*38.6*
	≥80	2,477	*20.8*	6,372	*16.9*
Townsend quintile	1 (least deprived)	2,037	*17.1*	8,735	*23.2*
	2	2,200	*18.5*	8,187	*21.8*
	3	2,380	*20.0*	7,420	*19.7*
	4	2,609	*21.9*	6,742	*17.9*
	5 (most deprived)	2,196	*18.5*	5,064	*13.5*
	missing	466	*3.9*	1,457	*3.9*
Smoking	Never	1,176	*9.9*	14,527	*38.6*
Highest ever	Trivial / light	1,010	*8.5*	2,409	*6.4*
recorded prior	Moderate	233	*2.0*	3,788	*10.1*
to index date	Heavy / very heavy	4,516	*38.0*	4,829	*12.8*
	Smoker but unknown quantity	2,236	*18.8*	8,710	*23.2*
	Missing smoking status	517	*4.3*	3,342	*8.9*

COPD, pneumonia and asthma

Cases were nearly four times as likely to have had a prior diagnosis of COPD overall (23% compared with 6% of controls) and across all diagnostic latency categories with the most marked difference in the 6 months prior to lung cancer diagnosis (3.4% compared with 0.4% of controls) (Table 4-2 and Figure 4-1).

The prevalence of pneumonia was also higher in cases than in controls, and displayed a similar pattern to that of COPD with more marked differences closer to the time of lung cancer diagnosis (Table 4-2 and Figure 4-2).

Asthma was more prevalent in cases than controls however this was less marked than for COPD or pneumonia. There was not a clear peak in asthma diagnoses just before lung cancer or index date as there was for COPD or pneumonia (Table 4-2 and Figure 4-3).

Table 4-2: Prior diagnoses of COPD, pneumonia and asthma in cases and controls

		Cases (11,888)		Controls (37,605)	
		n	**%**	**n**	**%**
COPD	No diagnosis prior to index date	9,131	*76.8*	35,319	*93.9*
Interval between	within 6 months	404	*3.4*	140	*0.4*
first diagnosis &	6 months up to 1 year	199	*1.7*	172	*0.5*
index date	1 year up to 5 years	1,033	*8.7*	947	*2.5*
	5 years up to 10 years	690	*5.8*	580	*1.5*
	10 years or more	431	*3.6*	447	*1.2*
Asthma	No diagnosis prior to index date	9,893	*83.2*	33,640	*89.5*
Interval between	within 6 months	111	*0.9*	108	*0.3*
first diagnosis &	6 months up to 1 year	71	*0.6*	130	*0.3*
index date	1 year up to 5 years	616	*5.2*	1,093	*2.9*
	5 years up to 10 years	620	*5.2*	1,188	*3.2*
	10 years or more	577	*4.9*	1,446	*3.8*
Pneumonia	No diagnosis prior to index date	10,819	*91.0*	36,540	*97.2*
Interval between	within 6 months	378	*3.2*	88	*0.2*
first diagnosis &	6 months up to 1 year	74	*0.6*	79	*0.2*
index date	1 year up to 5 years	318	*2.7*	422	*1.1*
	5 years up to 10 years	164	*1.4*	220	*0.6*
	10 years or more	135	*1.1*	256	*0.7*

COPD Chronic obstructive pulmonary disease

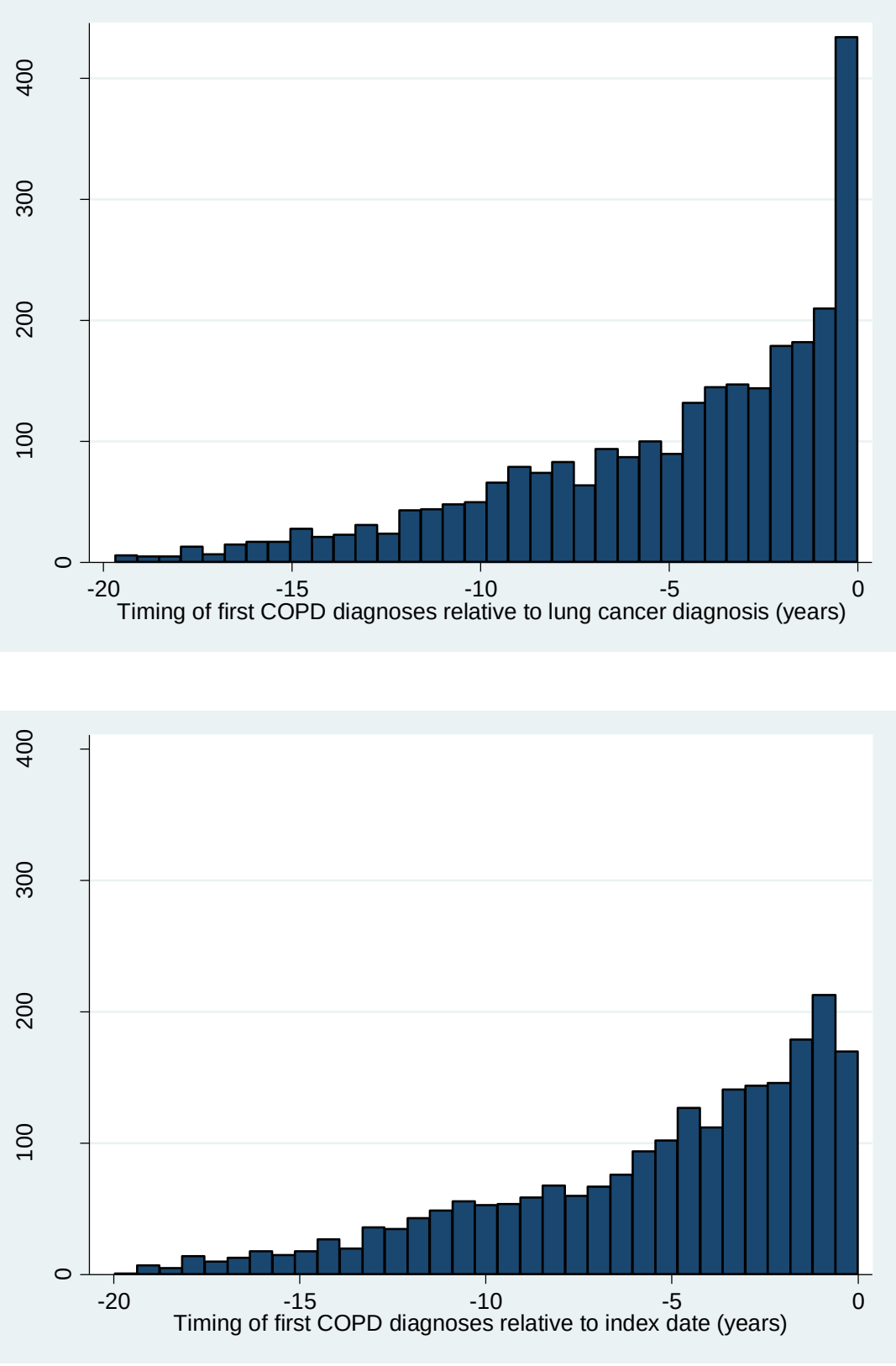
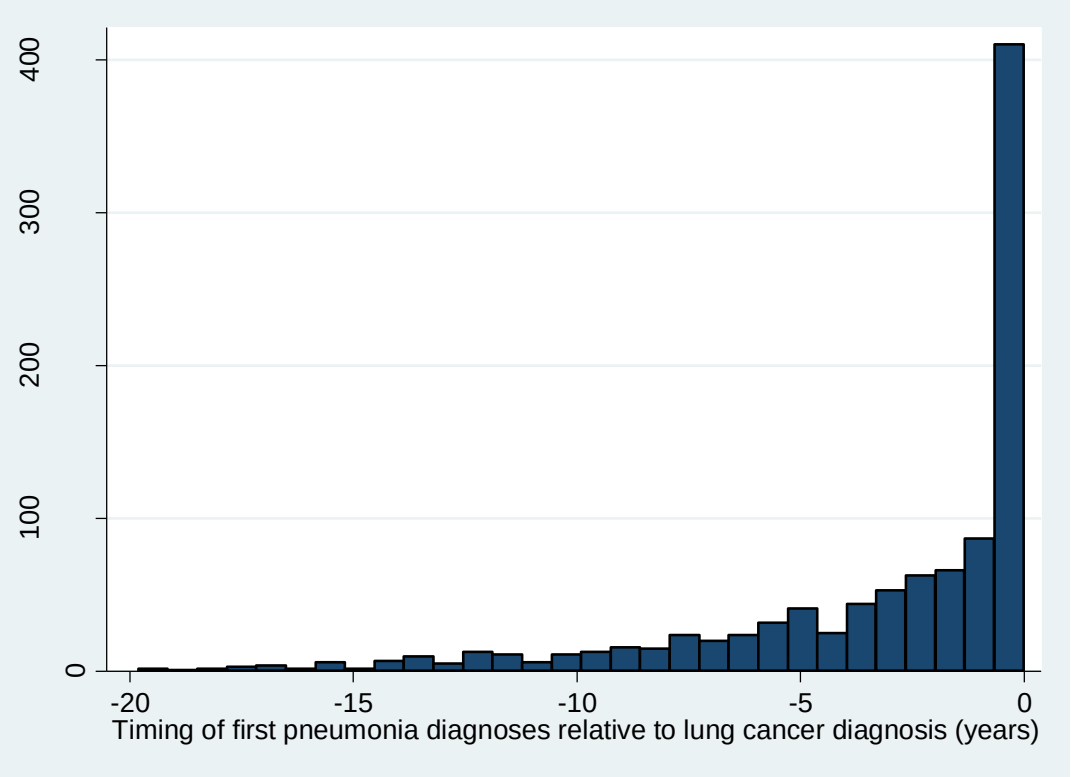


Figure 4-1: Timing of first diagnoses of COPD in cases and controls



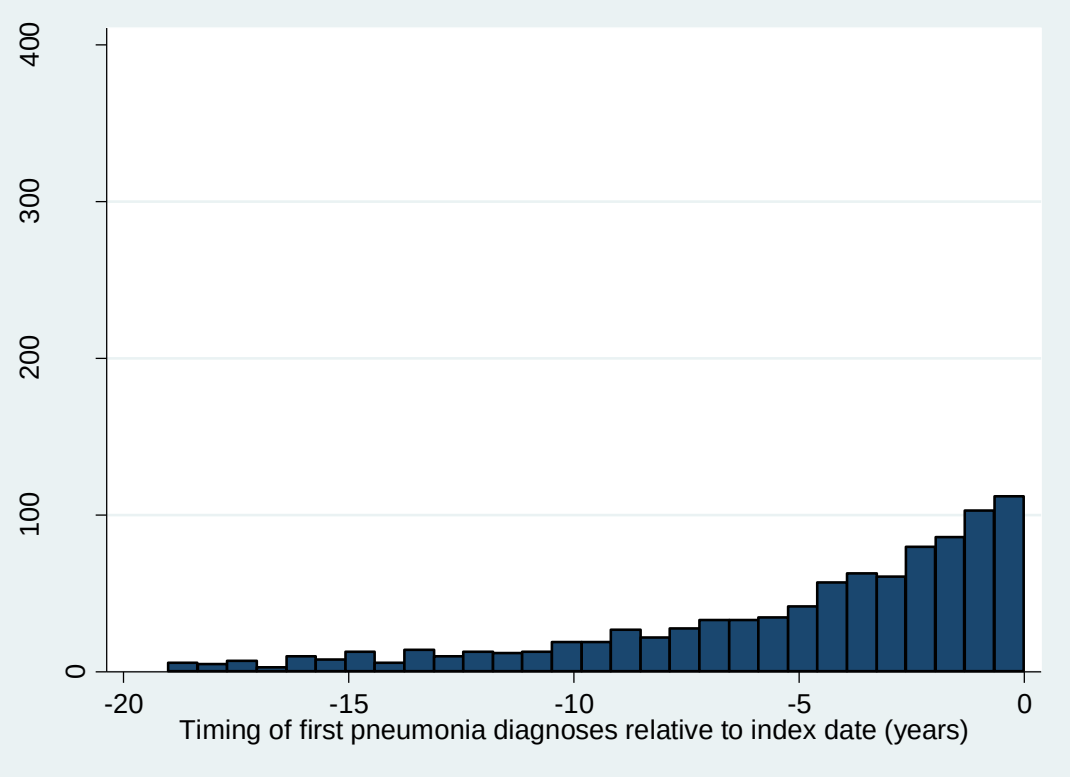
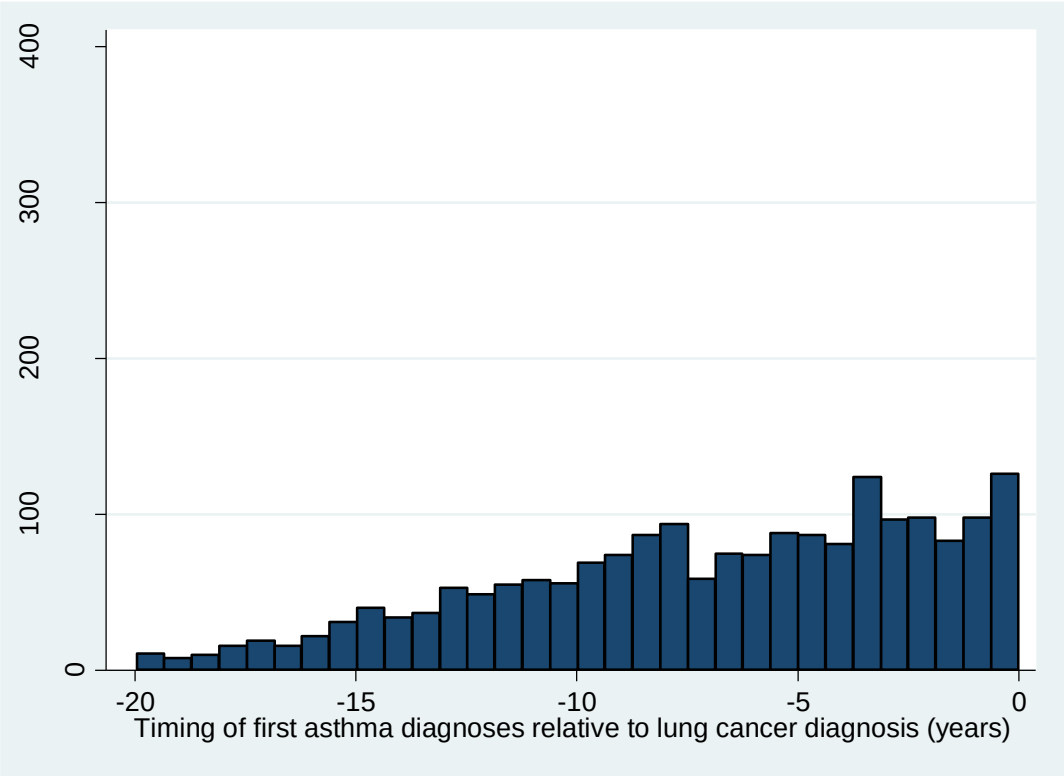


Figure 4-2: Timing of first diagnoses of pneumonia in cases and controls



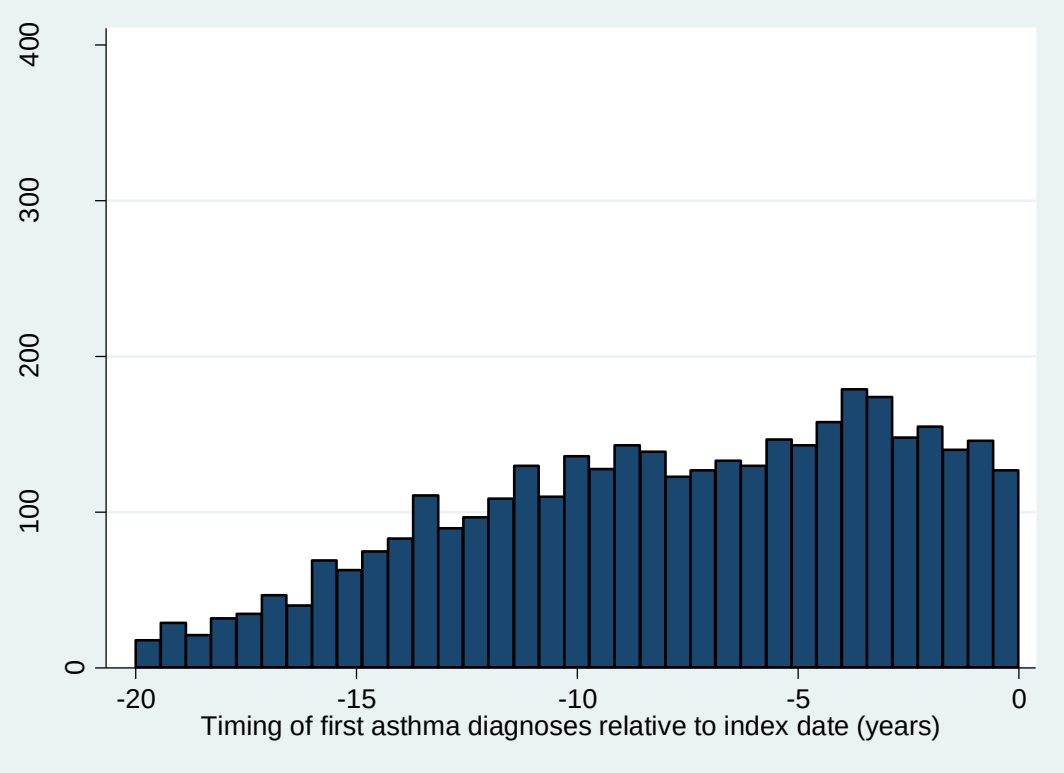


Figure 4-3: Timing of first diagnoses of asthma in cases and controls

4.3.1 Risk factors for lung cancer

Smoking

Smoking was strongly associated with lung cancer with odds in the heaviest smokers over 15 times those in never smokers (odds ratio (OR) 15.58, 95% CI 14.35-16.91). The association between socio-economic status (Townsend quintile) and lung cancer was confounded by smoking but adjusted odds ratios remained significantly increased in the most deprived group compared with the least deprived (OR 1.58, 95% CI 1.44-1.73) (Table 4-3).

COPD

COPD diagnoses made within 6 months of the index date were associated with an 11-fold increase in the odds of lung cancer compared with no prior COPD diagnosis, however this was heavily confounded by smoking and the odds ratio was 6.81 (95% CI 5.49-8.45) after adjusting for smoking and socio-economic status. The timing of diagnosis also had a considerable effect, with the adjusted odds ratio falling to 2.18 (95% CI 1.87-2.54) when diagnoses made within 10 years of the lung cancer or index date were excluded.

Pneumonia

A previous diagnosis of pneumonia was associated with increased odds of lung cancer with the pattern regarding diagnostic timing similar to that seen in COPD (Table 2). There was a very strong association with diagnoses of pneumonia made within 6 months of index date (OR 14.91, 95% CI 11.75-18.94) and there was evidence of confounding by smoking but to a lesser degree than in COPD (adjusted OR for diagnoses within 6 months 13.33, 95% CI 10.24-17.35). There remained an association when diagnoses of pneumonia made within 10 years of lung cancer or index date were excluded (adjusted OR 1.46, 95% CI 1.15-1.86).

Table 4-3: Odds ratios for lung cancer according to patient characteristics and previous respiratory diseases

N=49,493 *11,888 cases and 37,605 controls*		**Odds ratio (OR)** ***95% CI***		**Adjusted OR*** ***95% CI***	
Townsend quintile	1 (least deprived)	1.00		1.00	
	2	1.19	*1.11-1.28*	1.15	*1.06-1.24*
	3	1.48	*1.38-1.59*	1.28	*1.18-1.38*
	4	1.88	*1.75-2.03*	1.43	*1.32-1.55*
	5 (most deprived)	2.26	*2.08-2.44*	1.58	*1.44-1.73*
	Missing	1.65	*1.41-1.94*	1.20	*1.00-1.43*
Smoking	Never	1.00		1.00	
Highest ever recorded prior to index date	Light	6.00	*5.42-6.65*	5.88	*5.31-6.52*
	Moderate	9.67	*8.87-10.54*	9.33	*8.56-10.18*
	Heavy	15.58	*14.35-16.91*	14.88	*13.71-16.16*
	Unknown quantity	3.48	*3.20-3.78*	3.44	*3.17-3.74*
	Missing	1.79	*1.59-2.02*	1.76	*1.56-1.99*
COPD	None**	1.00		1.00	
Interval between first diagnosis & index date	within 6 months	11.47	*9.38-14.02*	6.81	*5.49-8.45*
	6 months up to 1 year	4.76	*3.85-5.89*	2.52	*2.00-3.19*
	1 year up to 5 years	4.34	*3.95-4.78*	2.48	*2.24-2.75*
	5 years up to 10 years	4.83	*4.29-5.44*	2.68	*2.36-3.05*
	10 years or more	3.74	*3.25-4.31*	2.18	*1.87-2.54*
Asthma	None**	1.00		1.00	
Interval between first diagnosis & index date	within 6 months	3.63	*2.77-4.76*	2.92	*2.15-3.97*
	6 months up to 1 year	1.94	*1.44-2.60*	1.51	*1.08-2.12*
	1 year up to 5 years	1.91	*1.72-2.12*	1.65	*1.20-1.51*
	5 years up to 10 years	1.83	*1.65-2.03*	1.43	*1.27-1.60*
	10 years or more	1.33	*1.20-1.47*	1.19	*1.06-1.33*
Pneumonia	None**	1.00		1.00	
Interval between first diagnosis & index date	within 6 months	14.91	*11.75-18.94*	13.33	*10.24-17.35*
	6 months up to 1 year	3.37	*2.42-4.70*	2.89	*1.99-4.18*
	1 year up to 5 years	2.59	*2.22-3.02*	2.16	*1.82-2.57*
	5 years up to 10 years	2.52	*2.04-3.10*	2.11	*1.66-2.67*
	10 years or more	1.68	*1.35-2.09*	1.46	*1.15-1.86*

OR, Odds ratio. CI, confidence interval. COPD, Chronic obstructive pulmonary disease. *Adjusted for smoking & Townsend quintile – odds ratios by smoking quantity are adjusted for Townsend quintile and those by Townsend quintile are adjusted for smoking quantity. ** No diagnosis prior to index date.

Asthma

The strength of association between a diagnosis of asthma and lung cancer was less than that of COPD or pneumonia, but there was still evidence of confounding by smoking and an effect of timing of diagnosis. After adjusting for smoking and socio-economic status, and when excluding diagnoses made within 10 years of lung cancer diagnosis odds of lung cancer were 1.19 (95% CI 1.06-1.33) times those in people without asthma.

4.3.2 Diagnostic overlap

Of patients with a prior COPD diagnosis, 40% (60% of cases and 33% of controls) also had a diagnosis of asthma, showing considerable diagnostic overlap. The effect of COPD on lung cancer remained after excluding the patients who also had asthma, with similar odds ratios to those shown in Table 4-3 across all diagnostic time windows (Table 4-4).

Table 4-4: Odds ratios for lung cancer in patients with record of COPD without a record of asthma

N for conditional logistic regression = 43,404		**Odds ratio (OR)** ***95% CI***		**Adjusted OR*** ***95% CI***	
No COPD diagnosis prior to index date		1.00		1.00	
Interval between first diagnosis & index date	**within 6 months**	14.38	*11.16-18.53*	7.89	*6.03-10.34*
	6 months up to 1 year	4.83	*3.69-6.32*	2.42	*1.80-3.25*
	1 year up to 5 years	4.93	*4.33-5.61*	2.81	*2.44-3.23*
	5 years up to 10 years	5.47	*4.55-6.58*	2.89	*2.37-3.53*
	10 years or more	3.96	*3.14-4.99*	2.23	*1.74-2.87*

Excludes 2,395 people with diagnoses of both COPD and asthma recorded
OR, Odds ratio. CI, confidence interval. COPD, Chronic obstructive pulmonary disease
**Adjusted for smoking & Townsend score.*

However when the effect of asthma on lung cancer risk was assessed excluding people who also had a diagnosis of COPD no independent association was found except in the most proximal diagnoses (Table 4-5).

Table 4-5: Odds of lung cancer in patients with record of asthma without a record of COPD

N for conditional logistic regression = 43,404		**Odds ratio (OR)** ***95% CI***		**Adjusted OR*** ***95% CI***	
No asthma diagnosis prior to index date		1.00		1.00	
Interval between first diagnosis & index date	**within 6 months**	3.39	*2.37-4.85*	3.22	*2.11-4.93*
	6 months up to 1 year	1.19	*0.76-1.86*	1.11	*0.66-1.86*
	1 year up to 5 years	1.10	*0.94-1.29*	0.99	*0.83-1.19*
	5 years up to 10 years	1.03	*0.88-1.20*	0.96	*0.81-1.14*
	10 years or more	0.74	*0.64-0.85*	0.78	*0.66-0.92*

*OR, Odds ratio. CI, confidence interval. *Adjusted for smoking & Townsend score.*
Excludes 2,395 people with diagnoses of both COPD and asthma recorded

4.3.3 COPD severity

There odds of lung cancer increased significantly with increasing severity of COPD (p=0.0221), however it was necessary to exclude over 50% of patients with COPD from this analysis due to missing data: Fifty-seven per cent of cases and 54% of controls had insufficient data to calculate percentage predicted FEV_1 when measurements taken within 6 months of index date were excluded. The adjusted odds of lung cancer for patients with missing lung function data was increased compared to all categories of severity suggesting that people with more severe disease may be less likely to have their lung function recorded.

Table 4-6: COPD severity based on records of lung function

5,043 people had COPD but for conditional logistic regression N=1,183		**Cases (n=2,757)**		**Controls (n=2,286)**		**Unadjusted OR**		**Adjusted OR***	
		n	**%**	**n**	**%**	***95% CI***		***95% CI***	
FEV_1	**>80%**	60	*2.2*	84	*3.7*	*1.00*		1.00	
	50-80%	499	*18.1*	467	*20.4*	1.36	*0.68-2.74*	1.35	*0.66-2.77*
	30-50%	472	*17.1*	373	*16.3*	2.09	*1.01-4.31*	1.94	*0.92-4.07*
	<30%	158	*5.7*	131	*5.7*	2.24	*0.95-5.30*	2.21	*0.92-5.31*
	Missing	1,568	*56.9*	1,231	*53.8*	2.94	*1.42-6.07*	2.92	*1.38-6.15*

COPD, chronic obstructive pulmonary disease. OR, odds ratio. CI, confidence interval.
****Adjusted for smoking***

4.4 Discussion

4.4.1 Main findings

There is a strong association between COPD and lung cancer but this is largely explained by the effect of smoking and is most apparent in recently diagnosed cases of COPD suggesting a strong element of ascertainment bias. The association between pneumonia and lung cancer followed a very similar pattern with a strong association for proximal diagnoses and less confounding by smoking. The effect of timing of diagnosis of asthma was similar to that observed with COPD and pneumonia, however after accounting for smoking, diagnostic overlap and ascertainment bias there was no evidence of an independent association between asthma and lung cancer.

4.4.2 Strengths

The strengths of this study are the large, unselected population on which it is based and the prospective recording of data, as discussed in section 2.1.5. Lung cancer diagnoses in THIN were found to be valid when compared with national registry data, (65) and whilst the validity of a GP diagnosis and recording of COPD in general practice data has not been tested to date, previous work has shown that demographics and smoking habits of patients with COPD in THIN are consistent with those in the UK population confirmed as having the disease. (138) Over 90% of people in this study had records of smoking status available and smoking prevalence in THIN has been shown to be comparable to that predicted by the General Household Survey. (139) In addition the strength of the association between smoking and lung cancer in this study was as expected.

4.4.3 Smoking and ascertainment bias

Further strengths of this study compared with previous work are that it incorporates both prospective recording of smoking data and the close examination of the timing of diagnoses of COPD in relation to lung cancer in

order to demonstrate the effects of ascertainment bias. These issues are discussed below in the context of other studies.

Smoking

Adequate adjustment for smoking and particularly for any modification of smoking behaviour after the onset of symptoms of lung cancer is difficult, with patient reported smoking being the only method by which this exposure can be measured. Retrospective collection of smoking data after the diagnosis of lung disease, particularly lung cancer, is subject to recall and reporting bias but even in studies such as this one that use prospectively recorded data, patients' current and past smoking habits, and particularly passive smoke exposure, may be inadequately reported or recorded.

Smoking has a massive effect on lung cancer risk and this is potentially why, even after adjusting for smoking habit and relative quantity smoked, results from this and other studies still suggest an association between lung cancer and other smoking related lung diseases (COPD and to a lesser extent pneumonia) but not with asthma (sufferers of which are less likely to be smokers). Residual confounding by smoking could also explain the graded increase in risk of lung cancer with each level of socioeconomic deprivation in this study and in previous research.(140, 141)

In an attempt to address the problem of adequate adjustment for smoking, Turner *et al* used data from 448,600 individuals who reported to be never smokers in the baseline survey of the United States Cancer Prevention Study II, which then ascertained cancer deaths over the following 20 years.(142) Information on prior lung disease was obtained at baseline from participant self-reports of doctor diagnoses. Lung cancer mortality was not associated with chronic bronchitis, but was with emphysema (hazard ratio (HR) 1.66 (95%CI 1.06-2.59)) or combined emphysema and chronic bronchitis (HR 2.44 (95% CI 1.22-4.90). However, of the 1,759 who died from lung cancer, those who initially

reported diagnoses of chronic bronchitis, emphysema or both were only 48, 20 and 8 people respectively.

A more recent study of women in Hong Kong did not detect an association between obstructive lung disease and mortality due to lung cancer in a subgroup of never smokers (HR 0.97, p = 0.909) but also had a very small number of exposed cases. (143) The use of lung cancer mortality as opposed to incident cases is potentially problematic when assessing whether COPD modifies susceptibility because those without COPD may be fitter and therefore more likely to have life-prolonging surgery or chemotherapy.

Ascertainment bias

By studying the timing of COPD diagnosis in relation to lung cancer, it was possible to clearly demonstrate the importance of clinical ascertainment bias that may result from those with symptoms of lung cancer undergoing more investigations and clinical assessment than those without, resulting in a diagnosis of COPD in the few weeks or months before the diagnosis of lung cancer is made.

In addition, people with any chronic lung disease, particularly COPD, will be monitored with more regular contact and investigations by health professionals, providing greater opportunity for a subsequent diagnosis of lung cancer. We know that most people with lung cancer present to health services quite late, and therefore it is feasible that the remaining association between lung cancer COPD diagnoses made 5 and 10 years prior could be explained by such ascertainment.

Some of the previous studies on this subject did assess diagnostic latency periods (summarised in Table 4-7), often in subgroup analyses, but they have relatively small numbers compared with the present study.(130, 144) In a Chinese study, ORs for lung cancer were 2.9 (95% CI 2.0-4.1) and 1.9 (1.2-3.1)

in people diagnosed with COPD 1-5 and 6-10 years before, which are similar to the results of this study, however there were only 74 and 32 cases in each of these groups.(144) In the study by Turner *et al* of never smokers, authors excluded deaths in the first 5 years follow-up, however as reported above, their overall number of exposed cases was very small.(142)

4.4.4 Limitations

COPD severity

Missing data makes the analysis according to COPD severity difficult to interpret as over half of all patients with COPD had insufficient data to calculate a percentage of predicted FEV_1. The adjusted odds of lung cancer for patients with missing lung function data (Table 4-5) was increased compared to all categories of severity suggesting that people with more severe disease may be less likely to have their lung function recorded.

There was a suggestion that people with more severe disease are more likely to develop lung cancer, however it is possible that ascertainment bias also affects this result if people with more severe disease are more likely than those with mild disease to be admitted to hospital and/or undergo investigations which result in the diagnosis of lung cancer.

Definition of COPD

Diagnoses of COPD, chronic bronchitis and emphysema may be based on symptoms, pulmonary function testing or radiological imaging, and a history of COPD may be identified by patient or physician-reported diagnoses, or by performing imaging or lung function tests on every participant.

The use of general practice records to identify prior lung disease in this study removed recall bias and any errors due to inaccuracy in patient's perceptions or knowledge of their prior lung disease, however it could be argued that a more objective measure of COPD such as airflow obstruction on spirometry or

radiographic evidence of emphysema is a more accurate method of defining COPD. Table 4-7 shows how definitions of COPD vary widely and that results also vary, but not consistently, according to the definition used.

Additional data fields

Histological type is not recorded in this general practice database however it would have been interesting to investigate whether the findings would have been modified by histological subtype given the different associations between smoking and squamous or adenocarcinoma. Cancer stage is also not recorded and therefore it was not possible to determine whether people with prior lung disease were diagnosed at an earlier or later stage than those without.

Occupation is recorded very infrequently in THIN so it was not possible to assess its effect; however two previous studies of benign lung diseases and lung cancer risk showed that adjusting for occupation or exposure to dusts or asbestos fibres made little difference.(145, 146)

4.4.5 Summary of previous studies

Table 4-7 summarises previous studies on COPD and risk of lung cancer, some of which have already been discussed. The methods of defining COPD, whether or not the authors considered the possibility of ascertainment bias and the way in which smoking data were collected are briefly described, as well as the overall study design and outcomes.

Table 4-7: Summary of previous studies investigating the association between COPD and lung cancer

Author *(Population)*	**Date published**	**Study design**	**Definition of COPD**	**Smoking data**	**Main findings after adjusting for smoking**	**Considered ascertainment bias?**
Tockman (147) *(North America)*	*1987*	Cohort of people with airflow obstruction (AFO) compared with people without AFO another cohort. Outcome: lung cancer deaths.	AFO defined as FEV_1 <60% predicted on entry to cohort.	Prospective on entry to cohort.	27 lung cancer deaths in AFO group, 14 in non-AFO group. Relative risk of lung cancer death: 2.57 in people with AFO compared to fev1>85%.	Chest x-ray performed on entry to cohorts with aim of excluding pre-existing lung cancer
Wu Williams(148) *(Chinese women)*	*1990*	Case control study: 965 female lung cancer cases, and 959 female general population controls.	Participant reported history of chronic bronchitis and/or emphysema	Retrospective participant interviews	Relative risk of lung cancer 1.4 (95% CI 1.2-1.8) in those with COPD compared with those without. In subgroup analysis effect only in squamous cell lung cancers, not in adenocarcinomas.	Yes – excluded lung disease within 3y of cancer diagnosis
Islam (149) *(United States)*	*1994*	Lung cancer deaths (n=77) in a cohort study. Assessed incidence of lung cancer in quartiles of baseline FEV_1	Per cent of predicted FEV_1 measured on entry to cohort study	Prospective on entry to cohort study	Among smokers, those in lowest quartile of FEV_1 had 2.7 times increased risk of lung cancer compared with highest quartile	Yes – patients who developed cancer within 1 year of entry were excluded
Wu (150) *(US women)*	*1995*	Case control study: 412 lung cancer cases and 1,253 population controls	Participant reported physician diagnoses of chronic bronchitis or emphysema	Retrospective participant interviews	Chronic bronchitis (but not emphysema – small numbers) was associated with increased risk of lung cancer (OR 1.60, 95% CI 1.1–2.4)	Yes – But small no.s when divided by latency category and associations no longer significant
Brownson (151) *(US women)*	*2000*	Case control study: 676 lung cancer cases from cancer registry and general population controls.	Participant reported physician diagnoses of chronic bronchitis or emphysema	Retrospective patient interview or questionnaire	Increased risk of lung cancer with chronic bronchitis (OR 1.7, 95% CI 1.2-2.3) and emphysema (OR 2.7, 95% CI 1.8-4.2).	Yes - but when proximal diagnoses were excluded only emphysema was significantly associated
Brenner (144) *(China)*	*2001*	Case control study, 886 cases and 1,968 controls randomly sampled from population census list	Patient reported physician diagnosis of chronic bronchitis or emphysema (COPD)	Retrospective patient interview	Increased risk of lung cancer in people with COPD: OR 1.4 (95% CI 1.1-1.8)	Yes – excluding COPD diagnoses 1-5 years before lung cancer made no difference

Kishi (129) *(US Screening trial)*	*2002*	Case control study: 24 cases of lung cancer identified by screening. Controls matched on smoking history.	Emphysema on CT and FEV_1 by spirometry on entry to trial.	Prospectively collected at interview on entry to trial	OR for lung cancer 9.6 (95% CI 1.5-60.1) if fev1<40% compared to >80% (although other degrees of AFO not significant)	Yes - Excluded year 1 of follow-up so all cancers diagnosed > 1y after COPD / emphysema
Mannino (152) *(United States)*	*2003*	1st national health & nutrition examination cohort: 113 lung cancers occurred in the 5,402 adults in the cohort.	Spirometry on entry to cohort	Prospective reports on entry to study	Moderate or severe AFO was associated with increased risk of lung cancer (HR 2.8 (1.8-4.4)	No – cancer diagnoses in early follow-up period do not seem to have been excluded
Littman (133) *(United States)*	*2004*	Analysis of 1,028 cases of lung cancer from CARET cohort study - all heavy smokers or asbestos exposed	Participant-reported physician diagnosis of chronic bronchitis or emphysema (COPD)	Prospective reports on entry to study	Those who developed lung cancer were more likely to report a history of COPD than controls (HR 1.29, 95% CI 1.09-1.53).	Yes – made no difference to results
Schabath (146) *(United States)*	*2004*	Case control study: 1,553 lung cancer cases and 1,375 healthy controls	Patient reported physician diagnosis of bronchitis or emphysema	Retrospective interviews	Emphysema (but not bronchitis) was associated with increased lung cancer risk (OR 2.87, 95% CI 2.20-3.76).	Yes - ORs consistent after exclusion of diagnoses made up to 10 years before
Wasswa-Kintu (153) *(Multiple studies)*	*2005*	Systematic review and meta-analysis of 8 studies of relationship between FEV_1 and lung cancer,	Mixed	Mixed	Risk of lung cancer increased with decreasing FEV1 in 4 studies which assessed FEV1 in quintiles.	Some but not all studies excluded initial follow-up period.
De Torres (154) *(European screening trial)*	*2007*	Analysis of 23 lung cancer cases from 1166 participants in a screening trial. Current or ex-smokers only.	Radiographic evidence of emphysema or AFO on spirometry on entry to trial	Prospective on entry to trial –	Emphysema on CT (RR, 2.51; 95% CI, 1.01 to 6.23) but not AFO (RR, 2.10; 95% CI, 0.79 to 5.58) was associated with increased risk of lung cancer	Yes –excluded cancer at baseline
Turner (142) *(United States – never smokers)*	*2007*	Analysis of lung cancer mortality among the never smokers in a previously established cohort.	Participant reported previous diagnoses of emphysema, chronic bronchitis or both.	Prospective reports on entry to study.	1,759 lung cancer deaths. Emphysema HR 1.66, chronic bronchitis 0.96 and both (COPD) 2.44, compared with people without each diagnosis.	Yes - first 1-5 years of follow-up excluded in a sensitivity analysis – no effect on results.
Purdue **(145)** *(Sweden - male construction workers)*	*2007*	Existing cohort with 834 lung cancer cases identified from cancer registry.	Spirometry on entry to cohort study to determine COPD diagnoses and severity	Prospective, reported at beginning of study	Increased rates of lung cancer for COPD (mild: RR 1.5, 95% CI 1.2 - 1.9; moderate/severe: RR 2.2, 95% CI 1.8 to 2.7) relative to normal lung function.	Yes - associations did not change with follow-up lag times of 5, 10 or 15 years after spirometry

Wilson (155) *(US screening trial)*	*2008*	99 lung cancer diagnoses identified from screening trial which only enrolled high risk patients - only smokers or ex-smokers.	Quantitative CT analysis of emphysema and lung function on entry to the trial.	Prospective prior to screening.	AFO was associated with increased risk of lung cancer (OR 2.09, 95% CI 1.33-3.27).Emphysema also increased risk of lung cancer (OR 3.56, 95% CI 2.21-5.73)	Yes – cases identified at initial screen were excluded
Yang (156) *(United States)*	*2008*	1585 lung cancer case-control pairs. Aimed to look at effects of alpha 1 antitrypsin deficiency (A1AT)	COPD diagnosis in medical notes, most also had spirometry to confirm	Medical records and retrospective interview	COPD was associated with a 3.9 fold increase in lung cancer risk. A1AT deficiency was also independently associated with increased risk of lung cancer	No
Schwartz (130) *(US women)*	*2009*	Case-control study of 562 women with lung cancer and population-based controls. Investigated risk of lung cancer associated with COPD.	Participant-reported history (obtained after lung cancer diagnosis) of emphysema, chronic bronchitis or COPD	Retrospective patient report after lung cancer diagnosis,	For combined obstructive lung disease OR 1.67 (1.15-2.41).	Yes - excluded diagnoses <1y before lung cancer and analysed according timing of diagnosis
Kiri (157) *(United Kingdom)*	*2010*	Used general practice research database to determine trends in lung cancer in patients with COPD compared with general population.	General practice records of COPD diagnoses	No smoking data.	Annual incidence rates of lung cancer were at least 4-fold higher in people with prior COPD compared with the general population	No
Maldonado(158) *(US screening trial)*	*2010*	Case control study of 64 screen detected cases of lung cancer 6 controls matched per case on age, sex and smoking. All smokers or ex-smokers.	Quantitative CT analysis of emphysema and lung function on entry to the trial.	Prospective prior to screening.	AFO was associated with an increase in risk of lung cancer (OR 1.15, 95% CI 1.00-1.32), but radiographic emphysema was not.	No – incident cases of lung cancer detected at first CT screen were included
Brenner (11) *(Multiple studies)*	*2011*	Systematic review and meta-analysis: 39 studies assessed effects of COPD, chronic bronchitis and/or emphysema on lung cancer risk.	Several different methods including lung function, participant report and emphysema on CT	Mixed	Relative risk of lung cancer with a previous history of COPD, chronic bronchitis or emphysema was 1.8 (95% CI 1.60-2.11).	Fewer than half of the studies included considered diagnostic latency
Leung (143) *(Hong Kong - women)*	*2012*	1,297 lung cancer deaths in cohort study of elderly people in a health maintenance programme.	Participant reports of physician diagnosed COPD.	Prospective, reported at beginning of study	In the overall analysis, obstructive lung disease was associated with lung cancer mortality (HR 1.86, $p< 0.001$) but not in never smokers	Yes – Potentially prevalent cases and deaths in initial 3 years excluded.

COPD Chronic Obstructive Pulmonary Disease; AFO Airflow obstruction; FEV_1 Forced expiratory volume in 1 second; US United States; CT Computerised tomography

4.4.6 Pneumonia and asthma

Prior diagnoses of pneumonia and asthma were assessed in addition to COPD to assess the specificity of these findings in relation to COPD and lung cancer.

Pneumonia

The association between pneumonia and lung cancer was even stronger than that between COPD and lung cancer for diagnoses made within 6 months. Likely explanations for this are ascertainment bias, as described above for COPD, reverse causation (lung cancer may lead to airway obstruction and distal infection and may also weaken immune response) and initial misdiagnosis (symptoms, clinical signs and findings on chest radiograph are often similar).

For diagnoses made over 10 years before lung cancer the association with pneumonia is less marked than with COPD. This is probably because the strength of association between smoking and pneumonia is not as strong as with COPD, and hence there is less residual confounding, and also because pneumonia is usually an acute illness and unless there is co-existing chronic disease does not result in on-going follow-up. Brenner *et al* reported very similar results: The combined relative risk of lung cancer in people who had had pneumonia was 1.43 times that of those who hadn't (95% CI 1.22-1.68). Twenty-two studies contributed to this analysis and whilst all adjusted for smoking many did not account for timing of diagnosis. The combined estimate from the 8 studies or subgroups of never smokers was similar to the overall figure.

Asthma

After removing cases that also had records of COPD, adjusting for smoking, and accounting for ascertainment bias, there was no evidence of an association between diagnoses of asthma and lung cancer in this study. This is consistent with some of the previous literature.(133, 159) This suggests that the link may

be limited to smoking related diseases (COPD and pneumonia) and further supports the hypothesis that the remaining association could be due to residual confounding by smoking. Based on these results it seems less likely that airway inflammation explains the increase in risk of lung cancer in people with COPD since airway inflammation is a prominent feature in asthma.

The finding of an association between proximal diagnoses of asthma and lung cancer suggests that ascertainment bias is not limited to the smoking related diseases and should be taken into account when considering whether any apparent association is causal. This may be part of the reason that some previous studies have reported an independent association between asthma and lung cancer.(132)

4.4.7 Clinical relevance

Despite the huge element of confounding in many of the initial studies which suggested COPD was an independent risk factor for lung cancer, biological studies are underway looking for evidence of a molecular link between the two diseases which could cause people with COPD to be at even higher risk of lung cancer than those who smoke exactly the same amount. (160) These data suggest, however, that the association between COPD or pneumonia and lung cancer is largely due to confounding by smoking and ascertainment bias.

There is an extremely strong unadjusted relationship between both COPD and pneumonia and lung cancer in the 6 months immediately prior to lung cancer diagnosis. This is useful in a clinical context with potential implications for patient selection in screening trials: this could facilitate recruitment of heavy smokers who are unwilling to admit their smoking status or patients for whom smoking data are unavailable or inaccurate, yet are at high risk of developing lung cancer. These results also support the current National Institute for Health and Clinical Excellence recommendation that all patients should have a chest radiograph looking for evidence of lung cancer at the time of COPD diagnosis

(136), and to reduce the disease burden, resources should probably remain focused on smoking cessation, novel therapies and early detection of lung cancer.

4.5 Chapter summary

In this chapter I have presented evidence of the strong association between COPD and lung cancer, but argued that this can probably all be explained by smoking and ascertainment bias. It is important that the scientific community consider this explanation for the apparent independent association between COPD and lung cancer so that resources in lung cancer can be appropriately allocated to evidence based interventions.

This work was published in the Journal of Thoracic Oncology in January 2013.(161) The interpretation of the data presented in this chapter and in the published work is that of myself and my co-authors (including my PhD supervisors), and not everyone will agree, as demonstrated by the correspondence to the journal following publication of our article.(162)

The studies in this and the previous chapter used primary care data to investigate factors which affect lung cancer prior to diagnosis. Diagnosis for patients in England occurs in secondary care and therefore the next four chapters use the linked HES-NLCA-ONS secondary care data to investigate factors which influence treatment and outcomes for people with lung cancer.

CHAPTER 5: VALIDATION OF RECORDS OF SURGICAL PROCEDURES

This chapter describes a validation study in which records of potentially curative thoracic surgical procedures in the HES and NLCA databases are compared, with the aim of determining the most appropriate definition of surgery for future studies. The chapter concludes with a description of some of the features of patients who had surgery for lung cancer and their survival, and a comparison with other published data.

5.1 Introduction

5.1.1 Background

The National Lung Cancer Audit (NLCA) has collected data on people with lung cancer in England since 2004 and contains information on demographics, diagnoses and treatments. These data have been used to provide evidence of inequalities in access to treatments, (60, 83) however the accuracy of treatment records in the NLCA, in particular whether a missing treatment date truly represents no treatment, has not been assessed.

The NLCA data were linked with inpatient admission data from Hospital Episodes Statistics (HES) for the initial purpose of assessing co-morbidity (as described in Section 2.3.3). Since the HES database also includes a code for every surgical procedure which takes place during an inpatient episode these can be compared with NLCA treatment records and patient features and outcomes can be analysed in order to assess the validity of treatment records in each database.

5.1.2 Rationale for this study

The case ascertainment and completeness of individual data fields in NLCA has improved substantially in the last few years and the database is now a valuable resource for epidemiological studies in lung cancer. It is important to determine the most accurate means of identifying exposures so that studies are consistent in their methods and are not affected by errors in data entry or recording bias.

In Chapter 6 I will describe a study in which factors associated with early mortality after surgery for lung cancer were explored and a new predictive score was developed. It was important to be confident that all cases included in that study underwent potentially curative surgery for lung cancer, hence the work presented in this chapter.

5.1.3 Aims of this chapter

The aim of the study described in this chapter was to assess the validity of records of potentially curative surgical procedures for NSCLC in the HES and NLCA datasets in order to agree a definition of surgery for future studies. This was done by:

1. Identifying patients with NSCLC who had a surgical procedure recorded in the NLCA, or in the linked HES data, or in both;
2. Examining and comparing the features of these patients (including survival) according to where surgery was recorded;
3. Describing the features of patients who had potentially curative surgery for NSCLC using the new definition and comparing these to the published literature.

5.2 Records of surgery in HES and the NLCA

5.2.1 Methods

Study population

The August 2011 NLCA-HES data extract was used (see section 2.2.4) but patients in the NLCA who were first seen after 31st March 2010 were excluded because HES data were not available after this date. Patients first seen prior to 2004 were also excluded. Cases of NSCLC were identified by excluding records where there was a clinical or histological diagnosis of mesothelioma or where histology was recorded as SCLC or carcinoid.

Patients with evidence of advanced disease (stage 3b or 4) were excluded to ensure as far as possible that the cases analysed had undergone surgery with curative intent.

Survival was assessed from the date of diagnosis or the start-date where this was missing (as described in section 2.2.5). Patients for whom a start-date could not be calculated and those with a date of death on or before their start-date were therefore also excluded.

Covariates

Stage, histology, lung function, performance status and socio-economic status (Townsend quintile) were defined as described in section 2.2.5.

Age refers to the NLCA variable age at time of diagnosis. Age and per cent of predicted FEV1 were studied as continuous variables, performance status was grouped as 0-1, 2, 3-4 or missing and stage as 1a - 1b, 2a - 2b, 3a or missing.

HES records of surgery

Surgical procedures are recorded in HES using OPCS-4 codes. Each of these codes is associated with an inpatient episode and the specific date of procedure is also recorded.

A list of OPCS-4 codes which, in a patient with a recent diagnosis of lung cancer, would be likely to represent an attempt at curative surgery was generated: two respiratory physicians (HP & Prof David Baldwin) independently rated a list of all thoracic surgical procedures before agreeing on the final code list which is given in Appendix E. This code list was merged with the HES database to identify every record which contained a relevant procedure. Procedures were categorised as pneumonectomy (highest priority), bi-lobectomy, lobectomy, segmentectomy / sleeve / wedge or other (lowest priority).

It was possible at this stage for an individual case to have more than one line of data, indicating that the patient underwent a procedure which would be consistent with lung cancer surgery on more than one occasion (possibly during more than one hospital episode). This situation could arise if a patient developed complications and required a further surgical procedure, or if they had undergone thoracic surgery prior to the diagnosis of lung cancer for a different indication. In an attempt to ensure that HES procedures were only included if they were for the current NLCA diagnosis of lung cancer, that they were performed with curative intent, and to exclude obvious errors in data entry, procedures performed more than 3 months before or more than 6 months after the NLCA start-date were not included.

If there were still multiple procedures for one patient the most complicated procedure type (highest priority as defined above) was used, followed by the procedure with the latest date.

NLCA records of surgery

Within the NLCA dataset there are four fields relating to surgery:

- the date that the decision to operate was made,
- the type of surgical procedure,
- the actual date of the procedure, and

- the trust where the procedure was carried out.

Date of decision to operate is not as useful as the other fields in this context as a decision to operate does not necessarily mean an operation took place; this analysis therefore focused on the remaining three variables. One might expect all four data fields to be complete for all patients who had a surgical procedure however this is not always the case because, as with any large dataset, data may be missing or contain errors. In addition, if the treatment plan changed for any reason the fields should be updated but it is possible that busy clinicians or administrators do not have time to keep track of changes and / or update the database in some cases.

The 14 different codes used by the NLCA to define type of procedure are given in Appendix E. For this study, extra-pleural pneumonectomy, de-bulking pneumonectomy, and pleurodesis were excluded as these do not represent potentially curative surgery for NSCLC. In records which had a code for one of these procedures the three fields of interest (type of surgical procedure, date of procedure, and trust where the procedure was carried out) were re-coded to missing.

Dates

HES records were only available up to March 31st 2010, and any procedures recorded in HES which were performed more than 3 months before or 6 months after the lung cancer diagnosis date were excluded (see above). Procedures recorded in the NLCA which were dated outside these time periods were also excluded to allow fair comparison (the procedure date, type and trust were re-coded to missing).

The entire record for any patient with a procedure (in either dataset) dated before January 1st 2004 or after March 31st 2010 was dropped from the analysis for this comparison study (re-coding these as missing as described above would

mean they were included in group 5 (see below) when in fact they may have had potentially curative surgery before or after the study period.

Definition of outcome

The ONS date of death (described in section 2.4) was used for survival analyses, patients who were still alive at the last ONS cross-check (8th August 2011) were censored on this date.

Statistical methods

Following the exclusions and re-coding described above, patients who had non-missing values in the fields relating to surgery in each dataset were identified and grouped as shown in Table 5-1. A Venn diagram was constructed to show the overlap between these groups.

Table 5-1: Criteria for groups which indicate where records of surgery were identified

Group	Criteria
1=Both	Date of surgery in HES AND Date of surgery in NLCA
1a	Date of surgery in HES AND Procedure type in NLCA NO date of surgery in NLCA
1b	Date of surgery in HES AND Trust of surgery in NLCA NO procedure type in NLCA NO date of surgery in NLCA
2=HES only	Date of surgery in HES NO reference to surgery in NLCA
3=NLCA only (date)	Date of surgery in NLCA NO reference to surgery in HES
4=NLCA only (procedure type or trust)	Procedure type and/ or trust of surgery in NLCA NO date of surgery in NLCA NO reference to surgery in HES
5=Neither	NO reference to surgery in either database

The features (age, stage, lung function, performance status, year of surgery, survival after diagnosis and 30- and 90- day post-operative mortality) of patients

in each of the groups were examined to identify any patterns. A Kaplan Meier survival curve was constructed to compare the overall survival of people in each of the groups in table 5-1.

5.2.2 Results

There were 133,689 patients in the NLCA database first seen between 1st January 2004 and 31st March 2010. After excluding 6,875 cases of mesothelioma and 13,553 histologically confirmed cases of small cell lung cancer or carcinoid, the remaining 113,261 cases were classified as NSCLC.

From the NSCLC population, 46,013 cases with advanced stage and 5,328 for whom a start date could not be calculated or was after the recorded date of death were excluded, as well as 945 with a procedure date outside the period of study. This left 60,975 records for the analysis.

HES records of surgery

There were 11,040 records which contained at least one of the OPCS-4 codes listed in Appendix E, dated between 1st 2004 and March 31st 2010, and less than 3 months before / 6 months after the NLCA start date. The distribution of procedure types is shown in table 5-2.

Table 5-2: Distribution of procedure types as recorded in HES

Procedure category	Frequency	Percentage
Pneumonectomy	1105	*10.0*
Bi-lobectomy	453	*4.0*
Lobectomy	7095	*64.3*
Segmentectomy or wedge	1661	*15.0*
Other	738	*6.7*

NLCA records of surgery

There were 9,373 records with a procedure date in the NLCA within 3 months before or 6 months after the start date, and between 1st 2004 and March 31st 2010. A further 75 records had a procedure type but no date, and 1,862 a trust of surgery but no procedure date or type.

Procedure types for those with a procedure date are shown in Table 5-3. In 16% (1,472) of these cases there was a procedure date recorded but the procedure

type field was blank. This is likely to be the reason that all other categories represent a slightly lower proportion of the total number of procedures when compared with the distribution in HES (Table 5-2).

Table 5-3: Distribution of procedure types recorded in NLCA

Procedure category	Frequency	Percentage
Pneumonectomy	756	*8.1*
Bi-lobectomy	215	*2.2*
Lobectomy	5,766	*61.5*
Wedge resection	854	*9.1*
Lung & chest wall resection	63	*0.7*
Multiple wedges	46	*0.5*
Sleeve resection	64	*0.7*
Segmental resection	130	*1.4*
Carinal resection	7	*0.1*
Missing	1,472	*15.7*

Comparison of databases

Procedure dates were recorded in both databases for 8,965 patients. Figure 5-1 shows the number of patients who had records of surgery in each group as defined in Table 5-1.

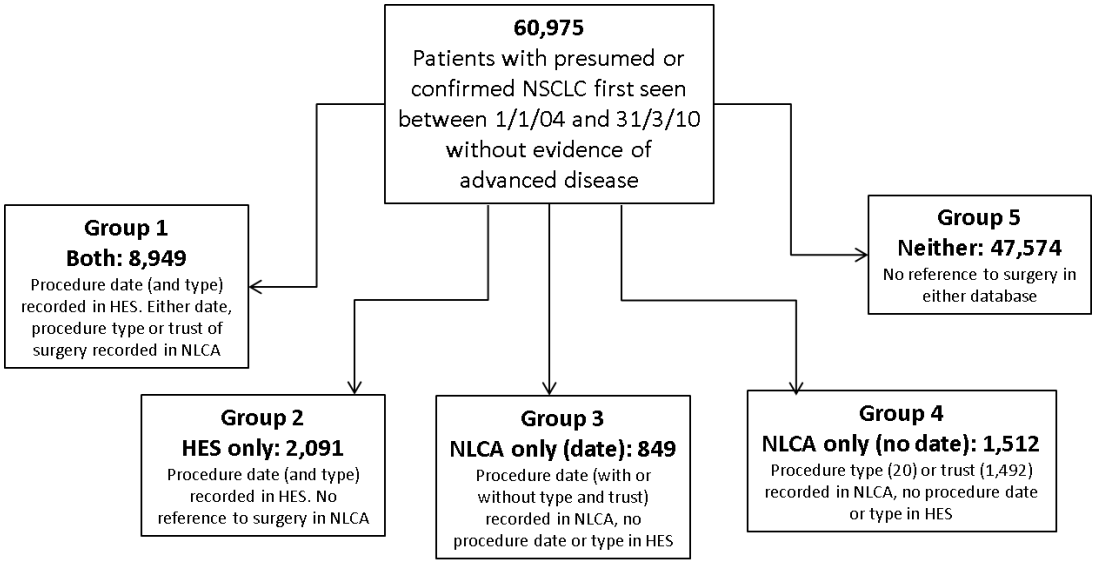
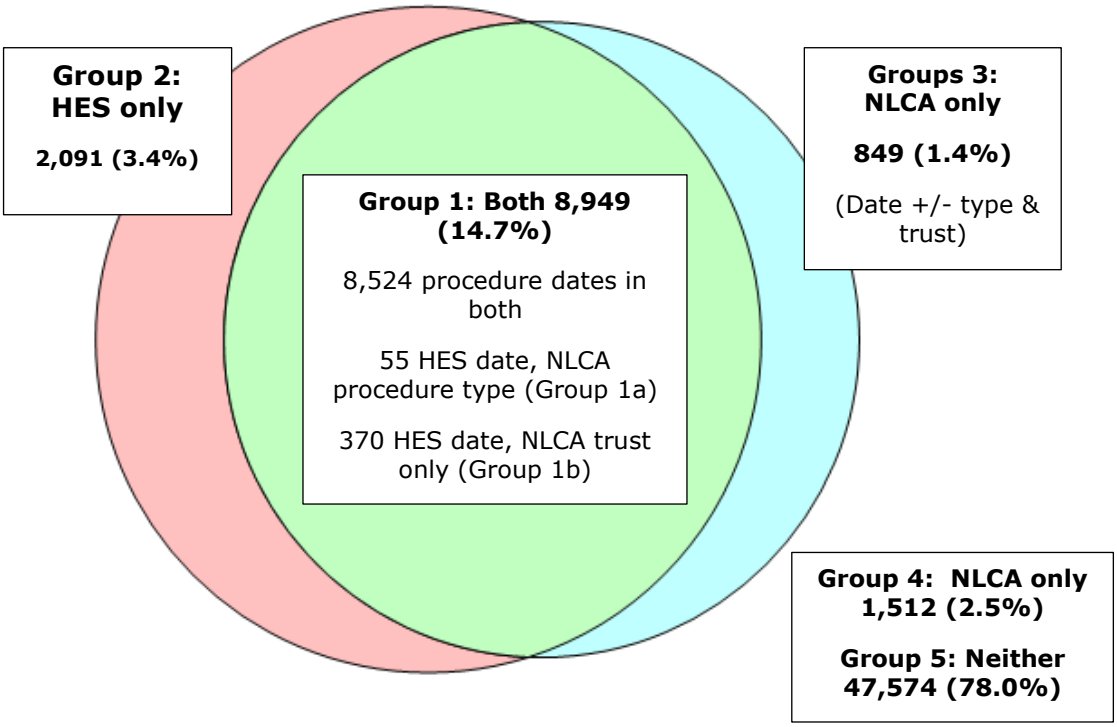


Figure 5-1: Records of procedures in HES and NLCA databases

The overlap between the records of surgical procedures in the databases is shown in the Venn diagram (Figure 5-2).



Percentages indicate proportion of overall population, N=60,975

Figure 5-2: Venn diagram depicting the overlap between records of surgical procedures in HES and the NLCA

Patient features

Features of patients in each of the groups are shown in table 5-4. Due to small numbers in groups 1a (n=55) and 1b (n=370) these were included with group 1.

The mean age of patients in groups 1-3 (who had a procedure date recorded in at least one database) was similar; patients in groups 4 and 5 were a few years older on average (Table 5-4). Patients with procedure dates recorded in both databases had better lung function than those with procedures recorded in one database only; lung function was considerably worse in those without a procedure date in either database. Performance status was also better in those who had a procedure recorded in both databases than any other group. A higher proportion of patients in groups 2 and 3 had a performance status of >1 (15% and 14% respectively) compared with group 1 (8%).

Patients with a date of surgery in the NLCA had fewer missing data on performance status, stage, and lung function than the other groups. It should be noted that these are all NLCA data fields.

Overall survival

Overall survival was longest in people who had a record of surgery in both databases (median 60 months / 5 years), followed by those with a record in HES (42 months) and those with a procedure date in the NLCA (20 months). The Kaplan Meier survival curve for all 5 groups is shown in Figure 5-3.

Overall, the features of patients in group 4 (procedure type or trust recorded in NLCA only) were similar to those of the people in group 5 (no record of surgery in either database). Survival was poor in both of these groups (median 6.7 and 9.5 months from diagnosis respectively).

Perioperative mortality

Early post-operative mortality was higher in the NLCA only (group 3: 5% 30-day mortality) and HES only (group 2: 5% 30-day mortality) compared with the

group where procedures were recorded in both databases (group 1: 3% 30-day mortality). The proportion of people who appeared to have died within 90-days of a procedure date was extremely high for those with a procedure recorded in the NLCA only (16% compared with 5% for group 1 and 9% for group 2).

Table 5-4: Characteristics of patients according to where surgical procedures were recorded

N=60,975	Record of surgical procedure				
	Group 1 Both	**Group 2 HES only**	**Group 3 NLCA only (date)**	**Group 4 NLCA only (no date)**	**Group 5 Neither**
	n=8,949	n=2,091	n=849	n=1,512	n=47,574
Mean age (years)	67.4	66.8	67.7	70.3	72.6
Mean % predicted FEV1	77.2	73.8	74.5	63.7	68.3
Missing FEV1 (% of total)	55.1	79.6	67.7	81.7	83.5
Stage (% of non-missing) 1a or 1b	66.3	57.0	60.3	44.9	35.7
2a or 2b	22.3	21.9	20.8	22.6	19.5
3a	11.4	21.1	18.9	32.4	44.8
Missing stage (% of total)	15.7	63.7	49.0	77.8	72.5
Performance status (% of non-missing) 0-1	92.4	84.6	85.7	71.0	47.6
2	6.3	11.2	9.2	17.4	24.3
3-4	1.3	4.2	5.1	11.6	28.1
Missing performance status (% of total)	29.0	61.1	37.3	62.8	49.4
Median survival (months)*	60.1	41.6	19.6	6.7	9.5
****Died within 30-days of surgery (%)**	2.7	4.6	5.4	-	-
****Died within 90-days of surgery (%)**	5.4	9.1	15.6	-	-

Survival is calculated from start date not date of procedure;FEV1 Forced expiratory Volume in 1 second; **Date of procedure as recorded in HES unless NLCA only

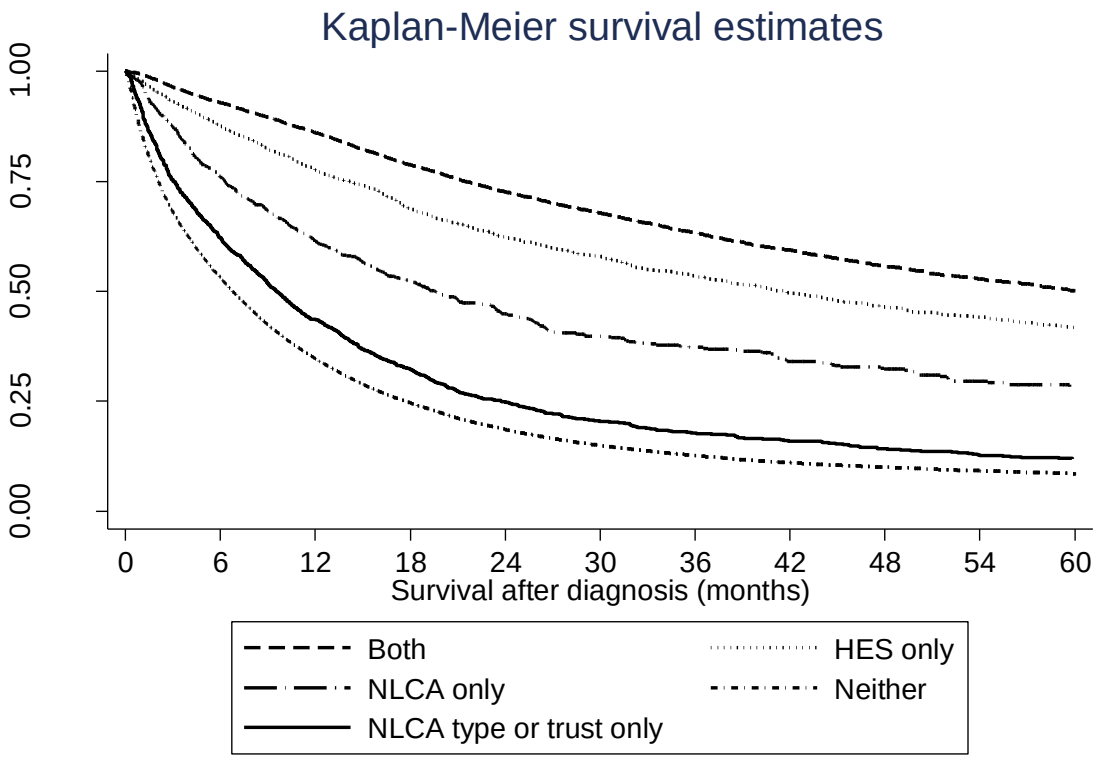


Figure 5-3: Kaplan Meier curve to show survival according to where surgery was recorded

Procedure records by year

Between 2004 and 2009 the proportion of patients with any record of surgery that had this recorded in both databases increased from 59% to 86% (Table 5-5). Over the same time period the proportion with a record in HES only decreased (33% to 9%) but the proportion with a record in the NLCA varied between 5% (2009) and 9% (2007) without a clear pattern.

The proportion with surgery recorded in both databases (group 1) decreased between 2009 and 2010 and the proportion in the NLCA only increased to 13%. (Table 5-5). This is likely to be because HES data entry stopped on 31st March 2010 and therefore may not be complete for procedures which took place shortly before this date, whereas NLCA data could be entered retrospectively after this date.

Table 5-5: Records of procedures in the NLCA and HES by year

Year (number of patients with any record of surgery)	**% recorded in both datasets** (Group 1)	**% recorded in HES only** (Group 2)	**% recorded in NLCA only** (Group 3)
2004 (n=459)	59.0	32.5	8.5
2005 (n=1301)	61.0	32.7	6.2
2006 (n=1721)	71.2	21.2	7.6
2007 (n=2076)	69.5	21.5	9.0
2008 (n=2410)	78.6	15.2	6.3
2009 (n=3080)	86.3	8.8	4.9
2010 (n=842)	78.7	8.2	13.1

5.2.3 Interpretation

NLCA only

The methods of data entry in HES and the NLCA differ, as described in Chapter 2. With this knowledge it was hypothesised that patients who had a record of a potentially curative surgical procedure in the NLCA but not in HES may not actually have had surgery. This may have been because their performance status deteriorated, the patient changed their mind, or new information became available showing the tumour to be technically inoperable after the initial treatment plan was made. In these situations it is possible that the NLCA record was not updated.

This hypothesis is supported by the observation that patients who only had a record of surgery in the NLCA and not in HES had considerably shorter survival (median 543 days) compared with those who had surgery recorded in both datasets (median 1839 days).

The analysis of 30-, 60- and 90-day post-operative mortality in these groups shows patients in the 'NLCA only' group to be more likely to die within all of these time periods but the greatest difference is evident within 90 days where 16.7% of patients with surgery recorded in the NLCA only died compared with

5.5% of those with surgery recorded in both datasets. This may indicate that some of these patients deteriorated rapidly and were too unwell to have surgery.

HES only

In contrast to the NLCA, HES codes were entered at the end of each hospital episode and were required for NHS trusts to bill the Primary Care Trust for the services they provide. For this reason it was felt that a procedure code in HES was likely to indicate that the procedure did actually take place. Figure 5-3 shows that 2,371 patients had a procedure recorded in HES but not in the NLCA. The 30- (and 90-) day mortality is higher in this group of patients, and the median survival is lower, than in the group for whom surgery is recorded in both datasets, however to a lesser degree than the NLCA only group. This may be because these patients were not expected to have surgery initially as they were considered borderline in terms of fitness and stage of tumour, and the NLCA database was not updated with this new information. This hypothesis is supported by the observation that the HES only group of patients were more likely to be performance status 2 or worse and stage 3a than those with a surgical procedure recorded in both databases (Table 5-4).

NLCA trust or procedure type only

The features of patients who only had a procedure type or trust of surgery recorded in the NLCA (no date in the NLCA and no reference to surgery in HES) were very similar to those of patients with no reference to surgery in either database. It is reasonable to conclude therefore that these patients did not have surgery.

5.2.4 Conclusion

Given the above observations, the definition of a surgical procedure was restricted to those recorded in HES (list of OPCS-4 procedure codes in Appendix E). Using this definition we can be most confident that these patients met the study criteria, however there may be other patients who also had potentially curative thoracic surgery for NSCLC in the study period, not captured by this definition; we expect this number to be low given the high level of case ascertainment of the NLCA and the incentives for recording procedures in HES.

5.3 Description of patients who had surgery and comparison with published data

5.3.1 Methods

Study population

The lung cancer population for this study was similar to that described in section 5.2. Patients with NSCLC first seen between January 1st 2004 and March 31st 2010 were identified using the NLCA database. In addition to those with advanced disease (stage 3b or 4), patients with an ICD-10 code in HES for metastatic disease which occurred prior to the procedure date were excluded. People with an age at diagnosis <30 were also excluded.

Definition of exposure

After the detailed analysis of records of surgical procedures in both databases, (described in section 5.2), the decision was made to define surgery for this analysis as a surgical procedure code which was:

1. recorded in HES;
2. in the list of OPCS-4 procedure codes consistent with potentially curative surgery for lung cancer (Appendix E);
3. dated within 3 months before and less than 6 months after the NLCA start date.

Patients were included in this analysis if they had a record of surgery by the above definition which occurred between 1st January 2004 and 31st March 2010.

Procedure type and date

The date and type of procedure were obtained from the HES database. If an individual patient had more than one appropriate procedure coded in HES (with either the same or different dates) the code for the highest priority procedure type and then the most recent date was used.

Where surgery was also recorded in the NLCA, the difference between the dates was calculated: if this was more than 10 days the patient was excluded from the analysis.

Demographics, co-morbidity and tumour features

All demographic data fields, histological subtypes, lung function, performance status and stage were obtained from the NLCA. Further information on these variables can be found in section 2.2.5.

Statistical methods

Demographic variables, patient fitness (performance status and lung function), tumour stage and procedure type were described using averages, histograms, proportions and simple tabulations as appropriate.

Kaplan Meier curves were plotted for the first year after the date of operation by age, stage, performance status and procedure type, and for the 5 years after surgery for the population overall and by stage. Date of death was obtained from ONS records and any record without a date of death was assumed to be alive at the last ONS cross-check and censored on this date (8th August 2011). These graphs were used to assist in determining the most appropriate time points at which to report risk factors for early post-operative death (Chapter 6).

Demographics, overall survival, and survival by stage were compared with previously published international data.

5.3.2 Results

These results describe 10,991 patients who underwent potentially curative surgery for NSCLC between 1st January 2004 and 31st March 2010. The process of determining the study population is shown in figure 5-4.

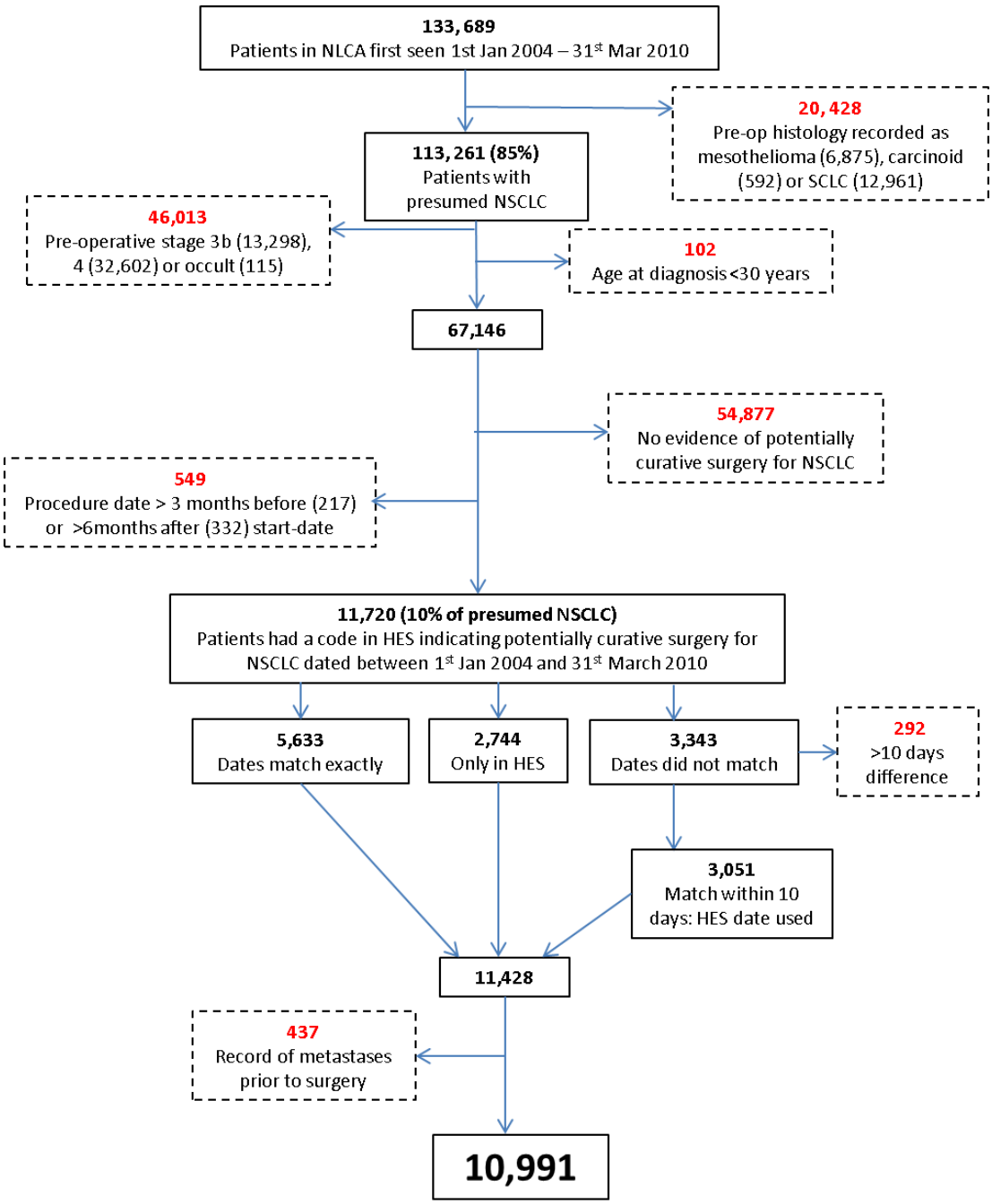


Figure 5-4: Process diagram for producing study population

Demographics

The majority (56%) of the patients were male and the mean age was 67 years, standard deviation 9.3 years.

Fitness

The most common performance status was 0 (31%), although performance status was not recorded in 38% of cases. Only 108 patients had a performance status of 3-4. The mean percentage of predicted FEV_1 was 77%, SD 20.8%, (figure 5.6) but FEV_1 was not recorded for 61% of patients.

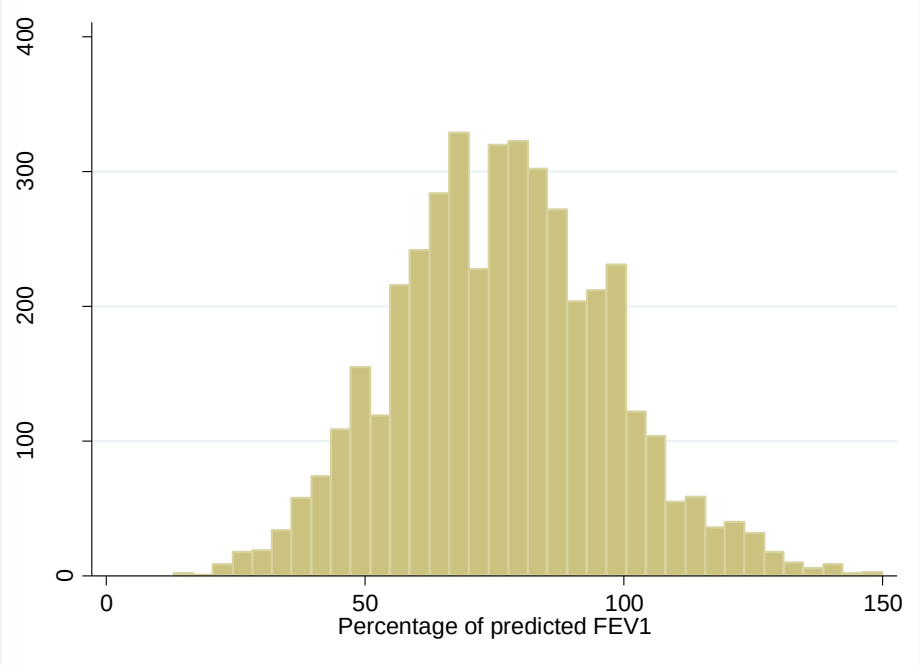


Figure 5-5: Distribution of lung function in patients who underwent surgery

Tumour features

Stage 1b was the most common stage (28%), although 26% did not have a pre- or post-operative stage recorded. The most common histological types were adenocarcinoma (31%) and squamous cell (28%); 21% did not have a record of pre or post-operative histology.

Procedure type

Most patients had a lobectomy (64%) and only 11% had a pneumonectomy.

Early post-operative mortality

Three per cent (334) of patients died within 30 days of their procedure.

The Kaplan Meier survival curves for the first 6 months post-operatively by age, stage, performance status and procedure type (figures 5-6 to 5-9) showed the rate of death to be slightly higher up to 90 days after surgery, but then more or less constant over the following 9 months (note altered scale on y-axes to show subtle differences in rate of deaths). This will be discussed further in the following chapter, section 6.3.2.

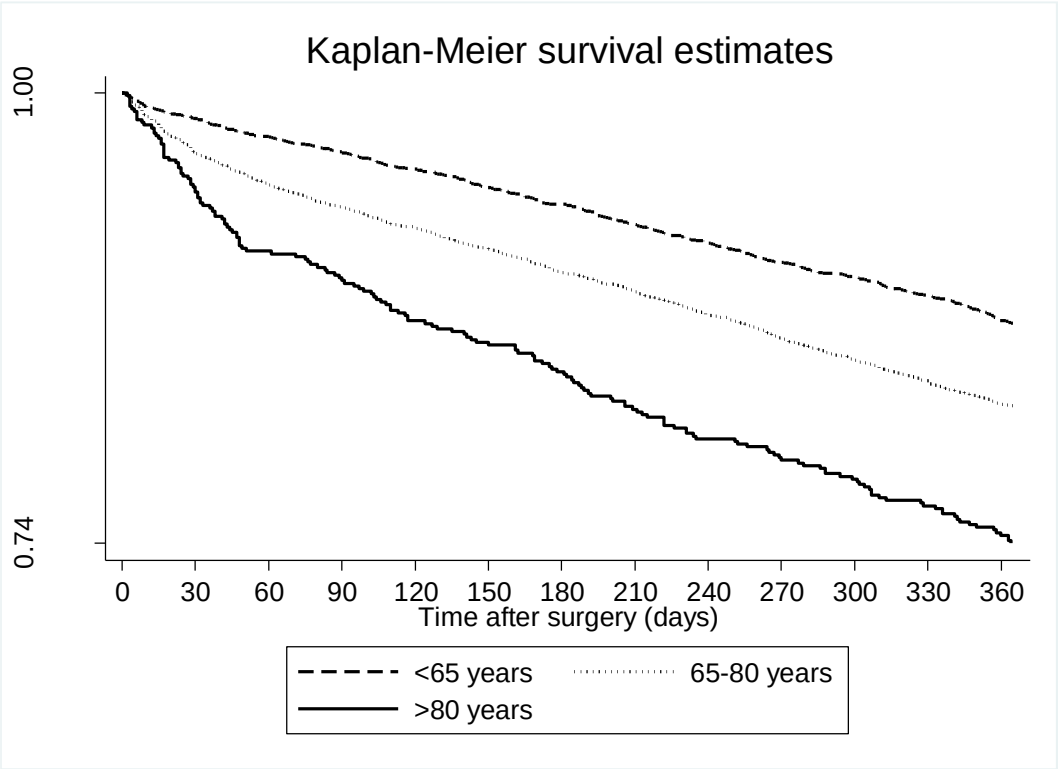


Figure 5-6: Kaplan Meier survival curve by age for first year after surgery

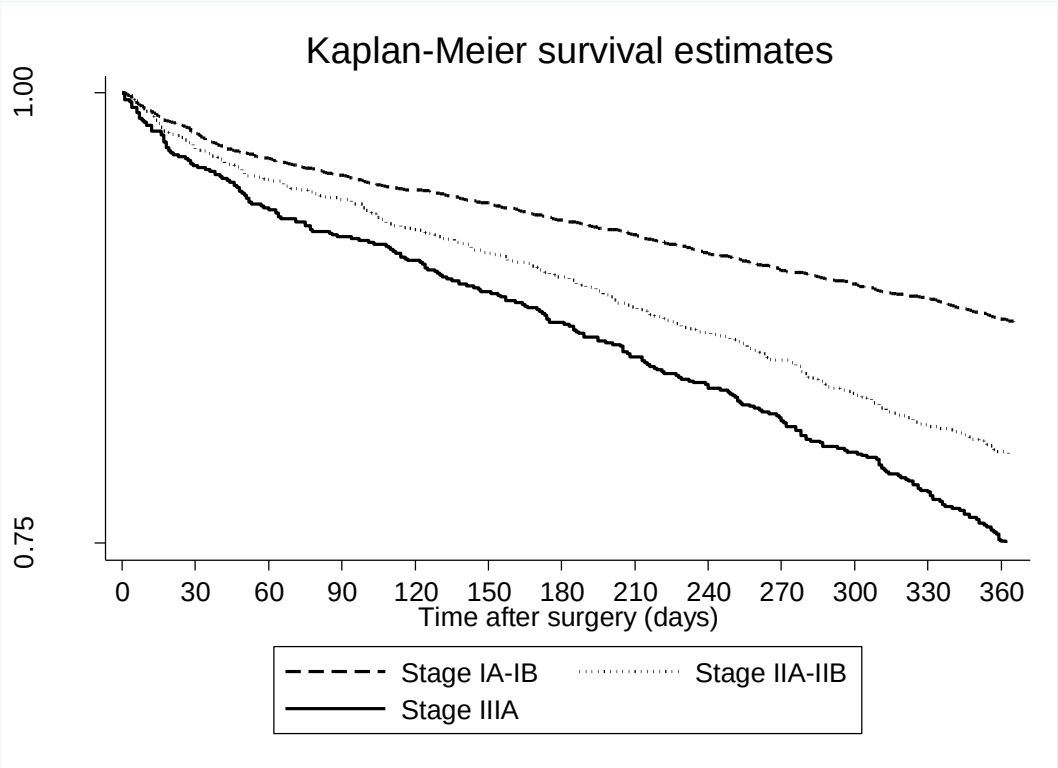


Figure 5-7: Kaplan Meier survival curve by stage for first year after surgery

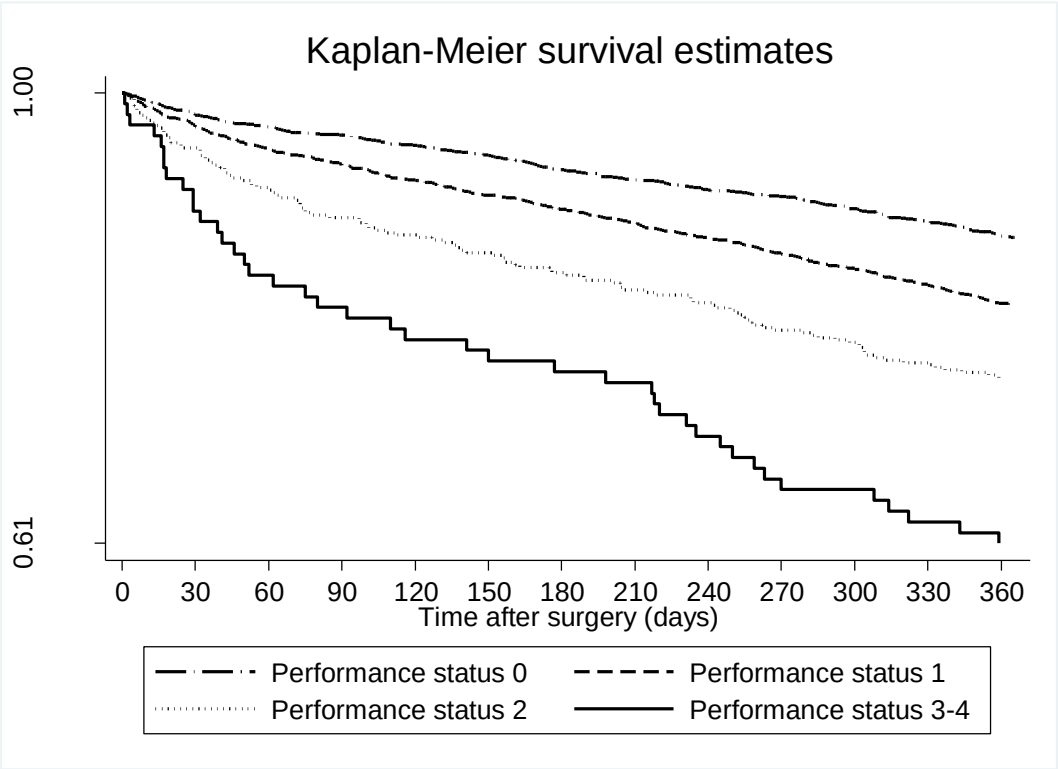


Figure 5-8: Kaplan Meier survival curve by performance status for first year after surgery

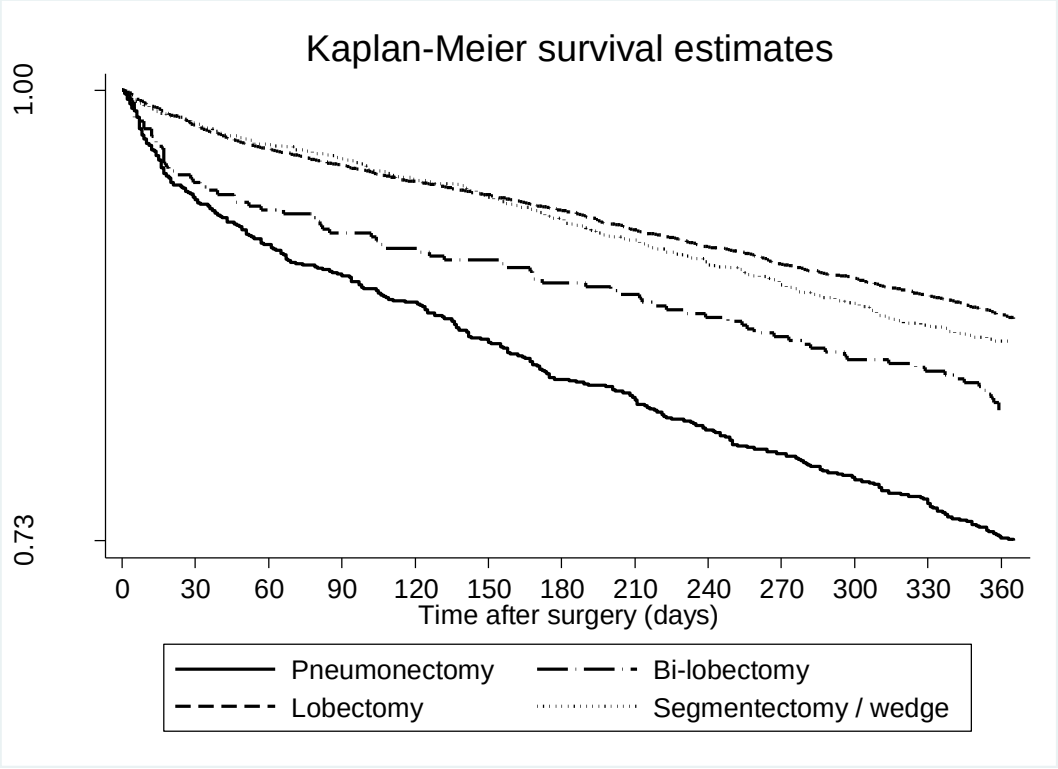


Figure 5-9: Kaplan Meier survival curve by procedure for first year after surgery

Long term survival

Overall 1-year and 5-year survival were 83% and 48% (figure 5-10).

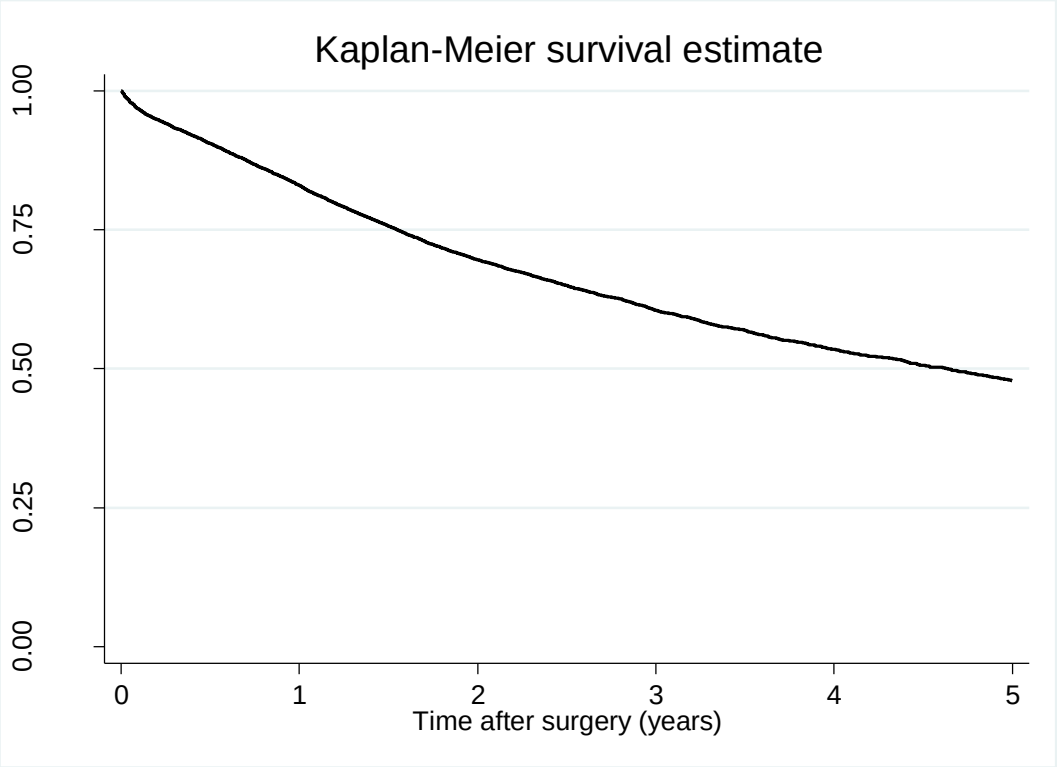


Figure 5-10: Survival after surgery for population overall

Five-year survival by stage is shown in table 5-6 and figure 5-11 (stage 2A was excluded from the figure because the survival curve overlapped almost entirely with stage 1B). The stage recorded was pre-operative unless this was missing from the NLCA database in which case post-operative stage was used.

Table 5-6: 1 and 5 year survival after surgery by stage

Stage	Proportion alive	
	1 year after surgery	5 years after surgery
IA	91%	60%
IB	85%	51%
IIA	86%	50%
IIB	79%	43%
IIIA	75%	34%
Missing	79%	43%

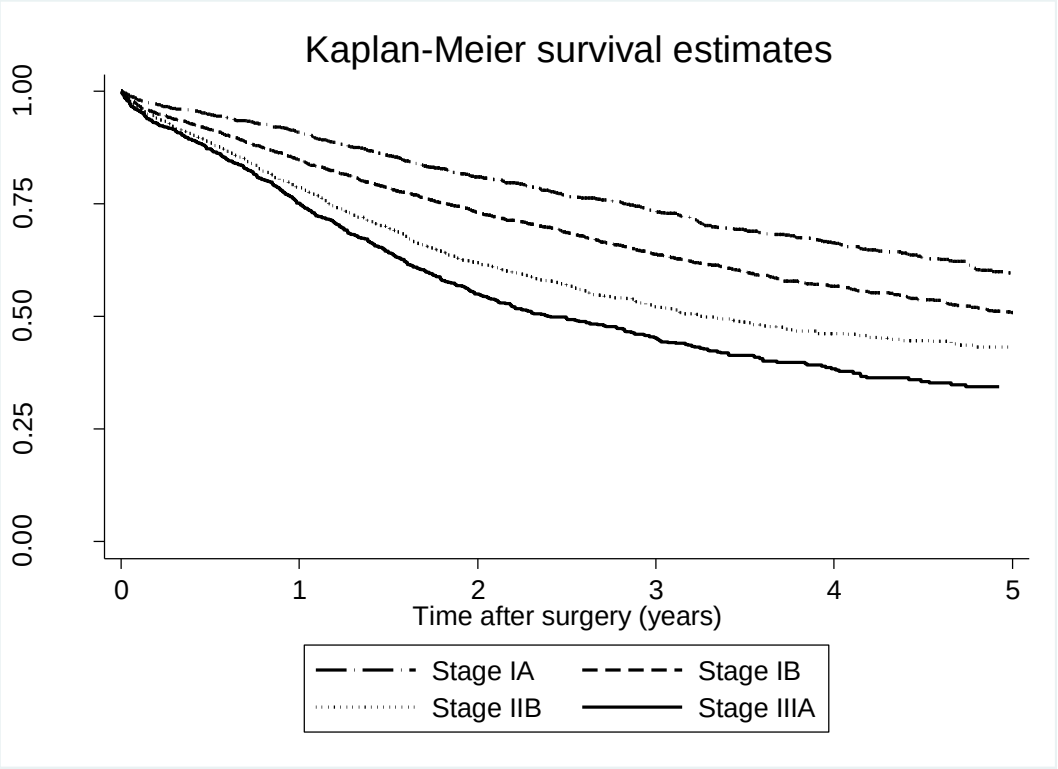


Figure 5-11: Survival after surgery by stage

5.3.3 Comparison with previously published data

The largest contemporary published series of operated lung NSCLC comes from the International Association for the Study of Lung Cancer (IASLC) staging studies.(19) This work resulted in the re-classification of the lung cancer TNM staging system from UICC version 6 (as used in this study) to version 7 based on 5 year survival after treatment.(16) Survival for stage Ia was better in the IASLC study than in the current study (73% vs. 60%), but worse for stage IIIa, with similar figures for other stages (Table 5-7). These differences are likely to be due to differences in the definition of staging: Pre-operative (clinical stage) was predominantly used in the current study whereas pathological (post-operative) stage was used the IASLC project. Pre-operative stage is likely to under-stage a proportion of patients. There may also be differences in populations studied: Australasia, North America and other countries in Europe contributed cases to the IASLC project and we know that survival from lung cancer in these countries is different to the UK, possibly due to differences in selection of patients and surgical techniques.(53, 72)

A further publication described some of the features of the surgically resected patients with NSCLC from the IASLC database revealing a slightly higher proportion of men than in the current study. This is consistent with other reports (Table 5-7) and is likely to reflect the years in which the studies were conducted (lung cancer incidence in males has been falling since the 1980s as described in Chapter 1).(163)

A few other published series of operated lung cancer which included 1- or 5-year survival figures are summarised in table 5-7. These studies were based on analyses of consecutive patients who underwent lung cancer resection at single institutions and are therefore much smaller than the current study. Nonetheless, the survival figures are reasonably similar.

Table 5-7: Patient features and survival from published series of operated NSCLC

	Current study	**Roth (164)**	**Van der Pijl (165)**	**Brim (166)**	**IASLC (19, 163)**	**Al Kattan (167)**
Country	England	Norway	Holland	Holland	Multi-national	England
Dates	2004-2010	1993-2006	2002-06	1989-2001	1990-2000	1987-88
Cases	10,991	148	126	766	9,137	200
Males	56%	68%	69%	78%	74%	71%
Age (mean)	67	67	63	65	-	64
Pneumonectomy	10%	-	25%	27%	-	29%
1y survival - overall	**83%**	**-**	**86%**	**-**	**-**	**-**
Ia	91%	89%	-	-	-	-
Ib	85%	78%	-	-	-	-
IIa	86%	-	-	-	-	-
IIb	79%	-	-	-	-	-
IIIa	75%	-	-	-	-	-
5y survival - overall	**48%**	**42%**	**-**	**40%**	**-**	**-**
Ia	*60%*	-	-	-	73%	60%
Ib	*51%*	-	-	-	54%	
IIa	*50%*	-	-	-	48%	30%
IIb	*43%*	-	-	-	38%	
IIIa	*34%*	-	-	-	25%	16%

IASLC: International Association for the Study of Lung Cancer. Reference numbers given in brackets after first author names.

5.4 Chapter summary

In this chapter I have described and compared the data available on thoracic surgical procedures in HES and the NLCA and used survival analyses to determine that a procedure code in HES is likely to be the most accurate means (within these datasets) of identifying people who had surgery with curative intent for lung cancer. I presented this work as a poster abstract at the British Thoracic Society Winter Meeting, London, 2013.(168)

A brief description of the features of patients who underwent surgery according to this definition has been given, and comparison with previously published data revealed similar survival. The following chapter uses this population to examine 30- and 90-day mortality after lung cancer resection, to determine the patient and tumour features associated with these outcomes and develop a predictive score for use in clinical practice.

CHAPTER 6: RISK FACTORS FOR EARLY DEATH FOLLOWING SURGERY FOR LUNG CANCER

This Chapter starts with a description of the history of risk models in thoracic surgical practice. This is followed by a study in which I investigate the factors associated with early death after lung cancer surgery using the NLCA-HES linked dataset, and use multivariate logistic regression to produce a new predictive model.

6.1 Introduction

6.1.1 Background

Following surgical resection 5-year survival can improve to between 25 and 70% for patients with non-small cell lung cancer (NSCLC), depending on the stage at presentation.(19) However access to this potentially life-saving treatment varies not only between countries but also between NHS Trusts in England: In one study patients first seen in a surgical centre (compared with a hospital without thoracic surgery on site) were 51% more likely to have surgery for NSCLC, even after accounting for variations in factors such as age, performance status and co-morbidity.(60) There is also evidence that trusts with higher resection rates have improved survival and it has been suggested that if all trusts increased their resection rates to this level the proportion of patients surviving lung cancer would improve.(169)

Thoracic surgery is not without risk even in young, relatively fit patients. (170) Factors which influence the decision whether or not to operate include the extent of disease (i.e. whether surgery is likely to provide a cure), patient fitness and the wishes of the individual patient. The extent of disease in which a cure is possible is debated and post-operative adjuvant chemotherapy is often recommended for patients with more advanced disease.(171) The possibility of earlier detection of lung cancer remains an important issue and screening trials and patient awareness campaigns are taking place in the hope that more patients will present at an earlier stage when potentially curative surgery is more likely to be an option. (67, 70) Estimating the level of risk associated with surgical resection in patients who have technically resectable NSCLC is therefore extremely important, but remains a challenge for clinicians who are faced with an aging population who often have multiple co-morbidities.

6.1.2 Rationale for this study

Whilst there are existing tools for the estimation of mortality risk from lung cancer surgery (see below), even in the development of Thoracoscore which is recommended by the British Thoracic Society,(25) only 57% of the thoracic surgical procedures were performed for cancer (not necessarily lung cancer) and the study database was limited to those French hospitals that chose to supply data.(170) The NLCA, in contrast, contains data on over 150,000 patients with primary lung cancer and has been shown to be representative of patients with lung cancer in England through comparison with national cancer registry data. (82) No previous study has looked at factors affecting early post-operative mortality in a UK population.

6.1.3 Aims of this chapter

The first aim of this chapter was to summarise the history and current practice of using risk prediction scores in the pre-operative assessment of patients (section 6.2). Following this, the aim was to use the NLCA, HES and ONS linked data to develop a risk prediction model specifically for lung cancer surgery. This was achieved as follows:

1. *A survival analysis to establish the important time windows on which to base early mortality estimates (sections 5.3 and 6.3).*
2. *Calculation of the proportion of patients who died in this early postoperative period (section 6.3)*
3. *A univariate and multivariate logistic regression analyses to determine risk factors for early post-operative mortality (section 6.3).*
4. *Construction of a predictive model using the results of multivariate analysis of risk factors for early mortality (section 6.4), and*
5. *Comparison of the new model with Thoracoscore (section 6.4). (170)*

6.2 History of surgical mortality risk assessment

A number of tools have been developed to assist clinicians with the estimation of perioperative mortality risk. Initially these were for general surgery and attempts were made to apply them to thoracic surgical cases, but more recently some have been developed specifically for thoracic surgery and one specifically for lung cancer. These predictive models, also known as surgical scoring systems, are described in chronological order.

6.2.1 American Society of Anaesthesiologists

The American Society of Anaesthesiologists (ASA) physical status classification system (figure 6-1) was published in 1941. (172) It is a subjective assessment of patient fitness and was originally designed as a tool for statistical studies; however several groups claim to be able to assign mortality risk estimates to the ASA grades.(173, 174) Surgical mortality risk is related to the type of procedure a patient has to undergo as well as patient fitness, but these studies were usually undertaken within one area surgery (e.g. gastrointestinal or orthopaedic) which to a certain extent removes the variation in risk according to type of procedure.

The classification is subjective and the definitions are ambiguous, which often leads to different grades for the same patient if assigned by more than one anaesthetist or surgeon. (174, 175) Despite these limitations, the ease of use of the ASA grade has meant that it is widely recorded and used (sometimes in combination with other factors such as patient's age) as a tool for comparative audit between centres and has been suggested as a tool for estimating early post-operative mortality risk.(173)

I	A normal healthy patient.
II	A patient with mild systemic disease.
III	A patient with severe systemic disease.
IV	A patient with severe systemic disease that is a constant threat to life.
V	A moribund patient who is not expected to survive without the operation.
VI	A declared brain-dead patient whose organs are being removed for donor purposes.

Figure 6-1: American Society of Anaesthesiologists (ASA) physical status classification system (172)

6.2.2 Goldman cardiac index

In 1977 Goldman published a scoring system to predict risk of cardiovascular complications in non-cardiac surgery.(176) Since deaths after major surgery can be often be attributed to cardiac complications the index has also been used to predict mortality risk: Figure 6-2 shows the factors included in the score, and the way in which it has been used to predict mortality.

Both the ASA grade and to a lesser extent the Goldman index were found to be predictive of perioperative mortality when tested in a series of 16,227 patients (215 of whom died within 4 weeks of operation) at a European tertiary care centre. (173) The majority of these cases had general, orthopaedic, vascular or neurosurgery with only 912 cases undergoing thoracic surgery. In thoracic surgery, particularly for lung cancer, respiratory disease may be equally or more important that cardiac disease and needs to be taken into account when estimating operative mortality risk.

Clinical finding	Score
Third heart sound (S3)	11
Elevated jugulo-venous pressure	11
Myocardial infarction in past 6 months	10
ECG: premature arterial contractions or any rhythm other than sinus	7
ECG shows >5 premature ventricular contractions per minute	7
Age >70 years	5
Intra-thoracic, intra-abdominal or aortic surgery	3
Poor general status, metabolic or bedridden	3

Score	Incidence of death	Incidence of severe CVS complication
>25	56%	22%
<26	4%	17%
<6	0.2%	0.7%

Figure 6-2: Goldman cardiac index, (176)

6.2.3 POSSUM

The physiological and operative severity score for the enumeration of mortality and morbidity (POSSUM) was published in 1991 as a tool for use in general surgical audit. (177) The score was developed by prospective analysis of 1,372 patients who underwent surgery at a single centre in Liverpool, England, 55 of whom died. Procedures carried out for trauma were excluded and the majority of operations were gastrointestinal, hepatobiliary or vascular.

The POSSUM score includes 12 physiological and 6 surgical factors with 4 possible grades for each (figures 6-3 and 6-4). The total physiological and surgical scores are combined and logistic regression analysis is used to give a percentage estimate of mortality and morbidity.

Brunelli *et al* applied the POSSUM score to 250 patients undergoing lung resection in their Italian centre between 1993 and 1996 to assess its use in thoracic as opposed to general surgery. They limited post-operative complications and deaths to those occurring within 30 days or prior to hospital discharge, replaced peritoneal soiling with thoracic cavity soiling and re-defined operative severity to cover the various types of thoracic surgery. There was good correlation between observed and predicted morbidity rates but death was included as a post-operative complication so there was no specific analysis of mortality.(178)

POSSUM was intended to assist surgeons with audit of their own practice by enabling them to account for the type of operations and fitness of their patients when comparing their practice with national standards. The total cannot be calculated until the procedure is complete (physiological factors are scored at the time of surgery and total blood loss cannot be scored until the operation is complete), rendering it less useful in providing patients with information regarding their operative risk prior to surgery.

	Score			
	1	**2**	**4**	**8**
Operative severity*	Minor	Moderate	Major	Major +
Multiple procedures	1		2	>2
Total blood loss (ml)	≤100ml	101-500	501-999	≥1000
Peritoneal soiling	None	Minor (serous fluid)	Local pus	Free bowel content, pus or blood
Mode of surgery	Elective		Emergency resuscitation of >2h possible ‡ Operation <24h after admission	Emergency (immediate surgery <2h needed)

**Surgery of moderate severity includes appendectomy, cholecystectomy, mastectomy, transurethral resection of prostate; major surgery includes any laparotomy, bowel resection, cholecystectomy with choledochotomy, peripheral vascular procedure or major amputation; major + surgery includes any aortic procedure, abdomino-perineal resection, pancreatic or liver resection, oesophagogastrectomy; definitions of surgical procedures with regard to severity are guidelines; not all procedures are listed and the closest should be selected; ‡ indicates that resuscitation is possible even if this period is not actually utilised.*

Figure 6-3: POSSUM Operative severity score, (177)

	Score			
	1	**2**	**4**	**8**
Age (years)	≤60	61-70	≥71	
Cardiac signs	No failure	Diuretic, digoxin, anti-anginal or hypertensive therapy	Peripheral oedema; warfarin therapy	Raised jugulo-venous pressure
Chest radiograph			Borderline cardiomegaly	Cardiomegaly
Respiratory history	No dyspnoea	Dyspnoea on exertion	Limiting dyspnoea (1 flight)	Dyspnoea at rest
Chest radiograph		Mild COAD	Moderate COAD	Fibrosis/consolidation
Systolic blood pressure	110-130mmHg	131-170 or 100-109	≥171 or 90-99	≤89
Pulse (beats/min)	50-80	81-100 or 40-49	101-120	≥121 or ≤39
Glasgow coma score	15	12-14	9-11	≤8
Haemoglobin (g/100ml)	13-16	11.5-12.9 or 16.1-17.0	10.0-11.4 or 17.1-18.0	≤9.9 or ≥18.1
White cell count (x 10^{12}/l)	4-10	10.1-20.0 or 3.1-4.0	≥20.1 or ≤3.0	
Urea (mmol/l)	≤7.5	7.6-10.0	10.1-15.0	≥15.1
Sodium (mmol/l)	≥136	131-135	126-130	≤125
Potassium (mmol/l)	3.5-5.0	3.2-3.4 5.1-5.3	2.9-3.1 or 5.4-5.9	≤2.8 or ≥6.0
Electrocardiogram	Normal		Atrial fibrillation (rate 60-90)	Any other abnormal rhythm or Q waves or ST/T wave changes

COAD, chronic obstructive airways disease.

Figure 6-4: POSSUM Physiological score - to be scored at the time of surgery (177)

6.2.4 E-PASS

The Estimation of Physiologic Ability and Surgical Stress, known as E-PASS, (179) was developed with the aim of reducing postoperative morbidity and mortality by giving an estimate of post-operative risk which would assist in selection of surgical technique. Data on 292 patients undergoing elective gastro-intestinal surgery at a single hospital in Japan between 1992 and 1995 were used to develop the score and it was evaluated in a series of 989 patients who had similar procedures at another Japanese hospital.

The multiple regression analysis included a total of 11 preoperative factors and six surgical factors. The overall score is termed the comprehensive risk score and it is calculated from the physiological risk score and the surgical stress score (figure 6-5). In order to use the score pre-operatively, individual surgeons are expected to estimate the likely blood loss, incision size and operation time based on previous experience in their centre.

The E-PASS score was applied to 282 patients with lung cancer and 458 patients who underwent elective thoracic operations by Yamashita and colleagues who found reasonable correlation between the comprehensive risk score and morbidity; however there were only 5 in-hospital deaths (0.7%) in the study period meaning assessment of mortality prediction was not possible and furthermore suggesting a very low-risk cohort of patients. (180)

Pre-operative risk score = - 0.0686 +

Age		*x 0.00345*
Severe heart disease (NYHA III or IV, or arrhythmia requiring mechanical support)	Presence =1 Absence=0	*x 0.323*
Severe lung disease (VC <60% or FEV1<50%)	Presence =1 Absence=0	*x 0.205*
Diabetes mellitus	Presence =1 Absence=0	*x 0.153*
Performance status (Defined by Japanese Society for Cancer Therapy)	0-4	*x 0.148*
ASA physiological status	1-5	*x 0.0666*

Surgical stress score = - 0.342 +

Blood loss / body weight (g/kg)		*x 0.0139*
Operation time (hours)		*x 0.0392*
Extent of skin incision	Minor=0, Laparotomy or thoracotomy=1, Both=2	*x 0.352*

Comprehensive score = -0.382 + (0.936 x physiological score) + (0.976 surgical score)

The authors suggest that a comprehensive risk score of 1.0 may be taken as a critical threshold at which homeostasis is maintained in surgical patients.

Figure 6-5: Equations for E-PASS scores,(179)

6.2.5 The European Society Subjective & Objective Scores

The European Society Subjective Score (ESSS), published in 2005, was the first score to be developed specifically for thoracic surgery, and was intended to enable fair comparative audit between thoracic surgeons and surgical centres, and to aid prospective clinical decision making. Data were obtained from multiple hospitals (27 units in 14 European countries) in the European Thoracic Surgery Database Project. (181) The inclusion of data from multiple sites was important as surgical skill, equipment and post-operative care (including availability of intensive care beds) vary between hospitals and results based on one centre cannot necessarily be extrapolated to others.

The main analysis included data from 3,426 patients who underwent any thoracic surgical procedure, 66 (1.9%) of whom died in hospital. The ESSS was developed using a randomly selected 60% of this cohort and tested using the remaining 40%. It combines Medical Research Council (MRC) dyspnoea score,

American Society of Anaesthesiologists (ASA) score, class of procedure and age, to give an estimate of the risk of in-hospital mortality.

The authors reported that the ESSS model "performed well at low risk, underestimated mortality at medium risk and overestimated mortality at high risk" (181). Recognising this, and the subjective nature of most of the parameters, the same group went on to develop a further model specifically for lung cancer resections based on the objective measures age and predicted post-operative lung function. This model (the European Society Objective Score (ESOS), figure 6-6) performed reasonably well in estimating mortality a though it was based on a small number of cases and was not fully tested because it was developed after the first part of the study had been completed.

Step 1: $\text{logit}_2 = -5.8858 + (\mathbf{0.0501 \times age}) - (\mathbf{0.0218 \times \text{predicted post-operative } FEV_1\%})$

Step 2: Predicted risk of in hospital death = $\exp(\text{logit}_2) / (1 + \exp(\text{logit}_2))$

Figure 6-6: European Society Objective Score, (181)

Brunelli and colleagues (who previously applied the POSSUM score to their patient group) assessed the performance of the ESOS scoring system for lung cancer resection using prospectively collected data on 695 procedures performed between 2004 and 2006 at three European centres.(182) They found that the score predicted mortality well with no significant differences between observed and ESOS-predicted mortality rates. They attempted a sub-group analysis of the highest risk patients however there were only 31 patients in this group and no deaths were reported.

6.2.6 Thoracoscore

In 2007 Falcoz and colleagues published a 9-factor scoring system (Thoracoscore) which estimated post-operative mortality in thoracic surgery.(170) Thoracoscore was developed using data entered from 59 French hospitals into the thoracic database Epithor. It is based on an analysis of 10,122 patients who underwent a thoracic surgical procedure between June 2002 and July 2005, 218 (2.2%) of whom died before being discharged from hospital. Age, sex, dyspnoea score, ASA score, performance status, priority of surgery (emergency or planned), diagnosis group, procedure class (pneumonectomy or other), and comorbid disease were found to be predictors of early post-operative death and thus comprise the score (figure 6-7). Co-morbidities were scored by number (none, 0-2 or >2). The score was tested using a further 5,061 patients (120 deaths) from the same database and found to perform well in estimating the mortality risk in this group of patients (c-index 0.85).

Thoracoscore was not designed to predict outcomes in lung cancer, although does include a field specifying whether the procedure was for benign or malignant disease. The study included patients undergoing thoracic surgery for a range of indications from the relatively minor spontaneous pneumothorax to complicated pneumonectomy for lung cancer. Patients with a malignant pathology were more than three times as likely to die in hospital as patients with a benign thoracic pathology. For patients with cancer, additional information was collected concerning pathologic staging, type of lymphadenectomy, type of histologic resection, and any adjuvant chemotherapy or radiotherapy received, however none of these fields were included in the final score. Results were only reported for pre-operative treatment, which did not have a significant impact on early mortality.

Thoracoscore was independently tested in 1,675 patients who underwent thoracic surgery in a New York hospital between 2002 and 2006.(183) It was

found to be a strong predictor of mid-term and in-hospital mortality, even when dyspnoea score (which was not recorded in this centre) was excluded. A UK study, however, found that Thoracoscore was not a significant predictor of in-hospital mortality.(184) There were only 16 deaths (2%) in the 703 patients studied so there may have been insufficient power to detect a significant association; further validation studies are needed to assess the performance of Thoracoscore in a population undergoing lung cancer resection.

Variable	Value	*Code*	*Coefficient*
Age (years)	<55	*0*	
	55-65	*1*	*0.7679*
	≥65	*2*	*1.0073*
Sex	Female	*0*	
	Male	*1*	*0.4505*
ASA score	≤2	*0*	
	≥3	*1*	*0.6057*
Performance status	≤2	*0*	
	≥3	*1*	*0.689*
MRC dyspnoea score	≤2	*0*	
	≥3	*1*	*0.9075*
Priority of surgery	Elective	*0*	
	Urgent or emergency	*1*	*0.8443*
Procedure class	Other [a]	*0*	
	Pneumonectomy	*1*	*1.2176*
Diagnosis group	Benign	*0*	
	Malignant	*1*	*1.2423*
Co-morbidity score*	0	*0*	
	1-2	*1*	*0.7447*
	>2	*2*	*0.9065*
Constant	-	-	*-7.3737*

MRC Medical Research Council, ASA American Society of Anaesthesiologists grade,
**Number of significant co-morbid conditions including: smoking, history of cancer, chronic obstructive pulmonary disease, diabetes mellitus, arterial hypertension, peripheral vascular disease, obesity and alcoholism; [a] Other includes mediastinoscopy or other mediastinal surgery, wedge resection, lobectomy or bi-lobectomy.*
Odds = exp (total of coefficients + constant); Probability of death = odds/(1+odds).

Figure 6-7: Thoracoscore: Prediction of risk of in-hospital mortality,(170)

6.2.7 Thoracic surgery for lung cancer

In 2011 Bernard et al published the first model to provide an estimate of mortality specifically following lung cancer resection. (185) This was also the first model to use lung function in any detail. It was developed using Epithor (the same database as Thoracoscore), and was based on data from 18,000 patients who had thoracic surgery for NSCLC. Six-hundred and ninety patients died within 30 days post-operatively or prior to hospital discharge.

Two models were evaluated with the only difference being that model 2 (figure 6-8) used number of co-morbidities whereas model 1 used presence or absence of individual conditions. They also included interaction terms making the model more complex but increasing accuracy by accounting for the difference in effect of pre-operative lung function depending on whether the operation was a pneumonectomy or not. The authors validated both models using the bootstrap sampling method (randomly selecting individuals from the dataset used to generate the score to create a test dataset) and found them to be predictive of mortality in this dataset with an area under the receiver operating curve of 0.78 (95% CI, 0.76–0.80) for model 2. They were, however, unable to test their models in an independent population.

This score is more complex than the widely used Thoracoscore, and is based on data from those institutions which chose to contribute to the Epithor database. It has only been tested in a subgroup of the population used to develop the model and perhaps for these reasons it has not been widely adopted in UK practice.

Variables	Categories	Coefficient
Sex	Female vs. male	*-0.745*
Age	Increasing years	*0.045*
Side	Left vs. right	*-0.42*
ASA score	Increasing units	*0.39*
Performance status	Increasing units	*0.3*
Body mass index (kg/m^2)	≤17	*Ref*
	18-21	*-0.89*
	22-26	*1.18*
	>26	*1.53*
FEV	Increasing %	*-0.01*
Lobectomy	Yes vs. no	*0.56*
Pneumonectomy	Yes vs. no	*1.09*
Pneumonectomy - FEV	Interaction	*0.01*
Side – pneumonectomy	Interaction	*-0.485*
Extended resection	Yes vs. no	*-0.9*
Extended resection – FEV	Interaction	*0.018*
Stage	III vs. (I or II or IV)	*0.47*
Stage	IV vs. (I or II or III)	*0.5*
Number of comorbidities	0	*Ref*
	1	*0.5*
	2	*0.81*
	3 or 4	*0.95*
Intercept (=constant)		*-6.64*

Figure 6-8: Logistic regression models including the number of co-morbidities per patient (model 2) for prediction of in-hospital mortality, (185)

6.2.8 Other studies of risk factors for mortality in lung cancer surgery

The majority of studies which have investigated risk factors for mortality following lung cancer resection incorporate similar patient and surgical factors as shown above. A few suggest that other factors should be taken into account where possible: In a retrospective study of 310 patients at their institution, Stolz et al found that coronary artery disease and respiratory failure (but not COPD, induction therapy, smoking habit or obesity) were statistically associated with an increased risk of death within 30 days of pneumonectomy for lung cancer. (186) One study suggested that adenocarcinoma is associated with reduced mortality in comparison with other histological subtypes, (187) however this may be due the type of patient who develops adenocarcinoma in comparison with squamous cell lung cancer, data on which were not available in this study based on cancer registry data. Exercise testing has been suggested as a potential predictor of

outcome, (188) but this is time consuming and not routinely performed pre-operatively and therefore data with which to assess its predictive value are not currently available. (189)

6.2.9 Post-operative morbidity

Mortality is not the only outcome of interest to patients with lung cancer, or their treating physicians. It is important to consider post-operative dyspnoea, of which several factors including pre-operative lung function have been suggested as predictors,(188, 190-193) and the effects of a sometimes prolonged admission to intensive care, both of which adversely affect quality of life. Several studies, including some of the mortality studies discussed above, have investigated the effects of various factors on morbidity; however compared with mortality it is less well defined. Further descriptions of post-operative morbidity are outside the scope of this thesis.

6.2.10 Summary

I have described a number of predictive models which have been used to estimate mortality after thoracic surgery, whether or not they were designed for this purpose. Some of the models performed well when tested using clinical data, however the numbers of procedures included in these validation studies are reasonably small, particularly for high risk patients and those with lung cancer as the indication for surgery.

There is currently no predictive score or tool to facilitate estimation of perioperative mortality risk based on thoracic surgery in a UK population, which is important given differences in healthcare systems, surgical expertise and patient demographics between countries. The most recent British Thoracic Society (BTS) guidelines on radical management of patients with lung cancer, (25) suggest using a score such as Thoracoscore when evaluating and consenting patients for surgery despite the limitations I have described.

6.3 Analysis of factors associated with early mortality following surgery for NSCLC

6.3.1 Aims

The aims of this section were to establish the important time windows on which to base early mortality estimates, calculate the proportion of patients who died in this early postoperative period and perform univariate and multivariate logistic regression analyses to determine risk factors for early post-operative mortality (*chapter aims 1-3*).

6.3.2 Methods

Study population

Patients were included if they had a record in the NLCA-HES linked database and were first seen between January 1st 2004 and March 31st 2010. People with NSCLC were identified and those with stage 3b or 4, or age less than 30 years, were excluded.

Definition of surgery

The definition of a potentially curative thoracic surgical procedure, following the work described in chapter 5, was an OPCS-4 code in the HES database for a thoracic surgical procedure which was likely to have been performed with curative intent for NSCLC (Appendix E). Procedures which took place before 1st January 2004 or after 31st March 2010 were excluded.

Procedures were categorised as pneumonectomy, bi-lobectomy, lobectomy, segmentectomy / wedge resection or other. Where more than one relevant procedure code was identified for an individual, the most extensive procedure was used; for the few patients who had codes for more than one procedure of the same type, the date of the most recent procedure was used. If recorded procedure dates differed between HES and the NLCA the difference between the

dates was calculated: if this was more than 10 days the patient was excluded from further analysis.

Patients with a procedure date more than 3 months before or 6 months after their lung cancer diagnosis were also excluded so that patient features, which are usually recorded at the time of diagnosis, might still be representative of the state of the patient at the time of surgery.

After the procedure date had been defined any record with a code for metastatic cancer prior to the date of surgery was excluded.

Definition of outcome

Office of National Statistics (ONS) dates of death and HES procedure dates were used to determine when a patient died in relation to their operation.

In order to ensure the analysis was based on the most appropriate postoperative period, the Kaplan Meier survival curves for death following surgery were inspected to determine the time during which mortality was highest (Chapter 5 figures 5-6 to 5-9). The rate of death was actually very similar over the six months following lung cancer resection, with only a slight increase observed in approximately the first 0-90 days. Surgical mortality is traditionally defined as within 30 days of the operation and therefore this was one of the outcomes used. A review of the literature, however, revealed that recovery usually takes at least 3 months.(194) It was therefore decided that deaths should be assessed within 30 and 90 days of surgery with an analysis comparing the patients who died within 30 days with those who died between 31 and 90 days.

Covariate definitions

Stage, histology, lung function, performance status and socio-economic status (Townsend quintile) were defined as described in section 2.2.5.

Statistical methods

The proportions of patients who died within 30 and 90 days of surgery were calculated. To determine whether there were marked differences in the features of patients who died within the first 30 days after surgery and those who died between 31 and 90 days, demographic, co-morbid, tumour, and procedure related features of patients who died in these two time periods were compared.

Logistic regression was used to estimate odds ratios (ORs) associated with demographic, co-morbid, tumour and procedure-related factors for death within 30 days and within 90 days. A multivariate model was built including all factors that were significantly associated (defined as $p<0.1$) with death in univariate analysis. The significance of each variable in the multivariate model was then assessed using a likelihood ratio test.

Interactions

Interactions between the following factors and death within 30 and 90 days of surgery were sought:

- Lung function and procedure type
- Side of surgery and procedure type

To improve the power for these analyses, procedure type was re-classified as pneumonectomy or non-pneumonectomy.

Sensitivity analyses

Previous studies and data reports from the NLCA have found a substantial proportion of missing data for performance status, stage, and lung function. (51, 55, 60) Sensitivity analyses were therefore planned restricted to records with complete data for these three variables. Since Charlson Index was a derived field no missing values were generated; other fields were considered less important and any analysis restricted to records which were entirely complete was not considered feasible. Because of the reduced power in this smaller dataset

Charlson index was re-coded as a binary variable (0-1, ≥2) and age as <55, 55-65, 66-75 and >75 years.

Stage may be recorded using either version 6 or version 7 of the UICC system from 2009 onwards. In case this affected the results a sensitivity analysis was performed restricted to cases where version 6 was used as this would be the majority of patients.

6.3.3 Results

There were 113,261 patients with NSCLC in the NLCA database first seen between January 2004 and March 2010; 46,013 with stage recorded as 3b or 4 and a further 102 who were aged less than 30 years at diagnosis were excluded. Of the remaining patients, 12,269 had an OPCS-4 code for a potentially curative procedure in HES between January 1st 2004 and March 31st 2010. Two-hundred and ninety-two patients where the difference between the procedure date recorded in the NLCA and that recorded in HES was >10 days, 437 who had an ICD-10 code for metastatic cancer recorded prior to the procedure date, and a further 549 where the procedure date was >3 months before or >6 months after the NLCA date of diagnosis were excluded, leaving 10,991 patients for analysis.

The study population and exclusions are the same as the final part of the previous chapter, and were shown in figure 5-4.

The majority (56%) of patients were male, 20% were aged 70-74 years and 31% had a performance status of zero (Table 6-1). The most commonly performed procedure was lobectomy (64%) with only 10% of patients having had a pneumonectomy. Twenty-eight per cent of patients had stage 1b NSCLC, although 26% did not have a pre- or post-operative stage recorded. The most common pre-operative histological subtype was adenocarcinoma (31%) followed by squamous cell (28%). Twenty-one per cent did not have a record of pre- or post-operative histology; in most cases this is likely to reflect missing data rather than the absence of histological confirmation of lung cancer.

Table 6-1: Proportions and characteristics of patients who died within 30-days and between 31-and 90 days of surgery

		Overall N=10,991		Died within 30 days (n=334)		Died between 31-90 days (n=313)		
		n	***%***	**n**	**%**[a]	**n**	**%**[b]	***Chi*** [2]
Sex	Female	4,824	*43.9*	107	*2.2*	103	*2.1*	
	Male	6,167	*56.1*	227	*3.7*	210	*3.4*	*p=0.813*
Age group	<55	1,008	*9.2*	12	*1.2*	23	*2.3*	
	55-59	1,090	*9.9*	21	*1.9*	24	*2.2*	
	60-64	1,847	*16.8*	31	*1.7*	32	*1.7*	
	65-69	2,128	*19.4*	56	*2.6*	49	*2.3*	
	70-74	2,226	*20.3*	84	*3.8*	78	*3.5*	
	75-79	1,828	*16.6*	88	*4.8*	62	*3.4*	
	80-84	730	*6.6*	34	*4.7*	35	*4.8*	
	85+	134	*1.2*	12	*9.0*	10	*7.5*	*p=0.331*
Ethnicity	White	8,983	*81.7*	254	*2.8*	256	*2.8*	
	Black	82	*0.7*	1	*1.2*	0	*0.0*	
	Asian	112	*1.0*	2	*1.8*	1	*0.9*	
	Other	77	*0.7*	1	*1.3*	2	*2.6*	*p=0.643**
	Missing	1,737	*15.8*	76	*4.4*	54	*3.1*	*p=0.317*
Townsend quintile	1	1,343	*12.2*	36	*2.7*	30	*2.2*	
	2	1,602	*14.6*	45	*2.8*	41	*2.6*	
	3	1,609	*14.6*	52	*3.2*	46	*2.9*	
	4	1,723	*14.6*	64	*4.0*	55	*3.4*	
	(most deprived) 5	2,074	*15.7*	74	*4.3*	58	*3.4*	*p=0.985**
	Missing	2,640	*18.9*	63	*3.0*	83	*4.0*	*p=0.328*
Performance status	0	3,422	*31.1*	72	*2.1*	60	*1.8*	
	1	2,815	*25.6*	84	*3.0*	93	*3.3*	
	2	465	*4.2*	23	*4.9*	28	*6.0*	
	3-4	108	*1.0*	11	*10.2*	9	*8.3*	*p=0.524**
	Missing	4,181	*38.0*	144	*3.4*	123	*2.9*	*p=0.523*
Per cent Predicted FEV_1	>80%	1,891	*17.2*	34	*1.8*	39	*2.1*	
	60-79%	1,499	*13.6*	42	*2.8*	45	*3.0*	
	40-59%	726	*6.6*	23	*3.2*	27	*3.7*	
	<40%	141	*1.3*	5	*3.5*	7	*5.0*	*p=0.975**
	Missing	6,734	*61.3*	230	*3.4*	195	*2.9*	*p=0.508*
Charlson index	0	5,456	*49.6*	128	*2.3*	130	*2.4*	
	1	2,791	*25.4*	87	*3.1*	84	*3.0*	
	2-3	2,233	*20.3*	95	*4.3*	81	*3.6*	
	≥4	511	*4.6*	24	*4.7*	18	*3.5*	*p=0.715*
Stage	IA	2,249	*20.5*	37	*1.6*	41	*1.8*	
	IB	3,064	*27.9*	87	*2.8*	82	*2.7*	
	IIA	334	*3.0*	6	*1.8*	3	*0.9*	
	IIB	1,494	*13.6*	55	*3.7*	47	*3.1*	
	IIIA	933	*8.5*	41	*4.4*	39	*4.2*	*p=0.812**
	missing	2,857	*26.0*	108	*3.8*	101	*3.5*	*p=0.903*
Side	Right	5,067	*46.1*	157	*3.1*	161	*3.2*	
	Left	3,930	*35.8*	105	*2.7*	114	*2.9*	
	Other	85	*0.8*	5	*5.9*	4	*4.7*	*p=0.875**
	Missing	1,909	*17.4*	67	*3.5*	34	*1.8*	*p=0.014*
Histology	Adenocarcinoma	3,406	*31.0*	60	*1.8*	75	*2.2*	
	Squamous cell	3,106	*28.3*	125	*4.0*	107	*3.4*	
	NSCLC NOS	1,833	*16.7*	62	*3.4*	62	*3.4*	
	Other	368	*3.3*	10	*2.7*	12	*3.3*	*p=0.356**
	Missing	2,278	*20.7*	77	*3.4*	57	*2.5*	*p=0.235*
Procedure	Segmentectomy / wedge	1,671	*15.2*	35	*2.1*	35	*2.1*	
	Lobectomy	7,051	*64.2*	160	*2.3*	165	*2.3*	
	Bi-lobectomy	431	*3.9*	25	*5.8*	13	*3.0*	
	Pneumonectomy	1,121	*10.2*	78	*7.0*	51	*4.5*	
	Other	717	*6.5*	36	*5.0*	49	*6.8*	*p=0.028*

Excluding missing; FEV_1 forced expiratory volume in 1 second; [a] % of patients in each subgroup who died within 30 days; [b] % of patients in each subgroup who died between 31 and 90 days.

Mortality

Three per cent of patients (334) died within 30 days of their procedure and a further 2.9% (313) between 31 and 90 days (therefore a total of 5.9% (647) died within 90 days). There were no statistically significant differences in patient, co-morbidity or tumour factors between patients who died within 30 days of their procedure and those who died between 31 and 90 days (Table 6-1). A higher proportion of those who died within 30 days had a pneumonectomy or bi-lobectomy compared with those who died between 31 and 90 days.

Given these findings, and the greater degree of accuracy due to a higher number of deaths, I have elected to report results for 90-day mortality in this section; the results for death within 30 days of surgery are similar and are shown in Table 6-3.

Within 90 days of surgery, males were more likely to die than females (7.1% vs. 4.4%) and the proportion of patients who died after pneumonectomy was higher than for lobectomy (11.5% vs. 4.6%) (Table 6-2). Sixteen per cent of patients over 85 years and 18.5% of those with performance status 3-4 died within this post-operative period. Age was strongly associated with post-operative mortality: Compared with a patient aged 70-74 years, the odds of death within 90 days of surgery for a patient aged >85 years were markedly increased, even after accounting for other demographic, tumour and co-morbidity factors (adjusted OR 2.84, 95% confidence interval (CI) 1.71-4.71) (Table 6-2). The next most strongly associated factors were procedure type and performance status. Significant associations were also observed with percentage predicted FEV_1, stage, Charlson index, Townsend score, ethnicity, histological subtype and sex.

Table 6-2: Risk factors for early post-operative death: Death within 90 days proportions and odds ratios

		Overall N=10,991		Died within 90 days of surgery (n = 647)					
		n	**%**	**n**	**%**	**OR**	***95% CI***	**AdjustedOR***	***95% CI***
Sex	Female	4,824	*43.9*	210	*4.4*	1.00		1.00	
	Male	6,167	*56.1*	437	*7.1*	1.68	*1.42-1.98*	1.37	*1.15-1.63*
							p<0.0001		*p=0.0004*
Age group	<55	1,008	*9.2*	35	*3.5*	0.46	*0.32-0.67*	0.46	*0.32-0.68*
	55-59	1,090	*9.9*	45	*4.1*	0.55	*0.39-0.77*	0.53	*0.37-0.75*
	60-64	1,847	*16.8*	63	*3.4*	0.45	*0.33-0.61*	0.44	*0.32-0.59*
	65-69	2,128	*19.4*	105	*4.9*	0.63	*0.49-0.82*	0.61	*0.47-0.79*
	70-74	2,226	*20.3*	162	*7.3*	1.00		1.00	
	75-79	1,828	*16.6*	150	*8.2*	1.14	*0.90-1.44*	1.19	*0.94-1.51*
	80-84	730	*6.6*	69	*9.5*	1.33	*0.99-1.79*	1.46	*1.07-1.98*
	85+	134	*1.2*	22	*16.4*	2.50	*1.54-4.06*	2.84	*1.71-4.71*
							*p<0.0001***		*p<0.0001***
Ethnicity	White	8,983	*81.7*	510	*5.7*	1.00		1.00	
	Black	82	*0.7*	1	*1.2*	0.21	*0.03-1.48*	0.21	*0.03-1.51*
	Asian	112	*1.0*	3	*2.7*	0.46	*0.14-1.44*	0.45	*0.14-1.46*
	Other	77	*0.7*	3	*3.9*	0.67	*0.21-2.14*	0.71	*0.22-2.31*
	Missing	1,737	*15.8*	130	*7.5*	1.34	*1.10-1.64*	1.46	*1.19-1.80*
							p=0.0030		*p=0.0005*
Townsend quintile	1	1,343	*12.2*	66	*4.9*	1.00		1.00	
	2	1,602	*14.6*	86	*5.4*	1.10	*0.79-1.53*	1.12	*0.80-1.57*
	3	1,609	*14.6*	98	*6.1*	1.25	*0.91-1.73*	1.35	*0.97-1.88*
	4	1,723	*14.6*	119	*6.9*	1.44	*1.05-1.96*	1.59	*1.16-2.19*
	(most deprived) 5	2,074	*15.7*	132	*6.4*	1.32	*0.97-1.78*	1.45	*1.06-1.99*
	Missing	2,640	*18.9*	146	*5.5*	1.13	*0.84-1.53*	1.13	*0.83-1.55*
							*p=0.0221***		*p=0.0028***
PS	0	3,422	*31.1*	132	*3.9*	1.00		1.00	
	1	2,815	*25.6*	177	*6.3*	1.67	*1.33-2.11*	1.38	*1.09-1.75*
	2	465	*4.2*	51	*11.0*	3.07	*2.19-4.31*	2.40	*1.68-3.41*
	3-4	108	*1.0*	20	*18.5*	5.66	*3.38-9.49*	4.08	*2.37-7.02*
	Missing	4,181	*38.0*	267	*6.4*	1.70	*1.37-2.11*	1.35	*1.06-1.73*
							*p<0.0001***		*p<0.0001***
Per cent predicted FEV1	>80%	1,891	*17.2*	73	*3.9*	1.00		1.00	
	60-79%	1,499	*13.6*	87	*5.8*	1.53	*1.12-2.11*	1.37	*0.99-1.90*
	40-59%	726	*6.6*	50	*6.9*	1.84	*1.27-2.67*	1.64	*1.12-2.41*
	<40%	141	*1.3*	12	*8.5*	2.32	*1.23-4.38*	2.07	*1.06-4.04*
	Missing	6,734	*61.3*	425	*6.3*	1.68	*1.30-2.16*	1.48	*1.13-1.95*
							*p=0.0002***		*p=0.0020***
Charlson index	0	5,456	*49.6*	258	*4.7*	1.00		1.00	
	1	2,791	*25.4*	171	*6.1*	1.31	*1.08-1.60*	1.20	*0.98-1.48*
	2-3	2,233	*20.3*	176	*7.9*	1.72	*1.41-2.10*	1.54	*1.25-1.90*
	≥4	511	*4.6*	42	*8.2*	1.80	*1.28-2.53*	1.53	*1.07-2.18*
							*p<0.0001***		*p<0.0001***
Stage	IA	2,249	*20.5*	78	*3.5*	1.00		1.00	
	IB	3,064	*27.9*	169	*5.5*	1.62	*1.24-2.14*	1.39	*1.05-1.84*
	IIA	334	*3.0*	9	*2.7*	0.77	*0.38-1.55*	0.67	*0.33-1.37*
	IIB	1,494	*13.6*	102	*6.8*	2.04	*1.51-2.76*	1.59	*1.16-2.19*
	IIIA	933	*8.5*	80	*8.6*	2.44	*1.77-3.36*	1.85	*1.32-2.60*
	missing	2,857	*26.0*	209	*7.3*	2.20	*1.68-2.87*	1.78	*1.32-2.41*
							*p<0.0001***		*p=0.0004***
Side	Right	5,067	*46.1*	318	*6.3*	1.00			
	Left	3,930	*35.8*	219	*5.6*	0.88	*0.74-1.05*		
	Other	85	*0.8*	9	*10.6*	1.77	*0.88-3.56*		
	Missing	1,909	*17.4*	101	*5.3*	0.83	*0.66-1.05*		
							p=0.1063		
Histology	Adenocarcinoma	3,406	*31.0*	135	*4.0*	1.00		1.00	
	Squamous cell	3,106	*28.3*	232	*7.5*	1.96	*1.57-2.43*	1.38	*1.10-1.73*
	NSCLC NOS	1,833	*16.7*	124	*6.8*	1.76	*1.37-2.26*	1.36	*1.05-1.76*
	Other	368	*3.3*	22	*6.0*	1.54	*0.97-2.45*	1.14	*0.70-1.84*
	Missing	2,278	*20.7*	134	*5.9*	1.51	*1.19-1.94*	1.08	*0.82-1.42*
							p<0.0001		*p=0.0420*
Procedure	Segmentectomy/wedge	1,671	*15.2*	70	*4.2*	0.90	*0.69-1.16*	0.80	*0.61-1.05*
	Lobectomy	7,051	*64.2*	325	*4.6*	1.00		1.00	
	Bi-lobectomy	431	*3.9*	38	*8.8*	2.00	*1.41-2.84*	1.94	*1.35-2.78*
	Pneumonectomy	1,121	*10.2*	129	*11.5*	2.69	*2.17-3.33*	2.81	*2.22-3.56*
	Other	717	*6.5*	85	*11.9*	2.78	*2.16-3.57*	2.12	*1.62-2.77*
							p<0.0001		*p<0.0001*

OR odds ratio; CI confidence interval; PS performance status; FEV_1 forced expiratory volume in 1 second. *ORs are adjusted for all other factors for which adjusted ORs are given. All p values calculated using likelihood ratio test; ** LRT p for trend.

Table 6-3: Factors associated with death within 30 days of surgery

		Overall N=10,991		Died within 30 days of surgery (n= 334)					
		n	%	n	%	OR	95% CI	Adjusted OR	95% CI
Sex	Female	4,824	43.9	107	2.2	1.00		1.00	
	Male	6,167	56.1	227	3.7	1.62	1.33-2.13	1.30	1.02-1.66
							p<0.0001		**p=0.0320**
Age group	<55	1,008	9.2	12	1.2	0.31	0.17-0.57	0.29	0.16-0.54
	55-59	1,090	9.9	21	1.9	0.50	0.31-0.81	0.46	0.28-0.76
	60-64	1,847	16.8	31	1.7	0.44	0.29-0.66	0.41	0.27-0.63
	65-69	2,128	19.4	56	2.6	0.64	0.45-0.91	0.61	0.42-0.87
	70-74	2,226	20.3	84	3.8	1.00		1.00	
	75-79	1,828	16.6	88	4.8	1.29	0.95-1.75	1.38	1.01-1.88
	80-84	730	6.6	34	4.7	1.25	0.83-1.87	1.43	0.94-2.17
	85+	134	1.2	12	9.0	2.51	1.33-4.72	2.84	1.47-5.48
							p<0.0001**		**p<0.0001****
Ethnicity	White	8,983	81.7	254	2.8	1.00		1.00	
	Black	82	0.7	1	1.2	0.42	0.06-3.06	0.44	0.06-3.22
	Asian	112	1.0	2	1.8	0.62	0.15-2.54	0.66	0.16-2.73
	Other	77	0.7	1	1.3	0.45	0.06-3.26	0.45	0.06-3.32
	Missing	1,737	15.8	76	4.4	1.57	1.21-2.04	1.73	1.32-2.26
							p=0.0088		**p=0.0017**
Townsend quintile	1	1,343	12.2	36	2.7	1.10		1.00	
	2	1,602	14.6	45	2.8	1.05	0.67-1.64	1.07	0.68-1.68
	3	1,609	14.6	52	3.2	1.21	0.79-1.87	1.33	0.85-2.06
	4	1,723	14.6	64	4.0	1.40	0.93-2.12	1.59	1.04-2.43
	(most deprived) 5	2,074	15.7	74	4.3	1.34	0.90-2.01	1.52	1.00-2.31
	Missing	2,640	18.9	63	3.0	0.89	0.59-1.34	0.84	0.54-1.29
							p=0.0558**		**p=0.0098****
PS	0	3,422	31.1	72	2.1	1.00		1.00	
	1	2,815	25.6	84	3.0	1.43	1.04-1.97	1.16	0.84-1.61
	2	465	4.2	23	4.9	2.42	1.50-3.91	1.84	1.12-3.03
	3-4	108	1.0	11	10.2	5.28	2.71-10.27	3.77	1.87-7.58
	Missing	4,181	38.0	144	3.4	1.66	1.25-2.21	1.33	0.96-1.85
							p<0.0001**		**p=0.0001****
Per cent predicted FEV1	>80%	1,891	17.2	34	1.8	1.00		1.00	
	60-79%	1,499	13.6	42	2.8	1.57	1.00-2.49	1.41	0.88-2.25
	40-59%	726	6.6	23	3.2	1.79	1.05-3.05	1.63	0.94-2.84
	<40%	141	1.3	5	3.5	2.01	0.77-5.22	1.96	0.73-5.28
	Missing	6,734	61.3	230	3.4	1.93	1.34-2.78	1.72	1.17-2.53
							p=0.0176**		**p=0.0895****
Charlson index	0	5,456	49.6	128	2.3	1.00		1.00	
	1	2,791	25.4	87	3.1	1.34	1.02-1.77	1.23	0.93-1.63
	2-3	2,233	20.3	95	4.3	1.85	1.41-2.42	1.66	1.25-2.19
	≥4	511	4.6	24	4.7	2.05	1.31-3.20	1.77	1.11-2.81
							p<0.0001**		**p=0.0002****
Stage	IA	2,249	20.5	37	1.6	1.00		1.00	
	IB	3,064	27.9	87	2.8	1.75	1.18-2.58	1.42	0.96-2.12
	IIA	334	3.0	6	1.8	1.09	0.46-2.61	0.90	0.37-2.19
	IIB	1,494	13.6	55	3.7	2.28	1.50-3.48	1.66	1.06-2.59
	IIIA	933	8.5	41	4.4	2.57	1.64-4.04	1.85	1.15-2.98
	missing	2,857	26.0	108	3.8	2.35	1.61-3.43	1.71	1.12-2.60
							p<0.0001**		**p=0.0143****
Side	Right	5,067	46.1	157	3.1	1.00			
	Left	3,930	35.8	105	2.7	0.86	0.67-1.10		
	Other	85	0.8	5	5.9	1.95	0.78-4.89		
	Missing	1,909	17.4	67	3.5	1.14	0.85-1.52		
							p=0.1615		
Histology	Adenocarcinoma	3,406	31.0	60	1.8	1.00		1.00	
	Squamous cell	3,106	28.3	125	4.0	2.34	1.71-3.19	1.57	1.14-2.18
	NSCLC NOS	1,833	16.7	62	3.4	1.95	1.36-2.80	1.67	1.01-2.12
	Other	368	3.3	10	2.7	1.56	0.79-3.07	1.12	0.56-2.25
	Missing	2,278	20.7	77	3.4	1.95	1.39-2.75	1.46	1.00-2.13
							p<0.0001		**p=0.0643**
Procedure	Segmentectomy/wedge	1,671	15.2	35	2.1	0.92	0.64-1.33	0.82	0.56-1.19
	Lobectomy	7,051	64.2	160	2.3	1.00		1.00	
	Bi-lobectomy	431	3.9	25	5.8	2.65	1.72-4.09	2.61	1.67-4.07
	Pneumonectomy	1,121	10.2	78	7.0	3.22	2.44-4.25	3.54	2.60-4.81
	Other	717	6.5	36	5.0	2.28	1.57-3.30	1.69	1.14-2.50
							p<0.0001		**p<0.0001**

OR odds ratio; CI confidence interval; PS performance status; FEV_1 forced expiratory volume in 1 second. *ORs are adjusted for all other factors for which adjusted ORs are given. All p values calculated using likelihood ratio test; ** LRT p for trend.

Interactions

No significant interactions were found for death within either 30 or 90 days, between procedure (pneumonectomy or not pneumonectomy) and FEV1 or side of surgery.

Missing data and sensitivity analyses

There were 3,319 patients with complete data on performance status, stage and lung function. The proportions of these patients who died within 30 and 90 days of surgery were slightly lower (2.5% and 5.1%) than in the overall population of 10,991, however in multivariate analysis age, procedure type, performance status, stage, Charlson index and FEV_1 were again found to be significantly associated with early post-operative death, with similar odds ratios to the initial analysis (Tables 6-4 and 6-5).

Repeat analysis excluding the 83 records (0.7%) which used staging version 7 to record stage and (also excluding the single record in which staging version was not recorded) produced results which were almost identical to those displayed in tables 1 and 2.

Table 6-4: Factors associated with death within 90 days of surgery for patients with records of performance status, stage, and lung function.

		Overall N=3,319		Died within 90 days of surgery (n=169)					
		n	**%**	**n**	**%**	**OR**	***95% CI***	**Adjusted OR***	***95% CI***
Sex	Female	1,482	*44.7*	57	*3.8*	1.00		1.00	
	Male	1,837	*55.3*	112	*6.1*	1.62	*1.17-2.25*	1.24	*0.88-1.75*
							p=0.0030		***p=0.2196***
Age group	<55	268	*8.1*	5	*1.9*	0.27	*0.11-0.68*	0.32	*0.12-0.83*
	55-59	272	*8.2*	8	*2.9*	0.43	*0.20-0.92*	0.40	*0.18-0.88*
	60-64	583	*17.6*	18	*3.1*	0.45	*0.26-0.78*	0.45	*0.25-0.80*
	65-69	665	*20.0*	29	*4.4*	0.64	*0.40-1.04*	0.61	*0.37-1.00*
	70-74	694	*20.9*	46	*6.6*	1.00		1.00	
	75-79	566	*17.1*	40	*7.1*	1.07	*0.69-1.66*	1.15	*0.73-1.82*
	80-84	236	*7.1*	18	*7.6*	1.16	*0.66-2.05*	1.43	*0.80-2.58*
	85+	35	*1.1*	5	*14.3*	2.35	*0.87-6.34*	3.17	*1.12-9.01*
							p<0.0001		***p<0.0001***
Ethnicity	White	2,769	*83.4*	146	*5.3*	1.00			
	Black	21	*0.6*	-	-				
	Asian	28	*0.8*	-	-				
	Other	24	*0.7*	-	-				
	Missing	477	*14.4*	23	*4.8*	0.91	*0.58-1.43*		
Townsend quintile	1	448	*13.5*	14	*3.1*	1.00			
	2	512	*15.4*	23	*4.5*	1.46	*0.74-2.87*		
	3	531	*16.0*	39	*7.3*	2.46	*1.32-4.59*		
	4	615	*18.5*	28	*4.6*	1.48	*0.77-2.84*		
	(most deprived) 5	772	*23.3*	38	*4.9*	1.60	*0.86-3.00*		
	Missing	441	*13.3*	27	*6.1*	2.02	*1.05-3.91*		
							p=0.3325		
PS	0	1,674	*50.4*	*51*	*3.0*	1.00		1.00	
	1	1,390	*41.9*	*97*	*7.0*	2.39	*1.69-3.38*	1.89	*1.31-2.72*
	2	215	*6.5*	*19*	*8.8*	3.08	*1.78-5.33*	2.39	*1.34-4.28*
	3-4	40	*1.2*	*2*	*5.0*	1.67	*0.39-7.13*	1.67	*0.27-5.11*
							p<0.0001		***p=0.0055***
Per cent predicted FEV1	>80%	1,508	*45.4*	58	*3.8*	1.00		1.00	
	60-79%	1,145	*34.5*	58	*5.1*	1.33	*0.92-1.94*	1.25	*0.85-1.84*
	40-59%	557	*16.8*	43	*7.7*	2.09	*1.39-3.14*	2.01	*1.30-3.10*
	<40%	109	*3.3*	10	*9.2*	2.53	*1.25-5.09*	2.78	*1.31-5.88*
							p=0.0001		***p=0.0004***
Charlson index	0-1	2,508	*75.6*	115	*4.6*	1.00		1.00	
	≥2	811	*24.4*	54	*6.7*	1.22	*1.03-1.44*	1.19	*1.00-1.42*
							p=0.0233		***p=0.0514***
Stage	IA	949	*28.6*	30	*3.2*	1.00		1.00	
	IB	1,237	*37.3*	63	*5.1*	1.64	*1.06-2.56*	1.42	*0.90-2.25*
	IIA	131	*3.9*	4	*3.1*	0.96	*0.33-2.78*	0.77	*0.26-2.27*
	IIB	614	*18.5*	41	*6.7*	2.19	*1.35-3.55*	1.70	*1.01-2.87*
	IIIA	388	*11.7*	31	*8.0*	2.66	*1.59-4.46*	2.18	*1.25-3.81*
							p=0.0001		***p=0.0085***
Side	Right	1,841	*55.5*	96	*5.2*	1.00			
	Left	1,365	*41.1*	64	*4.7*	0.89	*0.65-1.24*		
	Other	18	*0.5*	1	*5.6*	1.07	*0.14-8.12*		
	Missing	95	*2.9*	8	*8.4*	1.67	*0.79-3.55*		
							P=0.4988		
Histology (Pre-op)	Adenocarcinoma	1,155	*34.8*	43	*3.7*	1.00		1.00	
	Squamous cell	1,151	*34.7*	74	*6.4*	1.78	*1.21-2.61*	1.08	*0.72-1.63*
	NSCLC NOS	662	*19.9*	35	*5.3*	1.44	*0.91-2.28*	1.03	*0.64-1.66*
	Other	82	*2.5*	3	*3.7*	0.98	*0.30-3.24*	0.83	*0.25-2.82*
	Missing	269	*8.1*	14	*5.2*	1.42	*0.77-2.63*	1.37	*0.73-2.60*
							P=0.0538		***p=0.8873***
Procedure	Segmentectomy/wedge	441	*13.3*	16	*3.6*	0.88	*0.51-1.52*	0.76	*0.44-1.33*
	Lobectomy	2,275	*68.5*	93	*4.1*	1.00		1.00	
	Bi-lobectomy	131	*3.9*	12	*9.2*	2.37	*1.26-4.44*	2.37	*1.23-4.60*
	Pneumonectomy	335	*10.1*	40	*11.9*	3.18	*2.15-4.70*	3.36	*2.17-5.20*
	Other	137	*4.1*	8	*5.8*	1.46	*0.69-3.06*	1.10	*0.51-2.36*
							p<0.0001		***p<0.0001***

OR odds ratio; CI confidence interval; PS performance status; FEV_1 forced expiratory volume in 1 second. *ORs are adjusted for all other factors for which adjusted ORs are given. All p values calculated using likelihood ratio test; ** LRT p for trend.

Table 6-5: Factors associated with death within 30 days of surgery for patients with records of performance status, stage, and lung function.

		Overall N=3,319		Died within 30 days of surgery (n=82)					
		n	***%***	**n**	***%***	**OR**	***95% CI***	**Adjusted OR***	***95% CI***
Sex	Female	1,482	*44.7*	24	*1.6*	1.00		1.00	
	Male	1,837	*55.3*	58	*3.2*	1.98	*1.22-3.20*	1.47	*0.89-2.44*
							p=0.0038		***p=0.1244***
Age group	<55	268	*8.1*	2	*0.7*	0.25	*0.06-1.09*	0.32	*0.07-1.38*
	55-59	272	*8.2*	3	*1.1*	0.38	*0.11-1.28*	0.36	*0.10-1.27*
	60-64	583	*17.6*	9	*1.5*	0.53	*0.24-1.17*	0.54	*0.24-1.22*
	65-69	665	*20.0*	20	*3.0*	1.04	*0.56-1.96*	1.00	*0.52-1.91*
	70-74	694	*20.9*	20	*2.9*	1.00		1.00	
	75-79	566	*17.1*	19	*3.4*	1.17	*0.62-2.22*	1.26	*0.66-2.43*
	80-84	236	*7.1*	6	*2.5*	0.88	*0.35-2.22*	1.11	*0.43-2.86*
	85+	35	*1.1*	3	*8.6*	3.16	*0.89-11.19*	4.00	*1.04-15.38*
							p=0.0015		
Ethnicity	White	2,769	*83.4*	69	*2.5*	1.00			
	Black	21	*0.6*	-	-				
	Asian	28	*0.8*	-	-				
	Other	24	*0.7*	-	-				
	Missing	477	*14.4*	13	*2.7*	1.10	*0.60-2.00*		
Townsend quintile	1	448	*13.5*	7	*1.6*	1.00			
	2	512	*15.4*	12	*2.3*	1.51	*0.59-3.87*		
	3	531	*16.0*	20	*3.8*	2.47	*1.03-5.89*		
	4	615	*18.5*	13	*2.1*	1.36	*0.54-3.44*		
	(most deprived) 5	772	*23.3*	20	*2.6*	1.68	*0.70-3.99*		
	Missing	441	*13.3*	10	*2.3*	1.46	*0.55-3.88*		
							p=0.4889		
PS	0	1,674	*50.4*	22	*1.3*	1.00		1.00	
	1	1,390	*41.9*	47	*3.4*	2.63	*1.58-4.38*	2.26	*1.33-3.85*
	2	215	*6.5*	12	*5.6*	4.44	*2.16-9.10*	3.74	*1.74-8.06*
	3-4	40	*1.2*	1	*2.5*	1.93	*0.25-14.65*	1.93	*0.19-11.39*
							p=0.0002		***p=0.0046***
Per cent predicted FEV1	>80%	1,508	*45.4*	28	*1.9*	1.00		1.00	
	60-79%	1,145	*34.5*	29	*2.5*	1.37	*0.81-2.32*	1.26	*0.73-2.17*
	40-59%	557	*16.8*	20	*3.6*	1.97	*1.10-3.52*	1.76	*0.95-3.26*
	<40%	109	*3.3*	5	*4.6*	2.54	*0.96-6.72*	2.25	*0.79-6.39*
							p=0.0094		***p=0.0408***
Charlson index	0-1	2,508	*75.6*	56	*2.2*	1.00			
	≥2	811	*24.4*	26	*3.2*	1.20	*0.95-1.52*		
							p=0.1316		
Stage	IA	949	*28.6*	14	*1.5*	1.00		1.00	
	IB	1,237	*37.3*	29	*2.3*	1.60	*0.84-3.05*	1.32	*0.68-2.56*
	IIA	131	*3.9*	2	*1.5*	1.04	*0.23-4.61*	0.72	*0.16-3.32*
	IIB	614	*18.5*	21	*3.4*	2.37	*1.19-4.69*	1.53	*0.73-3.21*
	IIIA	388	*11.7*	16	*4.1*	2.87	*1.39-5.94*	1.93	*0.89-4.23*
							p=0.0020		***p=0.1093***
Side	Right	1,841	*55.5*	43	*2.3*	1.00			
	Left	1,365	*41.1*	32	*2.3*	1.00	*0.63-1.59*		
	Other	18	*0.5*	1	*5.6*	2.46	*0.32-18.90*		
	Missing	95	*2.9*	6	*6.3*	2.82	*1.17-6.80*		
							p=0.1820		
Histology (Pre-op)	Adenocarcinoma	1,155	*34.8*	16	*1.4*	1.00		1.00	
	Squamous cell	1,151	*34.7*	43	*3.7*	2.76	*1.55-4.93*	1.54	*0.84-2.84*
	NSCLC NOS	662	*19.9*	13	*2.0*	1.43	*0.68-2.98*	0.95	*0.45-2.04*
	Other	82	*2.5*	2	*2.4*	1.78	*0.40-7.88*	1.61	*0.35-7.30*
	Missing	269	*8.1*	8	*3.0*	2.18	*0.92-4.15*	1.93	*0.80-4.66*
							p=0.0063		***p=0.3348***
Procedure	Segmentectomy/wedge	441	*13.3*	7	*1.6*	0.88	*0.39-1.97*	0.76	*0.33-1.74*
	Lobectomy	2,275	*68.5*	41	*1.8*	1.00		1.00	
	Bi-lobectomy	131	*3.9*	6	*4.6*	2.62	*1.09-6.28*	2.33	*0.93-5.81*
	Pneumonectomy	335	*10.1*	25	*7.5*	4.39	*2.64-7.33*	4.07	*2.28-7.24*
	Other	137	*4.1*	3	*2.2*	1.22	*0.37-3.99*	0.92	*0.28-3.07*
							p<0.0001		***p=0.0001***

OR odds ratio; CI confidence interval; PS performance status; FEV_1 forced expiratory volume in 1 second. *ORs are adjusted for all other factors for which adjusted ORs are given. All p values calculated using likelihood ratio test; ** LRT p for trend.

Post-hoc analysis: Reference tables

Given the important effects of procedure type, age, and performance status (both in terms of significance and size of effect in the univariate model), the percentage of patients who died within 90 days of surgery was calculated stratified by procedure type, age, and performance status. Figure 6-11 shows these results as a simple cross tabulation, which could be used as a reference table for clinicians, giving them quick and easy access to data which could help them estimate of a similar patient's risk of dying within 90 days of surgery.

	Performance status					
Age	**0**	**1**	**2**	**0**	**1**	**2**
<70	**1%** (1-2) *1504*	**4%** (2-5) *907*	**8%** (3-12) *145*	**8%** (5-12) *296*	**12%** (7-16) *197*	**7%** (-3-17) *29*
70-80	**4%** (3-6) *777*	**7%** (5-8) *780*	**10%** (5-15) *120*	**20%** (12-28) *95*	**14%** (6-21) *87*	**24%** (1-46) *17*
>80	**8%** (3-12) *144*	**7%** (3-11) *200*	**23%** (7-39) *30*	**25%** (-14-64) *8*	**19%** (-3-40) *16*	**0**[#] - *0*
	LOBECTOMY			***PNEUMONECTOMY***		

95% confidence intervals shown in parenthesis; italics show total number of patients in each category; #No deaths recorded

Figure 6-9: Proportions of patients who died within 90 days of surgery for NSCLC

6.3.4 Discussion

For people undergoing surgical resection with curative intent for NSCLC, post-operative mortality was 3.0% within 30 days and 5.9% within 90 days. Patient demographic, co-morbidity and tumour-related features were similar in those who died within 30 days and those who died 31-90 days after surgery. Increasing age, performance status and procedure type were strongly associated with an increased risk of early post-operative death.

Strengths

The major strengths of this study are the large sample size, the nationally representative nature of the lung cancer cases included, and the use of English data to inform UK practice. No previous study has examined risk factors for post-operative mortality in NSCLC based on a UK population or UK surgical practice.

The definition of surgery was carefully considered prior to conducting the analyses and the decision was made to use surgical procedures recorded in HES rather than those recorded in the NLCA. All NHS trusts submit data to HES through clinical coding and since 2006 this has been used to generate tariffs for hospital services. It is unlikely, therefore, that procedures would take place without being coded in HES, and the accuracy of clinical coding in the NHS is audited annually.(91)

Limitations

The main weakness of this study is the amount of missing data. The variable with the greatest proportion of missing data was lung function (percentage predicted FEV1 was missing in 63% of cases) and it is possible that with more data on lung function this may prove to be a better predictor of mortality than performance status or co-morbidity score; this would be important as it varies more widely and is more objectively measured than performance status. All variables were analysed using a separate category for missing data, rather than

imputing figures, to ensure that the results were a true representation of the available data. The missing data must also be taken in context, and given the size of the study it was still possible to analyse data on lung function for over 4,500 individuals.

When using HES data to calculate Charlson index a score of zero was assigned to any individual who either had no records of hospital admission in the HES database, or had no record of an ICD-10 code relating to any of the diagnoses in the Charlson index.(90) This method may have missed diagnoses which were only recorded in primary care, however 95% of patients in this study had at least one complete inpatient episode prior to their procedure date and all relevant or major co-morbidities (all Charlson co-morbidities are major co-morbidities) should be recorded in each episode, particularly since the introduction of payment by results for NHS hospitals in 2006.

It was not possible to tell at which point in the patient pathway information on patient fitness was entered into the NLCA database and therefore we cannot be sure that performance status in particular reflects that of the patient at the time of surgery. The NLCA team suggest that clinicians or administrators enter the data at the time the patient is first discussed at an MDT meeting, but it is possible to update and replace entries after this point. It would, however, be unusual for a treatment with curative intent to take place a long time after a diagnosis of lung cancer was made and therefore we expect the majority of these data to reflect the patient's condition at the time of surgery.

Additional data

Whilst ethnicity is recorded in HES, the majority of patients in the English NLCA are white and therefore there were insufficient data to assess whether ethnicity affected the outcome. This must be considered when applying these results to ethnic minorities in clinical practice.

Data on trust factors such as case-load, number of specialist thoracic surgeons or number of intensive care beds were not available for this study. It would be interesting and important for future service provision to understand the effect of these factors on post-operative mortality.

This study is based on data from 2004 to April 2010. This is more recent than any previous study however surgical practice is changing with more video-assisted thoracic surgery (VATS) and minimally invasive procedures. (195, 196) In the future there is likely to be a need for a study investigating whether the factors which affect mortality following these procedures differ from the open procedures which currently dominate practice.

Previous studies

There are few previous reports of 90-day mortality following surgery for lung cancer. One Dutch group reported 3.9% 30-day and 6.8% 90-day mortality after lobectomy, bi-lobectomy, or pneumonectomy between 2000 and 2008.(197) These figures are very comparable to the results of the current study.

Almost as many people died between 31 and 90 days as in the first 30 days after lung cancer resection in our study. Many previous studies included death prior to hospital discharge in their definition of 30-day mortality to account for people who were alive longer than 30 days due to improvements in perioperative management and intensive care. Discharge practices, however, vary between hospitals, and there remain a substantial proportion of patients who die after discharge but before 90 days.(198) One of the important findings from this study is that the features of patients who die within the first 30 days of surgery are no different to those of patients who die between 31 and 90 days post-operatively. Since post-operative recovery takes several months, perhaps patients should be provided with an estimate of their risk of death within 3 months, instead or as well as within 1 month.

Clinical relevance

It is important that clinicians are aware of the factors (age, procedure type and performance status) which have the greatest effect on risk of early postoperative mortality, and that they provide their patients with sufficient information about risk for them to make an informed choice about treatment.

Using the proportion of people who died within 90 days of surgery two simple reference tables were developed which display the percentage of patients who died within this post-operative period according to age, performance status and the necessary procedure type. These are designed to provide clinicians with easy access to UK data on which an estimation of a patient's mortality risk can be based. These tables were developed using data on considerably more procedures which were performed specifically for lung cancer, and considerably more deaths, than previous studies,(181) however estimates for risk in subgroups with small numbers of procedures (the high risk categories) must be interpreted with caution. The population based data in these tables must also be considered in context if they are to be used to assist in the estimation of risk for an individual patient; co-morbidities, pre-operative lung function and stage also have an effect even after adjusting for performance status, and women have a slightly lower risk of early mortality than men.

With the wide availability of computers and portable technology such as smart phones, a more complex risk prediction tool could be used in the clinic setting with almost as much ease as these reference tables. In the next section I will therefore go on to produce a risk prediction model which includes several additional factors and should therefore estimate post-operative risk for an individual patient with a greater degree of accuracy – however this will require validation in an independent dataset.

6.4 Generation of a new risk prediction model

6.4.1 Aims

The aims of this section were to construct a predictive model using the results of multivariate analysis of risk factors for early mortality and to compare the new model with Thoracoscore (*chapter aims 4 and 5*).

6.4.2 Methods

The predictive score comprised the coefficients and constant from the logistic regression model described in section 6.3, restricted to the 3,319 records (169 deaths) with complete data on performance status, stage and lung function (table 6-4). Variables which were significantly associated ($p<0.1$) with 90-day mortality in the multivariate analysis were included in the final score. The calculation of percentage risk was based on the methods used by Falcoz et al. (170)

Ninety-day mortality was chosen because a substantial number of deaths occurred between 31 and 90 days, and although the traditional postoperative period is just 30 days there is evidence that recovery after lung cancer resection takes at least 3 months in the majority of cases.(194) There were no significant differences in the effects of the major contributing factors for death within 31-90 compared with 30 days and there were more deaths within 90-days increasing the power and therefore the accuracy of the model.

The predictive score was compared with Thoracoscore because it is recommended in national guidelines.(25) The coefficients comprising the final model were tabulated with those which make up Thoracoscore and inspected to determine whether there were any major differences in contributing factors or weighting. To demonstrate the function of the score, and to allow further comparison with Thoracoscore, both scores were applied to hypothetical low, moderate, and high risk patients.

6.4.3 Results

Age, procedure type, performance status, stage, Charlson index and were significantly associated with 90-day mortality in the multivariate model. Sex, ethnicity, Townsend quintile, side of surgery and histological subtype were not independently associated with 90-day mortality in this analysis but given the ease of accurately recording sex this variable was retained in the final model.

The coefficients and constant from the multivariate model are presented in Table 6-6. The probability of death within 90 days is calculated in two steps as follows:

1) Odds = exp (total of coefficients + constant)
2) Probability of death = odds / (1+odds)

I will refer to this new risk prediction model as the 'NLCA score' for the purpose of comparison with Thoracoscore and in the discussion.

Table 6-6: Coefficients from NLCA score and Thoracoscore

		Coefficient	
		Thoracoscore (170)	**NLCA data**
	Definition of early mortality	30 days or in-hospital	90 days
	Number of deaths	218	169
Age (years)	<55	-	-
	55-65	0.77	0.31
	>65	1.01	
	66-75		0.97
	>75		1.40
Sex	Female	-	-
	Male	0.45	0.23
ASA score	≤2	0	
	≥3	0.61	
Performance	0		-
status	≤2	-	
	1-2		0.68
	≥3	0.69	0.21‡
MRC dyspnoea	≤2	-	
score	≥3	0.91	
% predicted	>80%		-
FEV_1	61-80%		0.20
	40-60%		0.69
	<40%		0.95
Priority	Elective	-	
	Urgent/emergency	0.84	
Procedure class	Other [a]	-	
	(Bi-)lobectomy, wedge, or segmentectomy		-
	Other [b]		0.07
	Pneumonectomy	1.22	1.16
Diagnosis group	Benign	-	
	Malignant	1.24	
Comorbidity	0	-	
score*	1-2	0.74	
	≥3	0.91	
Charlson	0-1		-
index	≥2		0.33
Stage	1a		-
	1b		0.42
	2a or 2b		0.51
	3a		0.84
	Constant	**-7.37**	**-5.28**

*MRC Medical Research Council, ASA American Society of Anaesthesiologists grade, FEV_1 forced expiratory volume in 1 second; *Number of significant co-morbid conditions including: smoking, history of cancer, chronic obstructive pulmonary disease, diabetes mellitus, arterial hypertension, peripheral vascular disease, obesity and alcoholism; - indicates baseline group; ‡ Only 40 patients and 2 deaths in this group. [a] Other includes mediastinoscopy or other mediastinal surgery, wedge resection, lobectomy or bi-lobectomy. [b] Other includes procedures listed in Appendix E.*

Table 6-7 shows the predicted outcomes using the NLCA score and Thoracoscore for three hypothetical cases. Predicted mortality using the NLCA score was very similar to that of Thoracoscore for the low risk patient, slightly higher for medium risk and almost double for the high risk patient.

Table 6-7: Patient features and predicted outcomes

Description of patient	Estimated risk of death	
LOW RISK: 56 y/o female, MRC 1, ASA 2, PS 0, FEV_1 81%, hypertension, ex-smoker, NSCLC stage 1b, elective lobectomy	90-day mortality (NLCA score) In hospital mortality (Thoracoscore)	1.0% 1.1%
MODERATE RISK: 70 y/o male, MRC 2, ASA 3, PS 1, FEV_1 65%, COPD, hypertension, smoker, NSCLC stage 2b, elective lobectomy	90-day mortality (NLCA score) In hospital mortality (Thoracoscore)	6.4% 4.1%
HIGH RISK: 81 y/o male, MRC 4, ASA 3, PS 2, FEV_1 50%, COPD, ischaemic heart disease, hypertension, ex-smoker, diabetes, NSCLC stage 2b, elective pneumonectomy	90-day mortality (NLCA score) In hospital mortality (Thoracoscore)	43.0% 26.2%

y/o years old, MRC Medical Research Council dyspnoea score, ASA American Society of Anaesthesiologists grade, PS performance status, FEV_1 forced expiratory volume in 1 second, NSCLC Non-small cell lung cancer COPD Chronic obstructive pulmonary disease.

6.4.4 Discussion

This analysis shows that there are sufficient data in the NLCA to produce a model to estimate the risk of death within 90-days after lung cancer surgery. This is the first risk score to consider deaths which occur more than 30 days after surgery and the first to be developed based on a UK lung cancer population and UK surgical practice.

Strengths

The main strengths of the study are the large sample size and the representative nature of the study population. The data on lung function are particularly important as this is an objective measure of patient fitness in contrast to performance status, MRC dyspnoea score and ASA grade, all of which are subject to different interpretations and assignment by different clinicians.

Limitations

The main limitation of this study was the absence of a suitable independent dataset in which to test the performance of the score. Although it has not been validated as part of this study this will be possible in the future and is discussed further in Chapter 9 (section 9.2.2).

The components of the model were restricted to the data that were available in the NLCA and HES at the time of the study. There were insufficient data to calculate predicted postoperative lung function which was found to predict early mortality in the ESOS study,(181) and the ASA and MRC dyspnoea scores are not recorded in the NLCA. It is possible that these variables would have been significantly associated with 90-day mortality in the NLCA population since they were associated with in-hospital mortality in previous studies performed elsewhere in Europe.(170, 185) Smoking data are not available in the NLCA or HES data and it has been suggested that continued smoking may be associated with adverse outcomes after lung cancer surgery.(199)

Comparison with Thoracoscore

Most of the discrepancies between mortality risk estimates using the NLCA score and Thoracoscore (Table 6-7) are likely to be due to the longer time period in this study (deaths within 90 days compared with deaths in-hospital), and the differences in populations studied. The overall in-hospital mortality in the data on which Thoracoscore was based was 2.2%, compared with 5.1% 90-day mortality in this study. Thoracoscore,(170) was based on a larger number of procedures than this NLCA score, and a larger number of deaths (218 in-hospital deaths), but even though it is currently used by many clinicians in the UK to estimate perioperative mortality risks for patients with NSCLC, it was not restricted to patients with lung cancer and has not been validated in this population.

The other main differences between the two scores are the weighting and categories used for age, performance status and co-morbidity and the inclusion of ASA grade and MRC dyspnoea score, but not stage or lung function in Thoracoscore. Whilst data on both stage and lung function were available to the authors of Thoracoscore they did not find them to be significantly associated with in-hospital mortality, perhaps due to the better overall health of their cohort of patients.

Age was the most important factor in terms of size of effect on postoperative mortality in the current study. In the UK, almost three quarters of patients diagnosed with lung cancer between 2004 and 2010 were over 65 years of age,(55) and a recent study suggests that the number of older patients undergoing surgical resection for lung cancer is increasing.(76) The mean age of patients who had potentially curative resection for NSCLC in the NLCA database was 67 years and the effect of age appears to be most important over the age of 65 years. Thoracoscore uses 3 age categories: less than 55, 55-64 and greater than 64 years, and was developed using a dataset in which the mean age at operation was 54.7 years thus potentially underestimating the mortality in the older age group.

Neither score has been validated in a large study of lung cancer patients, and prospective evaluation needs to be undertaken as a matter of priority. If the NLCA score proves to be a useful clinical tool, care must be taken when applying it to people of non-white ethnicity given that they contributed such a small proportion of our study data (this is also likely to be the case for Thoracoscore although ethnicity is not reported in the publication, (170)).

Clinical relevance

The estimation of post-operative mortality risk is a crucial part of management of patients with NSCLC. In order to ensure that as many patients as possible are offered, and consider having, potentially life-saving surgery, estimates of risk

must be based on the best available evidence. It is also important to be aware of the patient's concerns and expectations regarding post-operative morbidity and mortality.

The use of early deaths after lung cancer surgery as a measure of performance of individual thoracic surgeons, which may be introduced for revalidation, has raised concerns over risk-averse patient selection by surgeons. It is clearly important that if mortality figures are to be used in this way they are adjusted for the factors which are strongly associated with early post-operative death and ideally fully adjusted. There is potential to use the NLCA risk prediction model to adjust mortality figures so that surgeons are not any less inclined to offer surgery to a patient of borderline fitness.

In addition to surgery, new techniques including stereotactic radiotherapy and radio-frequency ablation are starting to become available for the treatment of early stage lung cancer. These treatments are less invasive and it is thought that they provide the possibility of cure in patients for whom surgery may either have extremely high risk or for patients who are unwilling to accept the level of risk associated with whatever procedure they would require. A tool which estimates the risk of 90-day mortality following surgery in patients who are treated with stereotactic radiotherapy could prove valuable in comparing the observed outcome with the predicted outcome from surgery.

6.5 Chapter summary

In this chapter I have described the use of the linked NLCA-HES-ONS data to assess risk factors for early death after lung cancer resection and to develop a risk prediction model for use in clinical practice.

This work was published in *Thorax* in May 2013, (200) accompanied by an editorial on the topic of operative risk in lung cancer. (201) I also presented some of the data at the 11th Annual British Thoracic Oncology Group conference in January 2013, at the East Midlands Cancer Network meeting in November 2012, and as a poster abstract at the British Thoracic Society Winter Meeting in December 2012.

The next phase of this research should include a validation study using an independent dataset and I will discuss this further in Chapter 9. In the next two chapters I will use the same linked dataset and similar statistical techniques to investigate treatment decisions and survival for people with small cell lung cancer.

CHAPTER 7: VALIDATION OF RECORDS OF CHEMOTHERAPY AND RADIOTHERAPY

This chapter describes a validation study in which records of chemotherapy for small cell lung cancer in the HES and NLCA databases are compared with the aim of determining the most appropriate definition of chemotherapy for future studies. This is followed by an exploration of the data on radiotherapy in HES and the NLCA.

7.1 Introduction

7.1.1 Background

Small cell lung cancer (SCLC), as described in Chapter 1, makes up a relatively small proportion of the overall lung cancer burden in the UK and is declining in incidence as the prevalence of cigarette smoking declines. It remains, however, an aggressive disease which is often advanced at presentation and often rapidly fatal, despite treatment.

The NLCA has been used to investigate whether there are inequalities in care for people with SCLC, in particular in the receipt of chemotherapy (which is the mainstay of treatment for this disease).(83) Both the NLCA and HES contain data on treatment with chemotherapy, but to date these have not been validated or compared. The NLCA - HES linkage for the purpose of generating a measure of co-morbidity means it is now possible to compare records between the two databases and to analyse patient features and outcomes in order to assess the validity of treatment records in each database.

7.1.2 Rationale for this study

It is important to determine the most accurate means of identifying whether or not a patient was treated so that future studies are consistent in their methods and are not affected by errors in data entry or recording bias.

In Chapter 8 I will describe a study investigating which factors which were associated with increased likelihood of being treated with chemotherapy for SCLC and the effect that this had on survival. It was therefore important to use a variable which accurately identified people who received chemotherapy and those who did not.

7.1.3 Aim of this chapter

The aim of the main study described in this chapter was to assess the validity of records of chemotherapy for SCLC in the HES and NLCA databases in order to agree a definition for future studies. This was done by:

1. *Identifying patients with SCLC who had a record of chemotherapy in the NLCA, or in the linked HES data, or in both (section 7.2);*
2. *Examining and comparing the features of these patients (including survival) according to the database in which chemotherapy was recorded (section 7.2);*

7.1.4 Radiotherapy records

Radiotherapy is also an important part of the treatment for some people with SCLC. As part of this study I also attempted to compare records of radiotherapy in the NLCA and HES (section 7.3), however, as will be explained, the majority of radiotherapy records are not captured in the inpatient HES data and therefore it was not possible to assess the completeness of recording in the NLCA.

7.2 Records of chemotherapy in HES and the NLCA

7.2.1 Methods

Study population

The July 2013 NLCA-HES extract (see section 2.2.4) was used for this study. This included patients first seen between 1st January 2004 and 31st December 2011. In this extract HES data were available up to 31st March 2012 and although new patient records added to the NLCA after the end of 2011 were not included, data were entered for the existing patients up to June 2012.

To ensure each patient had a post-diagnosis follow-up period of at least 3 months, during which chemotherapy could have been given and recorded in both datasets, those diagnosed (or with a start date) after 31st December 2011 were excluded. Patients diagnosed before 2004, those for whom it was not possible to calculate a start date, and any records with a date of death on or before their start date were also excluded.

To allow fair comparison between HES and the NLCA, any cases with a record of chemotherapy in the NLCA after 31st March 2012 were excluded from this analysis.

Cases with histologically confirmed SCLC were identified using pre-treatment histology where this was available and post-treatment histology where the pre-treatment variable was missing.

Covariates

Histology, stage and performance status were defined as described in section 2.2.5. Age refers to the NLCA variable age at time of diagnosis. Route of referral was also obtained from the NLCA.

NLCA records of chemotherapy

Chemotherapy is recorded in the NLCA in four fields:

- the date on which it was decided that the patient should receive chemotherapy treatment,
- the date on which the first dose of chemotherapy was administered,
- the code for the hospital or trust where the patient received chemotherapy, and
- the reason that chemotherapy was given.

The reason for chemotherapy can be: chemotherapy alone, chemotherapy combined with radiotherapy, adjuvant chemotherapy post-surgery, or induction chemotherapy to downstage prior to surgery. There is only a date for the first and not subsequent doses, and there is no information regarding the number of cycles or which drugs were given.

HES records of chemotherapy

In the HES database there are two ICD-10 codes which relate to chemotherapy administration: Z51.1 (chemotherapy for neoplasm) and Z51.2 (other chemotherapy). There are also several OPCS-4 codes relating to chemotherapy procurement and delivery. The ICD-10 code should be recorded as a diagnosis and the OPCS-4 code should be recorded as a procedure for each episode where chemotherapy was given.

Following discussion with clinical coding staff at Nottingham University Hospitals NHS Trust, both the chemotherapy delivery and procurement OPCS-4 codes, as listed in Appendix F, but not the ICD-10 codes, were used to identify cases that had chemotherapy. Chemotherapy coding follows national guidelines and Nottingham University Hospitals follow a flow chart which is based on these guidelines (also shown in Appendix F).

Dates

Chemotherapy regimens for other cancers which were given before or after the diagnosis of lung cancer might have been found in HES but would not be recorded in the NLCA. Therefore HES episodes were excluded if they were more than 3 months before or 6 months after the NLCA start date. For continuity and to remove any possible errors in data entry, NLCA chemotherapy dates, trusts and reasons were deleted (re-coded to missing) if the date was outside this range.

Given that there may be several records of chemotherapy administration per patient in HES, after the above exclusions, the earliest date was considered to be the date of first dose (later dates were assumed to reflect subsequent doses or cycles).

Statistical methods

Records were grouped according to whether chemotherapy was recorded in both databases, in neither database or in one database only (as shown in Table 7-1). A Venn diagram was constructed to show the overlap between these groups.

NLCA only records were sub-divided into those with a chemotherapy date with or without trust and/or reason, trust with or without reason but no date, or reason but no trust or date. People with a record of chemotherapy in HES and any one of a date of first dose, trust of administration or reason for chemotherapy were considered to have chemotherapy recorded in both datasets (groups 1, 1a and 1b).

For each of groups 1-5 the average age and median survival from date of diagnosis was calculated. Kaplan Meier survival curves were plotted and examined for similarities and differences between the groups that might provide an insight into the validity of records in each dataset. The distributions of performance status, stage, and source of referral were tabulated for each of the

groups, as was the distribution of chemotherapy records according to year of diagnosis and region (cancer network) where the patient was first seen.

Table 7-1: Groups according to where records of chemotherapy were found

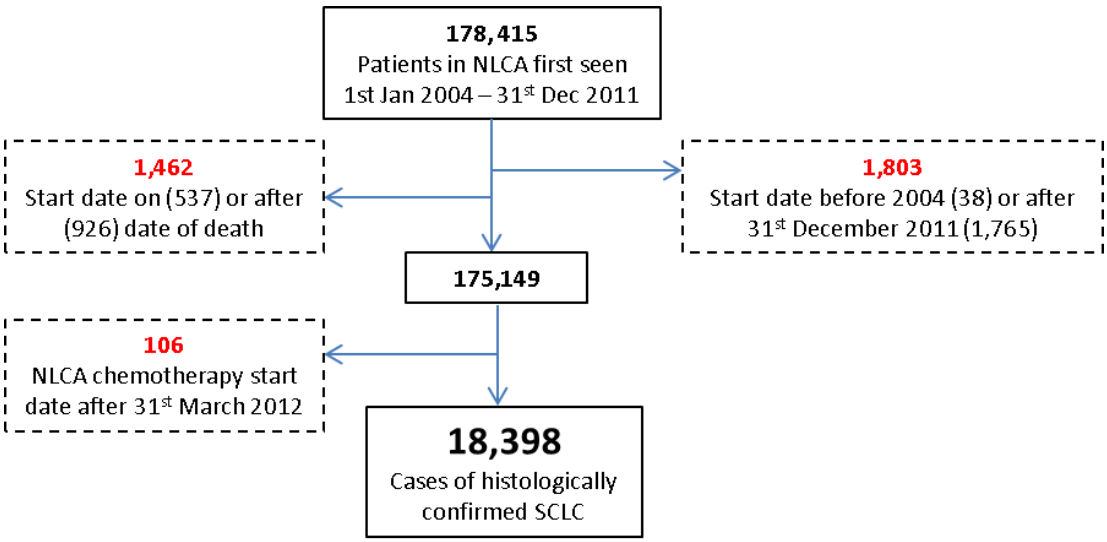
Group	Criteria
1=Both	Date of chemotherapy in HES AND Date of chemotherapy in NLCA
1a	Date of chemotherapy in HES AND Trust of chemotherapy in NLCA NO date of chemotherapy in NLCA
1b	Date of chemotherapy in HES AND Reason for chemotherapy in NLCA NO date or trust of chemotherapy in NLCA
2=HES only	Date of chemotherapy in HES NO reference to chemotherapy in NLCA
3=NLCA only (date)	Date of chemotherapy in NLCA NO reference to chemotherapy in HES
4=NLCA only (trust or reason)	Procedure type and/ or trust of chemotherapy in NLCA NO date of chemotherapy in NLCA NO reference to chemotherapy in HES
5=Neither	NO reference to chemotherapy in either database

Final definition of chemotherapy

As with the analysis of records of surgery (Chapter 5), patients with a record of having had chemotherapy in both HES and the NLCA were considered very likely to have actually had chemotherapy, and those without a record in either were considered unlikely to have received chemotherapy. Cases with records of chemotherapy in both datasets (group 1) and those in any other group(s) with similar features to this group will therefore be used to define receipt of chemotherapy for future studies.

7.2.2 Results

There were 178,415 records in the NLCA database with a date first seen between 1st January 2004 and 31st December 2011. After excluding 1,803 patients who were diagnosed outside the study period and 1,462 with start dates on or after the recorded date of death, 175,149 remained. One hundred and six cases had a record of chemotherapy in the NLCA which started after 31st March 2012 and of the remaining records 18,398 had histologically confirmed SCLC. The exclusions and derivation of the study population are shown in figure 7-1.



NLCA National Lung Cancer Audit; SCLC Small cell lung cancer

Figure 7-1: Exclusions and derivation of study population for chemotherapy record validation in SCLC

<u>NLCA records of chemotherapy</u>

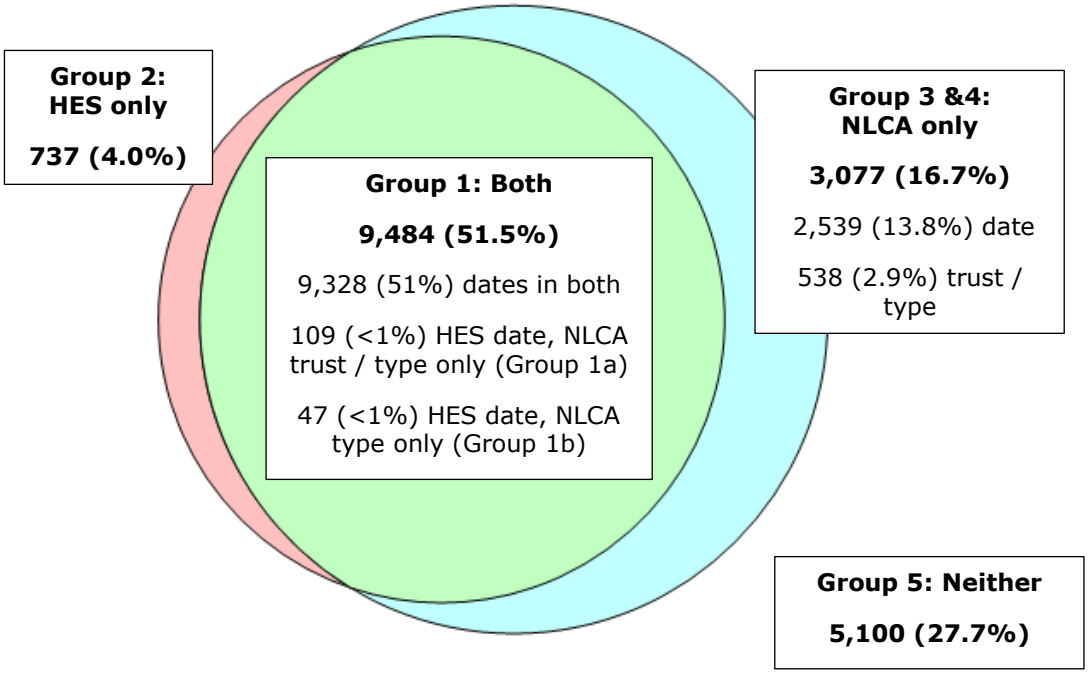
There were 11,867 records with a chemotherapy start date recorded in the NLCA which was less than 3 months before and 6 months after the date of diagnosis. A further 694 had a trust of administration or reason for chemotherapy but no date of first dose.

HES records of chemotherapy

There were 10,221 records which contained at least one of the OPCS-4 codes for chemotherapy (Appendix F), dated between 1st 2004 and March 31st 2010, and less than 3 months before / 6 months after the NLCA start date.

Comparison of databases

Chemotherapy was recorded in both databases in 9,484 (51.5%) of the 18,398 cases with SCLC; 5,100 (27.7%) had no record of chemotherapy in either database. Figure 7-2 shows the distributions of records of chemotherapy for SCLC in HES and the NLCA, and where these overlapped.



NLCA National Lung Cancer Audit; HES Hospital Episodes Statistics; percentages indicate proportion of overall small cell lung cancer population, N=18,398

Figure 7-2: Venn diagram depicting the overlap between records of chemotherapy in HES and the NLCA

Date of first dose

Of the 9,328 cases where the date of first chemotherapy dose was evident in both HES and the NLCA the two dates were exactly the same in 6,925 (74%)

cases. Of the remaining 2,403, in 1,652 (69%) cases the HES date was the later of the two dates. In 8,099 cases (87% of the overall number) the dates were within 1 week of each other.

Patient features

Overall the features of patients with chemotherapy recorded in HES only or the NLCA only (groups 2 and 3) were similar to those of patients with records in both databases (group 1). Patients in group 5 (who had no record of chemotherapy) and group 4 (who had a trust or reason for chemotherapy recorded in the NLCA but no date) were older than those in groups 1-3. The mean age was similar across groups 1-3 (Table 7-2).

There was a much higher proportion of patients with extensive stage and also a higher proportion with poor performance status (3-4) in groups 4 and 5 than in groups 1-3. A lower proportion of patients in groups 4 and 5 we referred by their general practitioner compared with groups 1-3. Performance status and stage were missing in a higher proportion (37.3% and 33.5% respectively) of those with chemotherapy records in HES only than in any other group (Table 7-2).

Table 7-2: Features of patients with small cell lung cancer according to where chemotherapy was recorded

	N=18,398	Record of chemotherapy				
		Both **Group 1** n=9,484	**HES only** **Group 2** n=737	**NLCA only (date)** **Group 3** n=2,539	**NLCA only (no date)** **Group 4** n=538	**Neither** **Group 5** n=5,100
Mean age (years)		66.3	66.1	67.1	70.9	72.8
Stage (% of non-missing)	**Limited**	37.6	39.4	36.8	24.2	21.5
	Extensive	62.4	60.6	63.2	75.8	78.5
Missing stage (% of total)		17.5	33.5	22.5	29.4	22.4
Performance status (% of non-missing)	**0-1**	67.2	63.7	60.9	36.1	24.8
	2	22.9	23.1	28.8	30.4	27.2
	3-4	9.8	15.2	10.3	33.5	48.0
Missing performance status (% of total)		17.3	35.7	18.7	27.3	23.6
Source of referral: (% of non-missing)	**Emergency admission**	13.9	13.7	12.3	17.4	22.8
	General Practitioner referral	56.2	53.5	54.9	37.6	38.7
	Consultant referral	18.6	17.9	19.6	24.7	20.1
	Other	5.4	9.0	6.5	8.1	6.4
	Emergency department	6.0	5.9	6.7	12.3	11.1
Missing source of referral (% of total)		4.4	8.0	5.4	8.0	4.9
Median survival after diagnosis (days) (IQR)*		268 (152-434)	242 (106-415)	230 (83-391)	36 (17-111)	32 (14-79)

NLCA National Lung Cancer Audit; HES Hospital Episodes Statistics; Groups 1-5 defined in text and table 7-1; IQR Interquartile range. *Survival is calculated from start date or date of diagnosis ; **Date of first chemotherapy dose as recorded in HES unless NLCA only.

Overall survival

Figure 7-3 shows the overall survival for patients with SCLC according to where chemotherapy treatment was recorded. Median survival for patients in groups 4 and 5 was similar (36 and 32 days respectively) and considerably shorter than that of patients in groups 1-3 (Table 7-2).

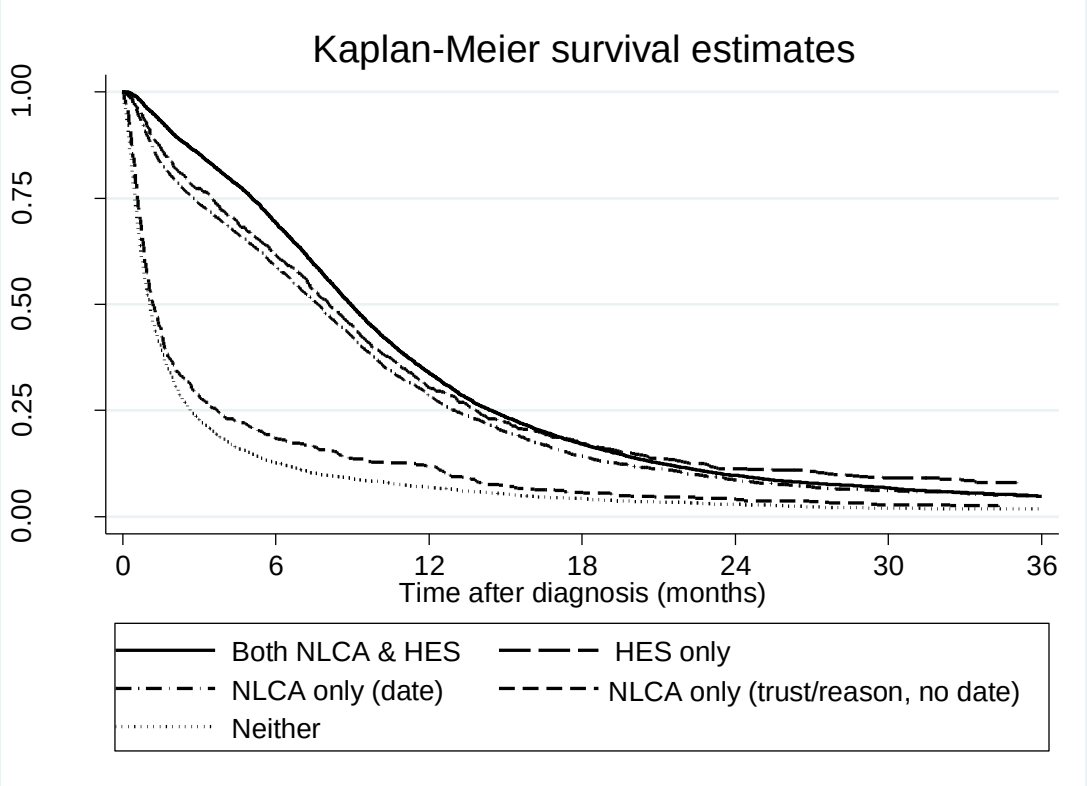


Figure 7-3: Survival after diagnosis by to chemotherapy records

Chemotherapy records by year and network

The proportion of patients with a record of chemotherapy in both datasets increased over time from 28% in 2004 to 57% in 2011. The proportion in HES only decreased over the same period from 6% to 3% but the proportion in the NLCA only decreased between 2004 and 2008 (31.5% to 11.6%) then stayed at this level between 2008 and 2011 (Table 7-3).

Table 7-3: Distribution of records of chemotherapy in people with small cell lung cancer by year of diagnosis

Year of Diagnosis (no. of patients)	**Record of chemotherapy (% of year total)**				
	Both databases Group 1	**HES only** Group 2	**NLCA only (date)** Group 3	**NLCA only (no date)** Group 4	**Neither** Group 5
2004 (537)	27.6	6.2	31.5	6.5	28.3
2005 (1,311)	35.7	5.0	24.3	10.8	24.2
2006 (1,819)	46.4	4.8	17.5	9.0	22.2
2007 (2,069)	54.0	5.3	12.1	3.3	25.3
2008 (2,690)	53.4	4.8	11.6	1.3	28.9
2009 (3,199)	53.7	3.1	11.7	0.8	30.7
2010 (3,297)	54.1	3.3	11.8	1.1	29.8
2011 (3,476)	56.7	3.0	11.7	0.9	27.7

NLCA National Lung Cancer Audit; HES Hospital Episodes Statistics
Groups 1-5 defined in text and table 7-1

Cancer networks were described in section 1.4. When the distribution of chemotherapy records was analysed by network first seen there were two outliers; these are highlighted in table 7-4. In network 3, an unusually low proportion (8.7%) of cases had chemotherapy records in both databases; however 51% had a record in the NLCA only which was higher than any other network. The proportion with no record of chemotherapy for this network (31%) was similar to the overall figure (28%). For patients first seen in network 21, there was also a low proportion with records of chemotherapy in both databases (12%) but a high proportion (55%) did not have a record of chemotherapy in either database.

Table 7-4: Distribution of records of chemotherapy in HES and the NLCA by year of diagnosis

Network first Seen* (No of patients)	Record of chemotherapy (% of network total)				
	Both	**HES only**	**NLCA only (date)**	**NLCA only (no date)**	**Neither**
	Group 1	Group 2	Group 3	Group 4	Group 5
1 (632)	65.7	2.5	6.7	-	25.2
2 (1,192)	34.1	5.7	23.9	8.4	27.9
3 (1,077)	8.7	0.4	51.2	8.6	31.1
4 (1,417)	55.9	2.3	16.5	1.6	23.7
5 (535)	64.7	3.4	6.5	5.2	20.2
6 (912)	65.7	7.0	4.2	2.5	20.6
7 (823)	48.0	3.5	17.1	2.1	29.3
8 (365)	55.3	8.2	7.1	2.7	26.6
9 (428)	34.3	3.6	22.6	1.2	38.3
10 (331)	48.6	6.0	8.2	3.0	34.1
11 (421)	66.0	2.9	5.2	4.0	21.9
12 (396)	58.6	4.6	11.1	6.8	18.9
13 (512)	60.2	9.8	2.7	3.1	24.2
14 (319)	59.6	6.9	2.2	0.6	30.7
15 (781)	58.3	3.3	9.7	-	28.7
16 (230)	64.8	3.5	4.8	-	27.0
17 (455)	50.1	5.7	19.8	0.2	24.2
18 (263)	60.1	3.4	1.1	7.2	28.1
19 (623)	46.2	4.7	11.1	0.2	37.9
20 (558)	54.5	2.2	18.1	1.4	23.8
21 (244)	12.3	-	31.6	0.8	55.3
22 (464)	42.9	1.5	16.6	5.4	33.6
23 (284)	51.4	7.0	9.5	2.1	29.9
24 (496)	51.4	5.0	10.5	0.2	32.9
25 (2,107)	60.1	2.8	6.9	2.4	27.8
26 (700)	53.3	6.4	7.3	2.0	31.0
27 (582)	46.7	2.4	15.6	4.1	31.1
28 (1,427)	59.9	4.5	10.2	1.3	24.1
Other (4)	50.0	-	-	-	50.0

**Network names replaced with numbers to preserve anonymity; NLCA National Lung Cancer Audit; HES Hospital Episodes Statistics; Groups 1-5 defined in text and table 7-1*

7.2.3 Interpretation and definitions

Cases which did not receive chemotherapy

The features of people with SCLC who had a trust and / or a reason but no date for chemotherapy coded in the NLCA (group 4) were very similar to those of

people who did not have any reference to chemotherapy in either database. This suggests that these people may not, have had chemotherapy and that the trust or reason may have been entered in error or not updated when treatment plans changed. These people were therefore considered not to have received chemotherapy for the rest of the work in this thesis.

Cases which received chemotherapy

The features of people who had a start date for chemotherapy in the NLCA only or a code (and date) for chemotherapy in HES only, had similar age, lung function and performance status to those who had records in both databases. Median survival was also similar, although slightly worse for those who only had chemotherapy recorded in the NLCA compared with those who had it recorded in both datasets. People will therefore be considered to have had chemotherapy treatment if a chemotherapy date was recorded in either database.

Date of first dose

Where chemotherapy is only recorded in one database that date will be used; where it is recorded in both databases the NLCA date will be used because in most cases the dates are the same, and in cases where they are different the NLCA date is usually the earlier of the two.

Trusts and networks

The proportion of patients without a record of chemotherapy in either database was unusually high in network 21. Reasons for this are unknown but may include a high proportion of patients being treated in the private sector and/or a policy of giving more chemotherapy in outpatients. Patients first seen at any of the trusts in this cancer network will be excluded from the analysis for the work in Chapter 8.

7.3 Radiotherapy in HES and the NLCA

7.3.1 Background

The use of radiotherapy in the treatment of SCLC is described in detail in section 1.3.2 but briefly it can be used in two ways: The first is combined with chemotherapy either at the same time or one after the other, in a relatively high dose over several sessions. This is termed chemo-radiotherapy or radical radiotherapy and the aim is to reduce the disease burden and prolong the patient's life; it is used for patients with limited stage disease where all of the tumour can be captured in one radiotherapy field.

Secondly radiotherapy can be used to treat or control symptoms which are directly related to the tumour. For example, a tumour in the lungs may cause the patient to cough up blood, or a tumour which has spread to the bones may be painful. This type of radiotherapy tends to be given at a lower dose and often only one or two sessions; this is termed palliative radiotherapy and is not intended to prolong the patient's life.

Chemo-radiotherapy has been shown to improve survival in clinical trials compared with chemotherapy alone, particularly when the two treatments are given at the same time.(38, 202, 203) In studies of chemotherapy it is therefore important to understand whether or not radiotherapy was also given. The quality and extent of radiotherapy data in the NLCA and HES databases is currently not known.

The aim of this study was

1. *to determine the completeness of radiotherapy data in the NLCA*
2. *to establish whether the treatment intent could be reliably determined from NLCA data, and*
3. *to compare NLCA radiotherapy records with those in the HES database.*

7.3.2 Methods

Study population

The same SCLC population was used for this study as in the chemotherapy study described in section 7.2. Patients with a record of radiotherapy (rather than chemotherapy) given before 2004 or after 31st March 2012 were excluded.

NLCA records of radiotherapy

Radiotherapy is recorded in the NLCA in several fields:

- date of decision to treat with radiotherapy,
- date radiotherapy was started,
- trust where radiotherapy was given,
- anatomical site of radiotherapy (trachea, lung, mediastinum, chest wall, brain, bone, skin, or other), and
- type or intention of radiotherapy (radical, CHART, chemo-radiotherapy, adjuvant or palliative).

There is only a date for the first radiotherapy dose and not subsequent doses (or fractions). Patients who respond to first line chemotherapy (with or without radiotherapy) may be offered prophylactic cranial radiation (PCI). This is intended to reduce the incidence of brain metastases and consequently improve survival.(14, 47, 48) In the NLCA there is a separate field for PCI and therefore any radiotherapy with anatomical site recorded as 'brain' in this field is likely to represent palliative radiotherapy to existing brain metastases.

HES records of radiotherapy

In HES there is a code for radiotherapy administration (Z50.0) but this would only be recorded in the inpatient HES database for those who had radiotherapy as an inpatient or those who attended a day-case unit; radiotherapy is usually administered on an outpatient basis. There are also OPCS-4 codes for radiotherapy delivery and planning. These are shown in Appendix F but again

would only be found in the HES data if radiotherapy was administered as an inpatient.

Following a discussion with clinical coders at Nottingham University Hospitals NHS Trust, OPCS-4 codes (and not ICD-10 codes) were used to identify radiotherapy in HES. The same methods were used as for the chemotherapy study described above to exclude any codes which were dated more than three months before or 6 months after the diagnosis of lung cancer and those which were prior to 2004.

Statistical methods

Records of radiotherapy in the NLCA were identified and the proportion of these for which the treatment intent and anatomical site was clear was calculated. Records were grouped according to whether radiotherapy was recorded in both databases, in neither database or in one database only (as shown in Table 7-2 above) and the intention was to compare features of patients in each of these groups.

7.3.3 Results

The study population was based on the exclusions shown in figure 7-1, excluding 375 individuals with records of radiotherapy given before 2004 (n=1) or after 31st March 2012 (n=374), rather than the 106 who received chemotherapy outside of this period. A total of 18,324 cases of histologically proven SCLC remained for the analysis.

NLCA records of radiotherapy

There were 4,103 cases of SCLC (22.4%) with a start date for radiotherapy in the NLCA. A further 934 had a radiotherapy trust or type (treatment intention) documented with no date. Of these 5,037 cases, 76% had a record of the type of radiotherapy and 48% had a record of the anatomical site. The distributions of these variables are shown in table 7-4. Eleven per cent of those where the treatment intention was 'radical' and seven per cent of those with treatment intent 'chemo-radiotherapy' had the anatomical site recorded as 'brain'.

Table 7-5: Treatment intention and anatomical site for patients with radiotherapy records in the NLCA

Radiotherapy		**% (N=5,037)**
Type / treatment intention	Radical	5.7
	CHART	0.1
	Chemo-radiotherapy	18.9
	Adjuvant	0.3
	Palliative	50.9
	Missing	24.1
Anatomical site	Brain	10.1
	Lung	37.7
	Missing	41.1
	Other	11.1

CHART Continuous Hyper-fractionated Accelerated Radiotherapy

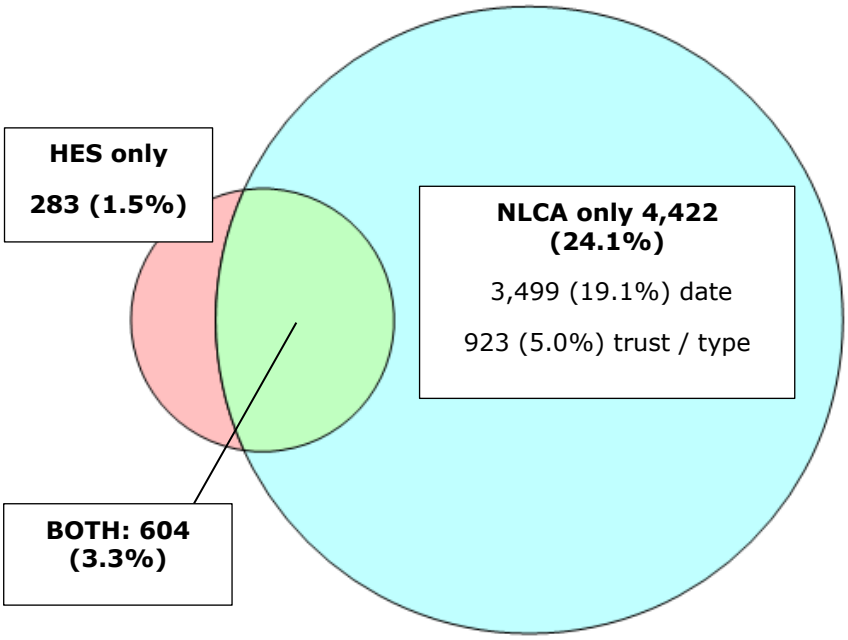
HES records of radiotherapy

There were only 887 records which contained one of the OPCS-4 codes for radiotherapy (Appendix F), dated between 1st 2004 and March 31st 2012, and

less than 3 months before / 6 months after the NLCA start date. An extra 259 records were identified using the ICD-10 codes for radiotherapy however the clinical coders had advised that OPCS-4 codes were more likely to be accurate.

Comparison of databases

The overlap of records of radiotherapy between the two databases is shown in Figure 7-4.



Percentages indicate proportion of overall SCLC population, N=18,324

Figure 7-4 Venn diagram depicting the overlap between records of radiotherapy in HES and the NLCA

7.3.4 Interpretation

Further comparison of patients with radiotherapy recorded in each dataset was not felt to be useful because the majority of patients only had a record of chemotherapy in the NLCA and the likely explanation that this was that radiotherapy was given to most patients on an outpatient basis. The group with radiotherapy recorded in both databases could not, therefore, be used as a comparator group to establish the validity of records in each dataset.

Definition for future studies

The reason for radiotherapy (whether it was with curative intent or for symptom control) is not recorded in HES but is indicated in the NLCA in 76% of cases. Any analysis of survival will be strongly affected by the treatment intention (which determines the dose and number of sessions) and therefore cases in the NLCA with a radiotherapy start date and treatment intent recorded as chemo-radiotherapy or radical radiotherapy will be used to identify people who received chemo-radiotherapy for SCLC in the following chapter. It must be acknowledged that this variable has not been validated and that an unknown proportion of the records with missing treatment intent also had chemo-radiotherapy but will not be captured by this definition.

7.4 Chapter summary

In this chapter I have described and compared the data available on chemotherapy and radiotherapy in HES and the NLCA and used survival analyses to determine that the most accurate means (within these datasets) of identifying people who had chemotherapy for SCLC is the presence of an OPCS-4 code in HES and/or a date of chemotherapy first dose in the NLCA. It was not possible to perform the complete validation study for radiotherapy however a definition of chemo-radiotherapy for SCLC was established. I presented some of this work as a poster abstract at the British Thoracic Society Winter Meeting, London, 2013.(204)

The following chapter uses this definition to perform a detailed analysis of the features of patients with SCLC who received chemotherapy, the number of cycles received and the effects that this treatment has on survival. A description of the patients who received chemotherapy according to the definition established in this is included as part of those results.

CHAPTER 8: TREATMENT DECISIONS AND OUTCOMES IN SMALL CELL LUNG CANCER

In this chapter I will describe a study of the characteristics of people with SCLC who received chemotherapy. I will then describe the use of HES data to determine the number of cycles of chemotherapy that patients were given, and the characteristics of patients who completed a 4-cycle course. This is followed by a survival analysis taking into account patient and trust-level factors, treatment with chemotherapy with or without radiotherapy, and the number of chemotherapy cycles.

8.1 Introduction

8.1.1 Background

The mainstay of treatment for people with small cell lung cancer (SCLC) is chemotherapy. The results of clinical trials show that treatment with a platinum agent combined with etoposide can result in a median survival of 8-12 months for people with extensive stage disease, (35-37) and up to 27 months for those with limited stage disease, particularly when combined with radiotherapy.(35, 38, 39) Since trials tend to include younger patients with relatively good performance status the median survival for the full spectrum of patients is likely to be less.

Prompt investigation, diagnosis and review by an oncologist who can initiate treatment are thought to be important since the rapid tumour growth means that it frequently spreads outside the lung and patients deteriorate quickly in terms of their fitness for treatment.

8.1.2 Rationale for this study

Current chemotherapy agents for SCLC have significant side effects (as described in section 1.3) therefore clinicians must carefully assess patients' fitness prior to chemotherapy and only treat those who are likely to benefit. It is also possible that inequalities in availability of and access to chemotherapy exist for SCLC. It is important to determine whether chemotherapy treatment rates vary by individual patient factors and/or organisational factors and how these are related to survival.

Data from clinical trials suggest that approximately 75% of patients complete the intended number of chemotherapy cycles, (205) but the proportion of patients in routine clinical practice for whom this is the case is not known. An understanding of the characteristics and outcomes for this group of patients

could also help clinicians identify people for whom the risks of starting chemotherapy may outweigh the benefits.

8.1.3 Aims of this chapter

This aim of the study in this chapter was firstly to update work by Dr Anna Rich who used an earlier version of the NLCA-HES database to examine the characteristics of people with SCLC who were treated with chemotherapy.(83) There are now almost double the number of SCLC cases in the database, and a new means of identifying people who had chemotherapy has been defined in this thesis (Chapter 7). I also aimed to use HES data to calculate the number of chemotherapy cycles received, to determine the factors which were associated with completing a course of chemotherapy, and examine survival according to number of cycles received. Finally, the aim was to assess organisational factors including time from diagnosis to treatment and whether the trust where the patient was first seen affected treatment and outcomes.

The specific aims of this chapter were therefore:

1. *to describe the characteristics of patients with SCLC who received chemotherapy (section 8.2);*
2. *to describe the characteristics of patients with SCLC who went on to complete a full course of chemotherapy (section 8.3);*
3. *to use logistic regression to determine the factors which were independently associated with a) receiving chemotherapy (section 8.2) and b) completing a course (section 8.3);*
4. *assess survival according to number of cycles completed (section 8.4); and*
5. *to quantify overall survival according to patient characteristics and the number of cycles completed (section 8.4).*

8.2 Characteristics of patients and factors associated with chemotherapy treatment

8.2.1 Aims

The aims of this section were to describe the characteristics of patients with SCLC who received chemotherapy and to use logistic regression to determine the factors which were independently associated with receiving chemotherapy (*chapter aims 1 and 3a*).

8.2.2 Methods

Study population

The July 2013 extract of the linked NLCA-HES data, which contained NLCA records for patients first seen between 1st January 2004 and 31st December 2011 with linked HES data up to 31st March 2012, was used for this study (further details in Chapter 2).

To give each patient a follow-up period of at least 6 months in both datasets, during which chemotherapy could have been given and recorded, those diagnosed (or with a start date) after 30th September 2011 were excluded. Patients diagnosed before 2006 were also excluded because case ascertainment in 2004 and 2005 was known to be lower than in recent years and the more recent data allowed a sufficiently large sample size even after excluding these records. This study is therefore based on a slightly smaller population than the validation study described in Chapter 7 (Figure 7-1).

Cases of histologically confirmed SCLC were identified and any patients who had a record of surgery were excluded.

Definition of exposure

First-line chemotherapy treatment was assessed by including any chemotherapy given in the first 6 months after a patient's diagnosis. Patients who had

chemotherapy were identified by the presence of a chemotherapy start date in the NLCA database or an OPCS-4 code for chemotherapy (Appendix F) in HES, based on the results of the work in Chapter 7.

In the Chapter 7 chemotherapy records were included if they were dated within 3 months before or 6 months after the date of diagnosis or start date. There were in fact very few cases with chemotherapy records dated before the start date therefore for clarity of methods only chemotherapy records dated on or after the start date were included. The time from diagnosis to treatment was calculated using the start date and the date of the first record of chemotherapy in HES. For patients with records of chemotherapy in the NLCA only the start date of chemotherapy was used.

Covariate definitions

Data from HES provided the information with which to calculate a Charlson co-morbidity index, as described in Chapter 2. This was calculated using HES episodes which started any time before the chemotherapy start date for patients who received chemotherapy and any time up to the date of diagnosis for those who did not receive chemotherapy.

Demographics, histology, stage and performance status were obtained from the NLCA and defined as described in section 2.2.5. Age refers to the NLCA variable age at time of diagnosis. Route of referral was also obtained from the NLCA.

Chemotherapy trusts

The system of hospital trusts in England was described in section 1.4. The NLCA records the trust at which the patient is first seen and the trust of chemotherapy treatment, which is often but not always the same. The number of patients that were both first seen and also treated with chemotherapy at each trust was plotted as a proportion of the total number of patients first seen at that trust who received chemotherapy at any trust (Figure 8-1, grey bars). For example:

Hypothetical Trust A first saw a total of 50 cases of SCLC in the study period and 35 (70%) of these patients received chemotherapy. Only 5 (14%) of these 35 patients received the chemotherapy on site at Trust A, with most receiving their chemotherapy at Trusts B, C or D; 14% would therefore be plotted for Trust A.

After inspecting the distribution of this variable, chemotherapy trusts were defined as those trusts that administered chemotherapy themselves in ≥75% of their treated cases (Trust A would not be considered a chemotherapy trust). Sensitivity analyses were performed using ≥50% and ≥90% as cut offs.

This variable was devised to assess whether this lack of services at a trust affects patient treatment decisions and outcomes. The assumption was that non-chemotherapy trusts did not have sufficient expertise and facilities to treat their own patients, and that they predominantly relied on the oncology services at other trusts for treating SCLC. As evidence that the non-chemotherapy trusts treated very few patients overall, Figure 8-1 also includes the total number of patients treated at each trust (including those referred for treatment from other trusts).

Statistical methods

The proportion of people who had chemotherapy according to the patient, tumour, and trust characteristics defined above was calculated and multivariate logistic regression was used to estimate the odds of receiving chemotherapy according to the same characteristics.

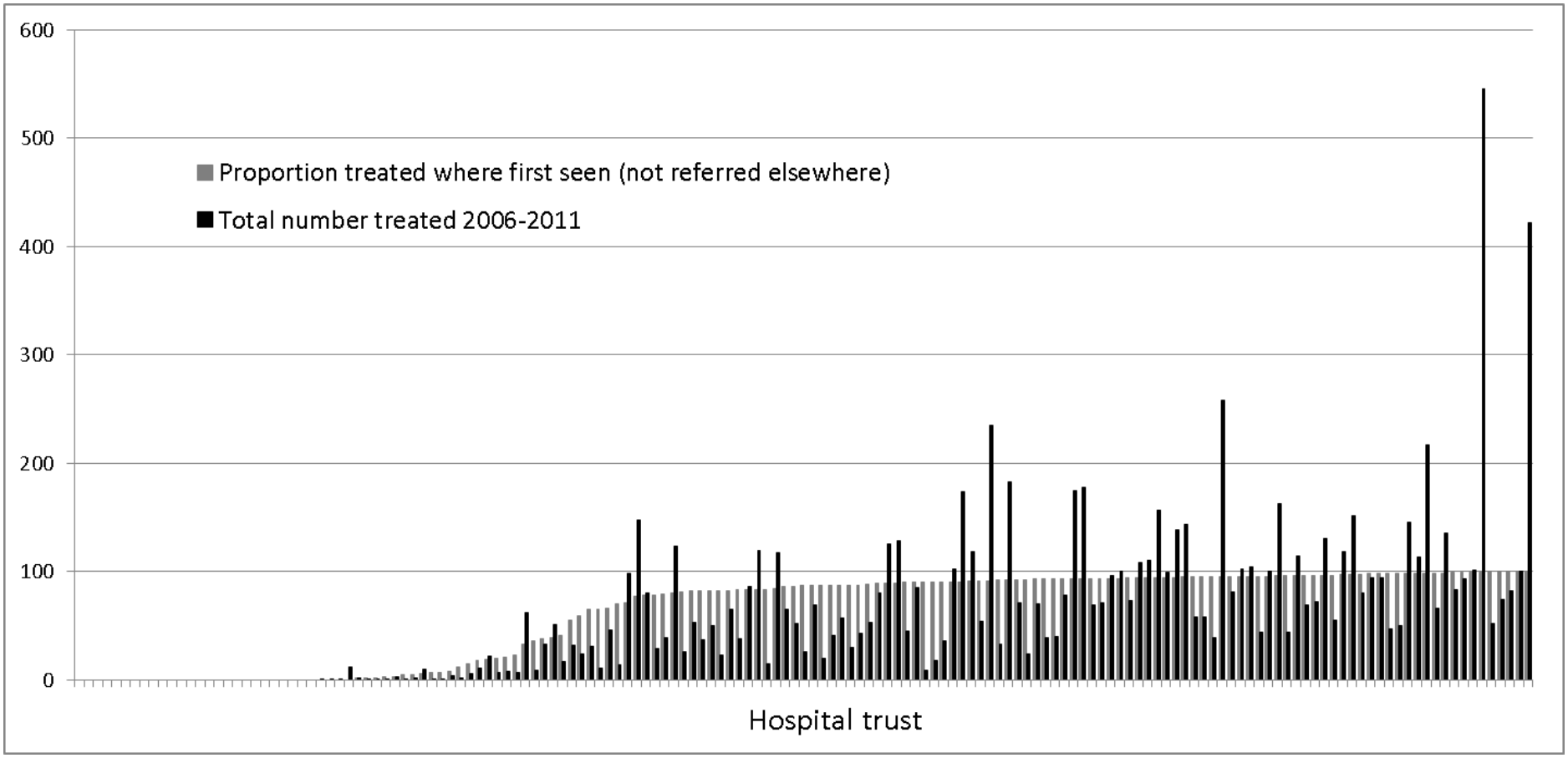


Figure 8-1: Proportion of patients with SCLC treated with chemotherapy at same trust as first seen, and total number of patients given chemotherapy at each trust 2006 - 2011.

8.2.3 Results

Study population

A total of 15,724 people with histologically confirmed SCLC diagnosed between 1st January 2006 and 30th September 2011 were identified from the NLCA. Exclusions included 289 people who had surgery and 119 with a treatment date prior to their diagnosis or after death (Figure 8-2: This differs from the study population in the previous chapter (shown in Figure 7-1) predominantly because of the exclusion of patients diagnosed before 2006 and after 31st September 2011). Trusts in one specific geographical area had an extremely high proportion of patients (65%) with no record of chemotherapy. This was believed to be due to a systematic error in data entry and therefore the 225 patients first seen in this group of trusts were excluded, leaving 15,091 people (96% of the original SCLC population) for analysis.

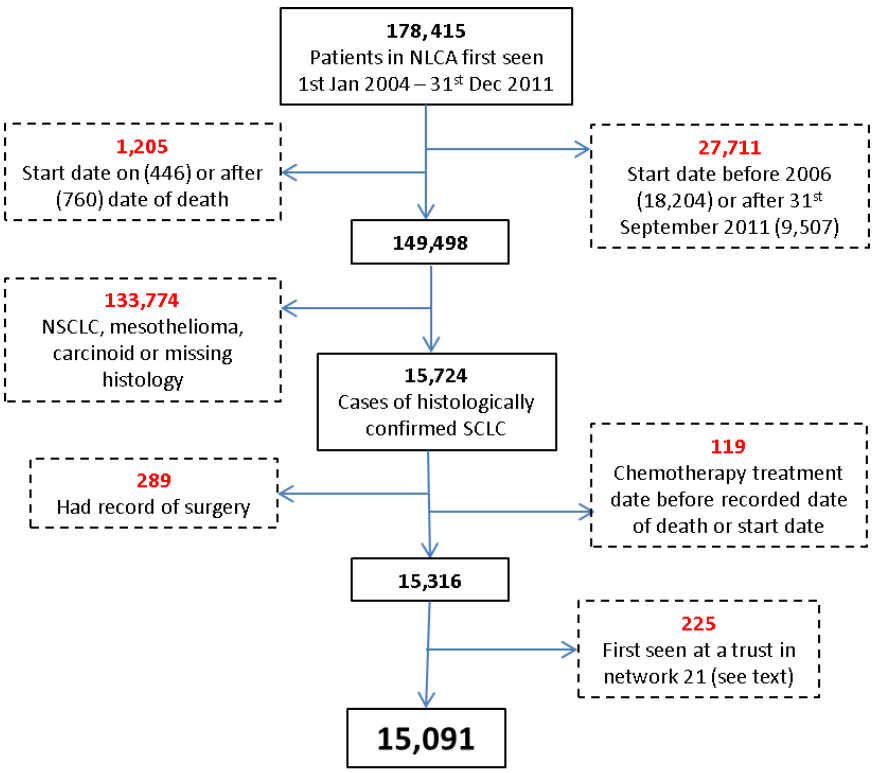


Figure 8-2: Study population and exclusions for study of chemotherapy in SCLC

Patient and tumour characteristics

The mean age at diagnosis was 68 years (standard deviation 9.7 years) and there were slightly more males than females (53% vs. 47%). More were in the most deprived Townsend quintile than any other (25% in quintile 5) (Table 8-1). Most people had performance status 0-2, with only 17% recorded as having performance status 3 or 4 at presentation (this is likely to be because all patients had histologically confirmed SCLC and must therefore have been fit enough for an invasive investigation prior to diagnosis). Fifty-five per cent had extensive stage disease at presentation (although stage was not recorded in 20% of cases).

Referral patterns

More patients (48%) were referred to the respiratory team by their general practitioner than any other route; 23% were referred via an emergency route.

Chemotherapy trusts

Seventy-three per cent of patients (n=11,032) were first seen at one of the 94 chemotherapy trusts; the 52 non-chemotherapy trusts first saw 27% of the population.

Features of patients who received chemotherapy

Seventy per cent of patients had a record of chemotherapy; 790 (7%) of these also had a record that was consistent with concurrent or sequential chemo-radiotherapy (Table 8-1). In the oldest age group (>84 years) 22% of patients received chemotherapy whereas 87% of the <55 year olds were treated. Of the 2,017 people with performance status 3, 41% had a record of chemotherapy. Sixty-six percent of patients first seen at a non-chemotherapy trust had a record of having had chemotherapy compared with 72% of those seen at a chemotherapy trust.

Table 8-1: Features of patients with SCLC who had chemotherapy

		Total N=15,091	%	Had chemotherapy n=10,582 (70%)	%
Sex	**Female**	7,126	*47.2*	5,021	*70.5*
	Male	7,965	*52.8*	5,561	*69.8*
Age group	**<55**	1,204	*8.0*	1,043	*86.6*
	55-59	1,567	*10.4*	1,314	*83.9*
	60-64	2,378	*15.8*	1,923	*80.9*
	65-69	2,766	*18.3*	2,133	*77.1*
	70-74	2,856	*18.9*	2,002	*70.1*
	75-79	2,384	*15.8*	1,435	*60.2*
	80-84	1,380	*9.1*	610	*44.2*
	>=85	556	*3.7*	122	*21.9*
Townsend quintile	**1 (Least deprived)**	2,111	*14.0*	1,499	*71.0*
(socio-economic status)	**2**	2,704	*17.9*	1,926	*71.2*
	3	2,958	*19.6*	2,083	*70.4*
	4	3,307	*21.9*	2,310	*69.9*
	5 (Most deprived)	3,738	*24.8*	2,631	*70.4*
	Missing	273	*1.8*	133	*48.7*
Performance status	**0**	2,121	*14.1*	1,909	*90.0*
	1	4,494	*29.8*	3,844	*85.5*
	2	3,072	*20.4*	2,094	*68.2*
	3	2,017	*13.4*	822	*40.8*
	4	567	*3.8*	68	*12.0*
	Missing	2,820	*18.7*	1,845	*65.4*
Charlson co-morbidity	**0**	4,899	*32.5*	3,986	*81.4*
index	**1**	2,644	*17.5*	1,996	*75.5*
	2-3	1,987	*13.2*	1,351	*68.0*
	>3	5,561	*36.8*	3,249	*58.4*
Stage	**Extensive**	8,293	*55.0*	5,474	*66.0*
	Limited	3,845	*25.5*	3,130	*81.4*
	Missing	2,953	*19.6*	1,978	*67.0*
Route of referral	**Emergency admission**	2,323	*15.4*	1,355	*58.3*
	General practitioner	7,267	*48.2*	5,624	*77.4*
	Consultant referral	2,729	*18.1*	1,869	*68.5*
	Other (inc private)	887	*5.9*	589	*66.4*
	Emergency department	1,120	*7.4*	630	*56.3*
	Missing	765	*5.1*	515	*67.3*
Trust first seen	**Non-chemotherapy trust**	4,031	*26.7*	2,673	*66.3*
	Chemotherapy trust	11,032	*73.1*	7,893	*71.5*
	Missing or trust which saw <20 cases	28	*0.2*	16	*57.1*

Time to treatment

Within the 10,582 cases treated with chemotherapy, the median time from diagnosis to initiation of treatment was 18 days (IQR 12-27) and this did not change between 2006 (19 days (13-29)) and 2011 (18 days (11-26)). For patients who first presented to a chemotherapy trust, the median time to

chemotherapy was 18 days (12-27); for those who presented to non-chemotherapy trusts this was 19 days (12-26).

Factors associated with chemotherapy treatment

After adjusting for other factors, patients were more likely to have chemotherapy if they were younger, had good performance status, limited stage disease and a low co-morbidity index (Table 8-2). People living in more socioeconomically deprived areas were less likely to be treated (likelihood ratio test for trend across Townsend quintiles in adjusted analysis p=0.0002) and compared with those referred by a general practitioner (GP) those referred to secondary care by any other route were less likely to get chemotherapy, even after adjusting for other patient features (Table 8-2).

For patients first seen at a chemotherapy trust the odds of being treated with chemotherapy were increased by 39% compared with patients seen at non-chemotherapy trusts (adjusted OR 1.39 (1.27-1.52) (Table 8-2). This difference persisted when ≥90% (adjusted OR 1.18 (1.08-1.28)) and ≥50% (adjusted OR 1.43 (1.30-1.58)) were used as the cut-off values for defining a chemotherapy trust.

Table 8-2: Odds ratios for receiving chemotherapy

		Odds Ratio (OR)			Adjusted OR*		
			95% CI			95% CI	
Sex	**Female**	**1.00**			**1.00**		
	Male	**0.97**	*0.90*	*1.04*	**0.98**	*0.91*	*1.07*
Age group	**<55**	**2.76**	*2.30*	*3.32*	**2.29**	*1.87*	*2.80*
	55-59	**2.22**	*1.89*	*2.59*	**1.88**	*1.58*	*2.23*
	60-64	**1.80**	*1.58*	*2.05*	**1.73**	*1.49*	*1.99*
	65-69	**1.44**	*1.28*	*1.62*	**1.38**	*1.20*	*1.57*
	70-74	**1.00**			**1.00**		
	75-79	**0.65**	*0.58*	*0.72*	**0.63**	*0.56*	*0.72*
	80-84	**0.34**	*0.30*	*0.39*	**0.34**	*0.29*	*0.40*
	>=85	**0.12**	*0.10*	*0.15*	**0.12**	*0.09*	*0.15*
Townsend quintile	**1 (Least deprived)**	**1.00**			**1.00**		
(socio-economic status)	**2**	**1.01**	*0.89*	*1.15*	**0.96**	*0.83*	*1.11*
	3	**0.97**	*0.86*	*1.10*	**0.94**	*0.81*	*1.09*
	4	**0.95**	*0.84*	*1.07*	**0.82**	*0.71*	*0.94*
	5 (Most deprived)	**0.97**	*0.86*	*1.09*	**0.83**	*0.72*	*0.96*
	Missing	**0.39**	*0.30*	*0.50*	**0.43**	*0.32*	*0.58*
Performance status	**0**	**1.00**			**1.00**		
	1	**0.66**	*0.56*	*0.77*	**0.86**	*0.72*	*1.02*
	2	**0.24**	*0.20*	*0.28*	**0.40**	*0.33*	*0.47*
	3	**0.08**	*0.06*	*0.09*	**0.14**	*0.12*	*0.17*
	4	**0.02**	*0.01*	*0.02*	**0.03**	*0.02*	*0.04*
	Missing	**0.21**	*0.18*	*0.25*	**0.31**	*0.26*	*0.37*
Charlson co-morbidity	**0**	**1.00**			**1.00**		
index	**1**	**0.71**	*0.63*	*0.79*	**0.89**	*0.78*	*1.02*
	2-3	**0.49**	*0.43*	*0.55*	**0.73**	*0.63*	*0.83*
	>3	**0.32**	*0.29*	*0.35*	**0.53**	*0.48*	*0.59*
Stage	**Extensive**	**1.00**			**1.00**		
	Limited	**2.25**	*2.05*	*2.47*	**1.63**	*1.46*	*1.83*
	Missing	**1.04**	*0.96*	*1.14*	**1.01**	*0.91*	*1.13*
Route of referral	**Emergency admission**	**0.41**	*0.37*	*0.45*	**0.68**	*0.60*	*0.77*
	General practitioner	**1.00**			**1.00**		
	Consultant referral	**0.63**	*0.58*	*0.70*	**0.91**	*0.81*	*1.02*
	Other (inc private)	**0.58**	*0.50*	*0.67*	**0.73**	*0.61*	*0.87*
	Emeergency department	**0.38**	*0.33*	*0.43*	**0.60**	*0.51*	*0.70*
	Missing	**0.60**	*0.51*	*0.71*	**0.79**	*0.66*	*0.96*
Trust first seen	**Non-chemotherapy trust**	**1.00**			**1.00**		
	Chemotherapy trust	**1.25**	*1.16*	*1.35*	**1.39**	*1.27*	*1.52*
	Missing or trust which saw <20 cases	**0.88**	*0.45*	*1.75*	**0.69**	*0.31*	*1.52*

*OR odds ratio; CI confidence interval; *Adjusted for all other variables*

8.3 Characteristics of patients and factors associated with completing a chemotherapy course

8.3.1 Aims

The aims of this section were to describe the characteristics of patients with SCLC who went on to complete a full course of chemotherapy and to use logistic regression to determine the factors which were independently associated with completing a course (*chapter aims 2 and 3b*).

8.3.2 Methods

Study population

This study was based on the population of people with SCLC diagnosed between 1st January 2006 and 30th September 2011, as described above but was restricted to people who had at least one record of chemotherapy in either the NLCA or HES database.

Chemotherapy cycles

For patients with a record of chemotherapy in HES it was possible to determine the number of cycles but for those with chemotherapy only recorded in the NLCA this was not possible. The main analysis was therefore restricted to people with at least one OPCS-4 code for chemotherapy in the HES database; people with chemotherapy recorded in the NLCA only were analysed as a separate group.

The number of OPCS-4 codes with associated dates at least 18 days apart (to avoid inclusion of subsequent doses in the same cycle but to allow for occasions where cycles were slightly shorter than the standard 21 days) was determined as an estimate of the number of cycles completed.

A full course of chemotherapy was defined as ≥4 cycles based on current UK recommendations.(14)

Covariates

Most covariates were defined in section 8.2.2. In addition a binary variable for time to treatment was generated using above or below the median number of days from diagnosis to first chemotherapy dose.

Patients who had radiotherapy in addition to chemotherapy were identified by the presence of a radiotherapy start date, and a record that it was given with as chemo-radiotherapy or with radical intent, in the NLCA database (further details in Chapter 7, section 7.3).

Statistical methods

The proportions of patients that received 1, 2, 3, 4, 5 and ≥6 cycles of chemotherapy were calculated. A binary variable for completion of chemotherapy was created dividing patients into those with records of 1-3 cycles and those with ≥4. The characteristics of patients in each group were examined.

The odds of completing a course of chemotherapy (receiving 4 or more cycles) were estimated and how this was associated with patient characteristics, time to treatment and receipt of chemo-radiotherapy.

8.3.3 Results

Study population

Of the 15,091 cases which met the inclusion criteria, 4,509 (30%) did not have a record of chemotherapy in either HES or the NLCA and 1,814 had a record of chemotherapy in the NLCA only. The analysis was therefore based on 8,768 cases of SCLC with at least one record of chemotherapy in HES (Figure 8-3).

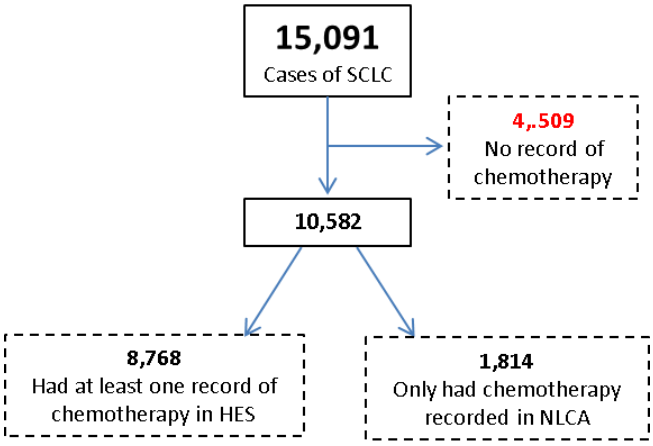


Figure 8-3: Study population for analysis of chemotherapy cycles

The study population was very similar to that described in section 8.2.2: most cases were male, 52% had extensive stage disease and the majority had a performance status of 0-2. Only 790 (9%) had records consistent with having received chemo-radiotherapy in the 6 months after diagnosis (Table 8-3).

Table 8-3: Characteristics of patients with SCLC and HES records of chemotherapy and of patients who completed ≥4 cycles

		Total N=8,768*	%	Received 1-3 cycles N=3,228 (37%)	Received ≥4 cycles N=5,540 (63%)	%
Sex	**Female**	4,166	*47.5*	1,525	2,641	*63.4*
	Male	4,602	*52.5*	1,703	2,899	*63.0*
Age group	**<55**	895	*10.2*	277	618	*69.1*
	55-59	1,106	*12.6*	351	755	*68.3*
	60-64	1,620	*18.5*	562	1,058	*65.3*
	65-69	1,777	*20.3*	608	1,169	*65.8*
	70-74	1,625	*18.5*	630	995	*61.2*
	75-79	1,147	*13.1*	502	645	*56.2*
	80-84	496	*5.7*	252	244	*49.2*
	>=85	102	*1.2*	46	56	*54.9*
Townsend quintile	**1 (Least deprived)**	1,241	*14.2*	440	801	*64.5*
(socio-economic status)	**2**	1,590	*18.1*	558	1,032	*64.9*
	3	1,743	*19.9*	641	1,102	*63.2*
	4	1,932	*22.0*	719	1,213	*62.8*
	5 (Most deprived)	2,152	*24.5*	804	1,348	*62.6*
	Missing	110	*1.3*	66	44	*40.0*
Performance status	**0**	1,626	*18.5*	383	1,243	*76.4*
	1	3,208	*36.6*	1,040	2,168	*67.6*
	2	1,664	*19.0*	761	903	*54.3*
	3	683	*7.8*	404	279	*40.8*
	4	53	*0.6*	33	20	*37.7*
	Missing	1,534	*17.5*	607	927	*60.4*
Charlson	**0**	3,318	*37.8*	1,020	2,298	*69.3*
co-morbidity index	**1**	1,641	*18.7*	565	1,076	*65.6*
	2-3	1,092	*12.5*	423	669	*61.3*
	>3	2,717	*31.0*	1,220	1,497	*55.1*
Stage	**Extensive**	4,550	*51.9*	1,838	2,712	*59.6*
	Limited	2,646	*30.2*	757	1,889	*71.4*
	Missing	1,572	*17.9*	633	939	*59.7*
Time to treatment	**<18 days**	4,305	*49.1*	1,532	2,773	*64.4*
	≥18 days	4,463	*50.9*	1,696	2,767	*62.0*
Route of referral	**Emergency admission**	1,157	*13.2*	531	626	*54.1*
	General Practitioner	4,666	*53.2*	1,500	3,166	*67.9*
	Consultant referral	1,556	*17.7*	606	950	*61.1*
	Other (inc private)	464	*5.3*	182	282	*60.8*
	Emergency Department	514	*5.9*	245	269	*52.3*
	Missing	411	*4.7*	164	247	*60.1*
Trust first seen	**Non-chemotherapy trust**	2,064	*23.5*	804	1,260	*61.0*
	Chemotherapy trust	6,688	*76.3*	2,417	4,271	*63.9*
	Missing / trust which saw <20 cases	16	*0.2*	7	9	*56.3*
Radiotherapy	**No chemo-radiotherapy**	7,978	*91.0*	3,059	4,919	*61.7*
	Chemo-radiotherapy	790	*9.0*	169	621	*78.6*

**Analysis restricted to people with SCLC who had record of chemotherapy in the Hospital Episodes Statistics database.*

Number of cycles

The proportions of patients that received 1, 2, 3, 4, 5 and ≥6 cycles of chemotherapy are shown in Table 8-4. Sixty-three percent of those who started chemotherapy received ≥4 cycles; 17% only had a record of 1 cycle.

Table 8-4: Number of chemotherapy cycles recorded for patients with SCLC

Number of cycles	Number of patients (N=8,768)	Percentage
1	1,515	17.3
2	805	9.2
3	908	10.4
4	2,091	23.9
5	981	11.2
≥6	2,468	28.2

Characteristics of people who completed a chemotherapy course

People who completed a chemotherapy course were more likely to be in younger age categories and in less deprived Townsend quintiles. A higher proportion of those with good performance status (PS) (76% of people with PS 1 compared with 41% of those with PS 3) and people with a low Charlson co-morbidity index (69% of people with CCI 0 compared with 41% of those with CCI >3) completed ≥4 cycles.

Sixty per cent of people with extensive disease completed their chemotherapy course compared with 71% of those with limited stage. People who were referred as a result of an emergency admission completed the course less frequently (54% of cases) than those referred by their GP (68%). People who had chemo-radiotherapy completed the course much more frequently than those who only had a record of chemotherapy (79% vs. 62%).

Factors associated with completing a chemotherapy course

Increasing age, performance status, stage and co-morbidity score, were independently associated with increased odds of completing chemotherapy, as was the GP route of referral compared with any other referral method. A

diagnosis to treatment interval of <18 days also increased the likelihood of completing a course compared with those who waited longer.

People who had chemo-radiotherapy were more likely to complete ≥4 cycles of chemotherapy than those who only had chemotherapy, even after adjusting for other patient characteristics (Table 8-5). There was some evidence that people first seen at a chemotherapy trust were more likely to complete a course than those seen at non-chemotherapy trusts even after adjusting for age, sex, socio-economic status, performance status, stage and co-morbidity score (OR 1.14, 95% CI 1.03-1.27).

Table 8-5: Factors associated with completing ≥4 cycles in patients with SCLC who started chemotherapy

		Odds ratio (OR) for completing ≥4 cycles	95% confidence interval (CI)		Adjusted OR**	95% CI	
Sex	**Female**	**1.00**			**1.00**		
	Male	**0.98**	*0.90*	*1.07*	**0.99**	*0.91*	*1.09*
Age group	**<55**	**1.41**	*1.19*	*1.68*	**1.29**	*1.07*	*1.54*
	55-59	**1.36**	*1.16*	*1.60*	**1.26**	*1.07*	*1.49*
	60-64	**1.19**	*1.03*	*1.38*	**1.15**	*0.99*	*1.34*
	65-69	**1.22**	*1.06*	*1.40*	**1.21**	*1.05*	*1.40*
	70-74	**1.00**			**1.00**		
	75-79	**0.81**	*0.70*	*0.95*	**0.83**	*0.71*	*0.98*
	80-84	**0.61**	*0.50*	*0.75*	**0.63**	*0.52*	*0.78*
	>=85	**0.77**	*0.52*	*1.15*	**0.80**	*0.53*	*1.21*
Townsend quintile	**1 (Least deprived)**	**1.00**			**1.00**		
(socio-economic status)	**2**	**1.02**	*0.87*	*1.19*	**0.99**	*0.84*	*1.16*
	3	**0.94**	*0.81*	*1.10*	**0.93**	*0.80*	*1.09*
	4	**0.93**	*0.80*	*1.08*	**0.90**	*0.77*	*1.05*
	5 (Most deprived)	**0.92**	*0.80*	*1.07*	**0.89**	*0.76*	*1.04*
	Missing	**0.37**	*0.25*	*0.55*	**0.37**	*0.25*	*0.56*
Performance status	**0**	**1.00**			**1.00**		
	1	**0.64**	*0.56*	*0.74*	**0.71**	*0.62*	*0.82*
	2	**0.37**	*0.31*	*0.42*	**0.45**	*0.38*	*0.52*
	3	**0.21**	*0.18*	*0.26*	**0.27**	*0.22*	*0.33*
	4	**0.19**	*0.11*	*0.33*	**0.23**	*0.13*	*0.41*
	Missing	**0.47**	*0.40*	*0.55*	**0.56**	*0.48*	*0.66*
Charlson	**0**	**1.00**			**1.00**		
co-morbidity index	**1**	**0.85**	*0.75*	*0.96*	**0.93**	*0.81*	*1.05*
	2-3	**0.70**	*0.61*	*0.81*	**0.81**	*0.70*	*0.94*
	>3	**0.54**	*0.49*	*0.61*	**0.69**	*0.62*	*0.77*
Stage	**Extensive**	**1.00**			**1.00**		
	Limited	**1.69**	*1.53*	*1.87*	**1.44**	*1.29*	*1.61*
	Missing	**1.01**	*0.89*	*1.13*	**0.95**	*0.84*	*1.08*
Time to treatment	**<18 days**	**1.00**			**1.00**		
	≥18 days	**0.90**	*0.83*	*0.98*	**0.81**	*0.74*	*0.89*
Route of referral	**Emergency admission**	**0.56**	*0.49*	*0.64*	**0.68**	*0.59*	*0.78*
	General Practitioner	**1.00**			**1.00**		
	Consultant referral	**0.74**	*0.66*	*0.84*	**0.85**	*0.75*	*0.97*
	Other (inc private)	**0.73**	*0.60*	*0.89*	**0.80**	*0.66*	*0.98*
	Emergency Department	**0.52**	*0.43*	*0.63*	**0.63**	*0.52*	*0.77*
	Missing	**0.71**	*0.58*	*0.88*	**0.82**	*0.66*	*1.01*
Trust first seen	**Non-chemotherapy trust**	**1.00**			**1.00**		
	Chemotherapy trust	**1.13**	*1.02*	*1.25*	**1.14**	*1.03*	*1.27*
	Missing / trust which saw <20 cases	**0.82**	*0.30*	*2.21*	**0.63**	*0.23*	*1.72*
Radiotherapy	**No chemo-radiotherapy**	**1.00**			**1.00**		
	Chemo-radiotherapy	**2.29**	*1.92*	*2.72*	**1.74**	*1.44*	*2.09*

**Analysis restricted to people with SCLC who had record of chemotherapy in the Hospital Episodes Statistics database; OR odds ratio, CI confidence interval; **Adjusted for age, sex, socio-economic status, co-morbidity index, performance status and stage*

8.4 Factors associated with survival in people with SCLC

8.4.1 Aims

The aims of this section were to assess survival according to number of cycles completed and to quantify overall survival according to patient characteristics and the number of cycles completed (*chapter aims 4 and 5*).

8.4.2 Methods

Study population

The survival analyses included all of the SCLC cases identified in section 8.2 with the same exclusions (Figure 8-2).

Statistical methods

Kaplan-Meier survival curves were plotted and median survival calculated from the date of diagnosis (or start date) for people with limited and extensive stage disease, overall and according to the number of cycles of chemotherapy given in the 6 months after diagnosis. To minimise the effects of immortal time bias, survival for those who had chemotherapy was also plotted from the end of the last chemotherapy cycle.

Cox regression was used to estimate survival from the date of diagnosis according to patient and tumour characteristics, whether or not chemotherapy or chemo-radiotherapy were given, the route of referral and whether or not the patient was first seen in a chemotherapy trust. Hazard ratios were calculated for each of these adjusting for age, sex, performance status, co-morbidity and stage. Survival was also estimated from the end of the last chemotherapy cycle (last chemotherapy cycle start date plus 21 days) according to the number of cycles given (1-3 or ≥4), time from diagnosis to treatment and trust first seen. This survival analysis was also performed from the end of the last chemotherapy cycle.

Patients were followed-up to death or they were censored on 31st March 2013, the date of ONS death tracing for this dataset. Interactions were sought between stage and all other variables in the effect on survival. The proportional hazards assumption was checked by inspecting Nelson-Aalen plots.

8.4.3 Results

Median survival

Median survival from diagnosis for all patients (N=15,091) was 6.2 months (IQR 1.5-12.4); this was 11.2 months (5.4-20.3) and 4.2 months (1.1-9.3) for people with limited and extensive stage disease respectively.

For those who did not receive chemotherapy (n=4,509) median survival was 2.6 months for limited and 0.9 months for extensive stage disease (Table 8-6), compared with 12.9 months (7.7-22.5) and 7.3 months (3.5–11.5) for those treated with chemotherapy.

For people with limited stage disease who completed their course of chemotherapy (n=1,889), median survival from diagnosis was 15.4 months (10.1-26.8); for those with extensive stage disease who completed the course (n=2,712) this was 9.6 months (IQR 7.2-14.0).

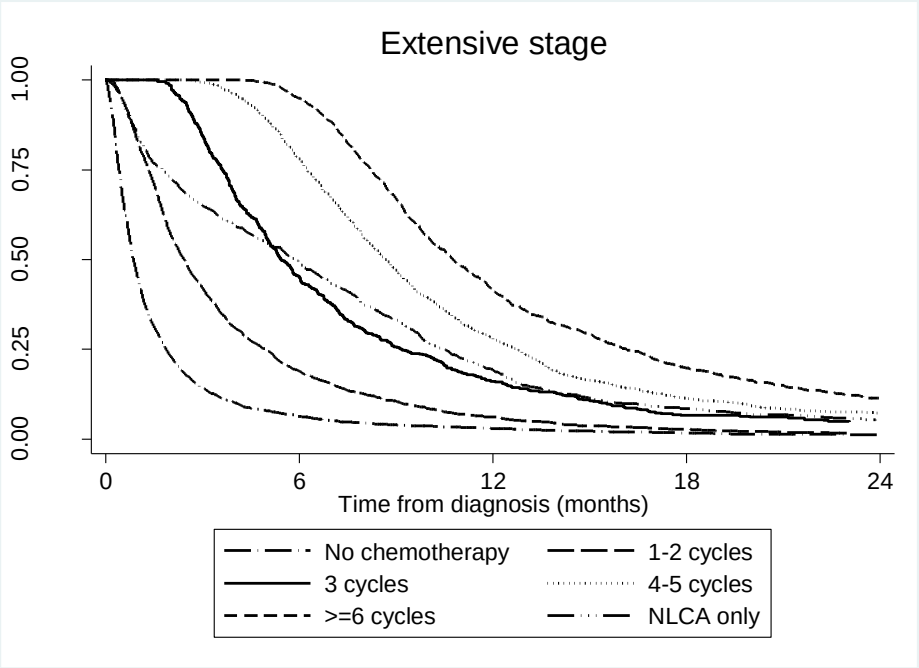
Survival by number of cycles

Survival after diagnosis for people with SCLC according to their stage at presentation and number of cycles received is shown in Figures 8-4 and 8-5. Figures 8-6 and 8- 7 show the same survival analysis restricted to people with a record of chemotherapy in HES from the end of the last chemotherapy cycle. People who received 1 or 2 cycles are grouped together, as are those who received 4 or 5 cycles because their survival curves were almost identical.

Table 8-6: Median survival in days according to stage and number of cycles of chemotherapy

Cycles	No. of patients (%) N=15,091	Median survival from diagnosis (IQR)	No. of patients (%) N=7,866*	Median survival from completing chemotherapy (IQR)
Overall				
0	4509 *(29.9)*	30 (14-74)		
1-2	2320 *(15.4)*	85 (43-185)	1595 *(20.2)*	70 (24-191)
3	908 *(6.0)*	209 (125-377)	818 *(10.4)*	126 (50-300)
4-5	3072 *(20.4)*	316 (213-527)	3000 *(38.1)*	189 (95-405)
≥6	2468 *(16.4)*	377 (270-612)	2453 *(31.2)*	224 (114-453)
NLCA only†	1814 *(12.0)*	230 (78-396)		
Extensive stage (N=8,293)				
0	2891 *(34.0)*	26 (12-56)		
1-2	1399 *(16.9)*	73 (40-149)	925 *(23.8)*	53 (20-126)
3	439 *(5.3)*	163 (108-276)	386 *(9.7)*	82 (31-192)
4-5	1432 *(17.3)*	260 (187-384)	1391 *(35.0)*	135 (68-256)
6	1280 *(15.4)*	324 (247-485)	1270 *(32.0)*	168 (88-333)
NLCA only†	924 *(11.1)*	177 (52-315)		
Limited stage (N=3,845)				
0	715 *(18.6)*	79 (25-218)		
1-2	465 *(12.0)*	141 (58-366)	346 *(16.7)*	134 (50-383)
3	295 *(7.7)*	301 (177-552)	276 *(13.3)*	213 (97-461)
4-5	1112 *(28.9)*	420 (280-758)	1098 *(53.1)*	300 (154-648)
6	777 *(20.2)*	505 (342-843)	774 *(37.5)*	347 (183-690)
NLCA only†	484 *(12.6)*	356 (199-586)		

NLCA only†: record of chemotherapy but insufficient data to calculate number of cycles
**Excludes 4,509 patients who did not receive chemotherapy, 1,814 NLCA only, and 902 that died within 21 days of starting a chemotherapy cycle.*



People with 'NLCA only' records had chemotherapy but there were insufficient data to calculate number of cycles

Figure 8-4: Kaplan Meier curve for people with extensive stage SCLC showing survival after diagnosis according to the number of chemotherapy cycles they received

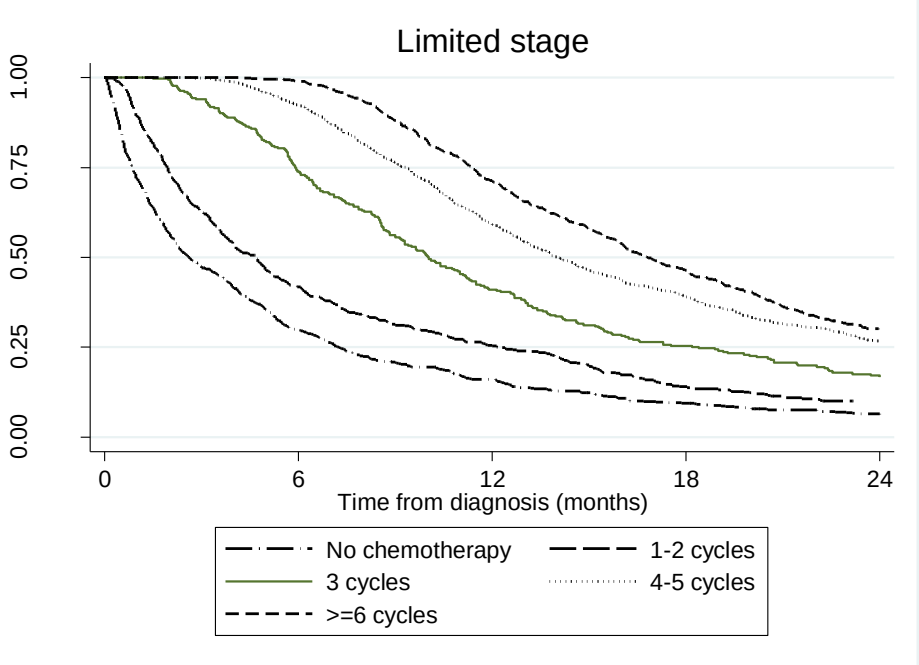


Figure 8-5: Kaplan Meier curve for people with limited stage SCLC showing survival after diagnosis according to the number of chemotherapy cycles received

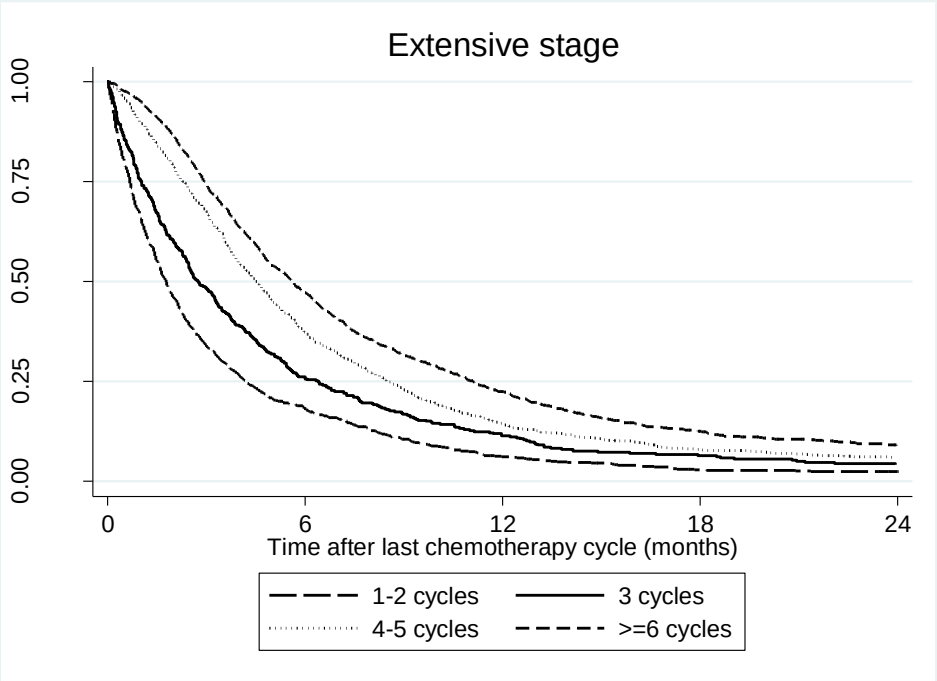


Figure 8-6: Kaplan Meier curve for people with extensive stage SCLC showing survival after finishing chemotherapy according to the number of cycles they received

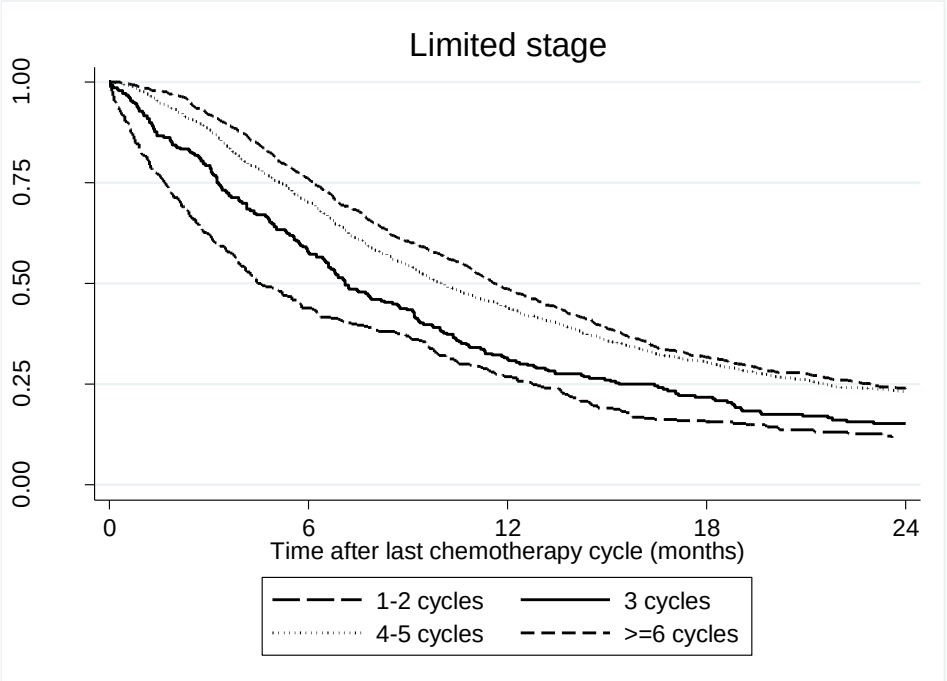


Figure 8-7: Kaplan Meier curve for people with limited stage SCLC showing survival after finishing chemotherapy according to the number of cycles they received

Factors associated with survival

The Cox regression analysis from time of diagnosis showed that for people who completed chemotherapy the risk of death was 75% lower than for those who received no chemotherapy (adjusted HR 0.25, 95% CI 0.24-0.27) (Table 8-7).

Survival did not vary significantly by Townsend quintile or according to whether the patient was first seen at a chemotherapy trust. There was evidence that people diagnosed with SCLC as a result of an emergency admission were more likely to die than people referred by their GP (HR 1.69, 95% CI 1.61-1.78); this association was attenuated but not fully explained by age, sex, PS, stage and co-morbidity index (adjusted HR 1.30, 95% CI 1.23-1.37). Those who waited ≥18 days between diagnosis and treatment had a lower risk of dying compared with those who waited <18 days (adjusted HR 0.84, 95% CI 0.81-0.88).

There was no evidence that stage modified the effect of any of the variables on survival and there was no evidence to reject the proportional hazards assumption. The results were very similar when survival for those who had chemotherapy was assessed from the end of their last recorded cycle (Table 8-8).

Table 8-7: Hazard ratios for death for people with SCLC (analysis from time of diagnosis)

Survival from diagnosis		Total patients N=15,091	Deaths n=10,041	%	Hazard ratio (HR)	95% CI		Adjusted HR*	95% CI	
Townsend quintile	**1 (Least deprived)**	2,111	1,981	*93.8*	**1.00**			**1.00**		
	2	2,704	2,558	*94.6*	**1.02**	*0.96*	*1.08*	**1.03**	*0.98*	*1.10*
	3	2,958	2,788	*94.3*	**1.02**	*0.96*	*1.08*	**1.01**	*0.95*	*1.07*
	4	3,307	3,109	*94.0*	**1.01**	*0.95*	*1.06*	**1.03**	*0.98*	*1.09*
	5 (Most deprived)	3,738	3,517	*94.1*	**0.99**	*0.94*	*1.05*	**1.02**	*0.96*	*1.08*
	Missing	273	262	*96.0*	**1.69**	*1.49*	*1.93*	**1.39**	*1.22*	*1.58*
Route ofreferral	**Emergency admission**	2,323	2,246	*96.7*	**1.69**	*1.61*	*1.78*	**1.30**	*1.24*	*1.37*
	General Practitioner	7,267	6,738	*92.7*	**1.00**			**1.00**		
	Consultant referral	2,729	2,566	*94.0*	**1.19**	*1.14*	*1.25*	**1.01**	*0.96*	*1.06*
	Other (inc private)	887	836	*94.3*	**1.18**	*1.10*	*1.27*	**1.03**	*0.96*	*1.11*
	Emergency Department	1,120	1,086	*97.0*	**1.66**	*1.56*	*1.77*	**1.29**	*1.21*	*1.37*
	Missing	765	743	*97.1*	**1.26**	*1.17*	*1.36*	**1.12**	*1.04*	*1.21*
Radiotherapy	**No chemotherapy**	4,509	*4,448*	*98.6*	**1.00**			**1.00**		
	Chemotherapy only	9,590	*8,962*	*93.5*	**0.28**	*0.27*	*0.29*	**0.35**	*0.34*	*0.37*
	Chemo-radiotherapy	992	*805*	*81.1*	**0.16**	*0.15*	*0.18*	**0.26**	*0.24*	*0.28*
Chemotherapy cycles	**No chemotherapy**	4,509	4,448	*98.6*	**1.00**			**1.00**		
	1-3 cycles	3,228	3,115	*96.5*	**0.45**	*0.43*	*0.47*	**0.54**	*0.52*	*0.57*
	≥4 cycles	5,540	4,936	*89.1*	**0.20**	*0.19*	*0.21*	**0.25**	*0.24*	*0.27*
	NLCA record only	1,814	1,716	*94.6*	**0.31**	*0.29*	*0.33*	**0.39**	*0.37*	*0.41*
Time to treatment	**<18 days**	5,137	4,834	*94.1*	**1.00**			**1.00**		
	≥18 days	5,445	4,933	*90.6*	**0.80**	*0.77*	*0.83*	**0.84**	*0.81*	*0.88*
	No chemotherapy	4,509	4,448	*98.6*	**3.30**	*3.17*	*3.44*	**2.64**	*2.52*	*2.76*
Trust first seen	**Non-chemotherapy trust**	4,276	4,011	*93.8*	**1.00**			**1.00**		
	Chemotherapy trust	10,779	10,171	*94.4*	**1.01**	*0.97*	*1.04*	**0.99**	*0.95*	*1.02*
	Missing / trust which saw <20 cases	36	33	*91.7*	**0.99**	*0.70*	*1.39*	**0.93**	*0.66*	*1.32*

*HR hazard ratio; NLCA National Lung Cancer Audit; *Adjusted for sex, age, performance status, co-morbidity index and stage*

Table 8-8: Survival for people with SCLC who had chemotherapy from end of last chemotherapy cycle

Survival analysis from end of chemotherapy		**Total patients**	**Deaths**		**Hazard**			**Adjusted**		
		N=7,866 (a)	**n=7,149**	**%**	**ratio (HR)**	*95% CI*		**HR***	*95% CI*	
Time to treatment	**<18 days**	3,794	3,524	*92.9*	**1.00**					
	≥18 days	4,072	3,625	*89.0*	**0.84**	*0.80*	*0.88*	**0.88**	*0.84*	*0.93*
Chemotherapy cycles	**1-3 cycles**	2,413	2,300	*95.3*	**1.00**					
	≥4 cycles	5,453	4,849	*88.9*	**0.58**	*0.55*	*0.60*	**0.61**	*0.58*	*0.64*
Radiotherapy	**No chemo-radiotherapy**	7,095	6,538	*92.1*	**1.00**					
	Chemo-radiotherapy	771	611	*79.2*	**0.57**	*0.53*	*0.62*	**0.72**	*0.66*	*0.78*
Trust first seen	**Non-chemotherapy trust**	1,867	1,675	*89.7*	**1.00**					
	Chemotherapy trust	5,984	5,461	*91.3*	**1.05**	*1.00*	*1.11*	**1.04**	*0.98*	*1.09*
	Missing / trust which saw <20 cases	15	13	*86.7*	**0.85**	*0.49*	*1.47*	**0.76**	*0.44*	*1.31*

*(a) this excludes the 1,814 cases with chemotherapy only recorded in the NLCA and 902 who died before completing their final cycle of chemotherapy; HR hazard ratio; *Adjusted for sex, age, performance status, co-morbidity index and stage*

8.5 Discussion

The studies described in this chapter used current English data to describe which patients with SCLC received chemotherapy treatment and provide real-life estimates of survival. There is evidence that a patient's chances of receiving and/or completing chemotherapy treatment were related not only to their fitness but also socio-economic status, the trust at which they were first seen, the time taken from diagnosis to first treatment, and their route of referral to a lung cancer specialist.

8.5.1 Strengths & Limitations

The main strengths of these studies are the large sample size and the validity of the database. (82) Mortality data from the ONS ensured that the outcome data were both accurate and complete. Chemotherapy records were validated in Chapter 7. The main limitations are the lack of detailed trust level data, and the limited number of years of therapy data available at present.

It was necessary to restrict the analysis of chemotherapy cycles to patients with chemotherapy records in HES because the NLCA only records a single chemotherapy start date. This will have excluded patients who had chemotherapy administered as an outpatient rather than inpatient day-case. Those with a record in the NLCA only were identified as such and analysed in a separate group; the survival curve in figure 8.4 gives no suggestion that these patients belong predominantly to any one of the cycle groups, and gives reassurance that these are missing from HES in a random manner.

It is possible that some patients started their chemotherapy during an inpatient admission and received further doses as an outpatient; in this case they would be misclassified as having only received one cycle. Some reassurance that this is an infrequent occurrence comes from the clinical experience of the oncologist who co-authored the publication arising from this work: not only would it be fairly unusual to treat SCLC before a patient has been discharged from hospital

because if they needed to stay in hospital it is unlikely that they would be fit enough for chemotherapy, but furthermore the proportion of patients that only received one cycle of chemotherapy (17%) correlates with clinical experience taking into account that some of these patients will have died before a second cycle could be given (personal communication Dr Vanessa Potter, Consultant Oncologist, Nottingham University Hospitals NHS Trust - November 2013).

Inpatient HES data do not capture the majority of radiotherapy episodes and therefore this work relied solely on the NLCA for radiotherapy data and may have underestimated the number of patients treated.

8.5.2 Comparison with trial data

This assessment of chemotherapy within observational population-based data cannot be used to directly assess effectiveness, but comparisons with trial data are still valuable:

Number of cycles

There are few studies examining the number of cycles of chemotherapy actually given for SCLC outside of clinical trials. Burgers et al, found no significant difference in the proportion of patients who received 4 or more cycles of chemotherapy within a trial (49/60) compared with outside of a trial (35/46) in their UK hospital.(206) However, they only included patients who were eligible for one of these two SCLC chemotherapy trials.

In this study, completion of 4 cycles was chosen to represent a complete course of chemotherapy, but practice does differ in this respect. Median survival was longer after completion of 6 or more cycles of chemotherapy compared with 4-5 cycles, but it cannot be concluded that 6 cycles are preferable to 4 as the fitness of the patients at the point of finishing 4 cycles and reasons for discontinuing treatment are unknown.

Survival

Median survival for 105 people with limited stage SCLC treated with etoposide and carboplatin in a randomised trial by the Norwegian Lung Cancer Study Group was 14.5 months from diagnosis, and 8.4 months for 113 people with extensive stage.(35) This is similar for patients in the present study (12.9 and 7.3 months).

In this study, 71% of patients with limited stage who started chemotherapy and 60% of those with extensive stage received 4 or more cycles. This is lower than in the Norwegian trial where 70% of patients overall received 5 cycles; patients treated with chemotherapy in that trial were fitter and younger than those in the NLCA (median age 64 vs. 67 years).

8.5.3 Clinical relevance

Route of referral

Those who were diagnosed as a result of an emergency hospital attendance were less likely to start chemotherapy, less likely to complete a course, and less likely to survive than those referred by a GP, even after adjusting for patient fitness and stage. There is likely to be an element of residual confounding by patient fitness in these estimates but other studies have also found this group to have poor survival.(207, 208) Given that almost a fifth of the patients in this study presented by an emergency route (which is similar to or lower than other estimates for lung cancer overall in the UK,(207)), this is an extremely important group for UK clinicians to target if overall survival from lung cancer is to improve.

Time to initiation of treatment

The average time from diagnosis to first chemotherapy dose was 18 days with 25% of patients waiting more than 27 days. A period of longer than 18 days was associated with improved survival and this is likely to reflect the impact of

prioritising patients who are unwell with aggressive disease and poor performance status. Those potentially fitter patients who waited longer than 18 days were, however, less likely to complete 4 cycles of chemotherapy than those who were treated more quickly. Completing ≥4 cycles was a strong predictor of better survival however the direction of causation is very uncertain and it may well be that non-completion of chemotherapy is due to death, rather than the other way around. Despite this uncertainly earlier treatment for all may improve survival. The 2011 UK guidelines on SCLC recommend that all patients are assessed by an oncologist within a week of deciding to recommend treatment.(14)

Trust features

People first seen at chemotherapy trusts did not survive longer than those seen at non-chemotherapy trusts despite being more likely to receive chemotherapy. Of patients first seen at a non-chemotherapy trust 66.3% had chemotherapy compared with 71.5% of those seen a chemotherapy trust. It is possible that treating an average of 5% more of their patients was not enough to translate into an overall survival benefit for chemotherapy trusts, but also that the additional patients treated were less fit and higher risk in ways that are difficult to measure.

The only other work on trust features and rates of treatment in SCLC is a previous study using the NLCA which reported increased odds of receiving chemotherapy in patients who were first seen in trusts which entered >5% of lung cancer patients into clinical trials. (83) Whilst 33 of the 94 chemotherapy trusts were classified as high trial participation centres in that study, the remaining 61 chemotherapy trusts did not have high trial participation; 7 of the 52 non-chemotherapy trusts were classified as high trial participation centres (personal communication Dr Anna Rich, May 2013).

8.5.4 Conclusion

These national data reflect the decisions that were made about chemotherapy treatment in clinical practice in England. This study has provided real-life measures of survival in those treated with chemotherapy taking into account patient and tumour characteristics and described which patients were less likely to complete a full course of treatment based on key socio-demographic and clinical features. It was not possible to determine whether 6 cycles of treatment are better than 4 cycles due to the influence of immortal time bias and a lack of longitudinal data on patient fitness and reasons for stopping treatment.

There is variation in the time from diagnosis to initiation of treatment, and some evidence of inequalities in access to treatment with particularly poor outcomes for those who present via the emergency route. This further supports the need for initiatives that improve early presentation and diagnosis. The finding that chemotherapy trusts treat a greater proportion of patients but that this does not show a survival benefit requires further work to clarify the reasons but it may be that better selection of patients for treatment and continuation of treatment is key.

8.6 Chapter summary

In this chapter I have used the NLCA-HES linked data to provide evidence of inequalities in access to chemotherapy for SCLC including particularly low treatment rates and poor survival for people diagnosed as a result of an emergency admission. I was able to demonstrate that the beneficial effects of chemotherapy in terms of survival were similar in this unselected population to those reported in clinical trials, but also that completing chemotherapy is strongly associated with improved survival and that over a third of patients do not complete 4 cycles.

I have described a novel way of defining a chemotherapy trust and found that potentially modifiable organisational factors which affect whether or not patients receive chemotherapy include the route of admission and whether or not a trust has the facilities to administer chemotherapy on site; however the consequent effect on survival is unclear and requires further investigation perhaps with longer follow-up.

This work has been accepted for publication in the British Journal of Cancer and is currently in press.

CHAPTER 9: ONGOING RESEARCH

In this chapter I will describe ongoing research relevant to this thesis. In particular I will describe a qualitative study which I am currently working on exploring patients' and clinicians' attitudes to the risks surrounding treatments for lung cancer, and two studies which I have helped to design that will test the NLCA early surgical mortality score in independent datasets.

9.1 Attitudes to risk in lung cancer surgery

When deciding whether or not a patient should have surgery it is important to recognise that the estimation of risk is not the only aspect which should be considered. The decision will also be influenced by the patient's and healthcare professionals' perceptions of these risks. Very little work has been done on the attitudes of patients and clinicians in this area and therefore a qualitative study was set up to explore these issues:

A qualitative study to map attitudes to risks surrounding treatment for lung cancer

9.1.1 Background & rationale

As has been described elsewhere in this thesis, if NSCLC is detected at an early stage then 5-year survival is much better than the overall figure of 8-9%, predominantly because of surgical resection. The overall proportion of lung cancer patients in England and Wales having an operation to try to cure their cancer varies between hospitals from less than 5% to more than 25%, (51) and the UK average is lower than other parts of Europe and North America

There are many points to consider when assessing a patient who has technically resectable lung cancer for surgery. One aspect is the likelihood that they will survive the operation and the immediate postoperative period; this was the subject of Chapter 6. However there is also a need for research which explores what level of risk is acceptable to people with lung cancer, and whether this differs from the level of risk which clinicians are willing to accept for their patients.

It is possible that given the poor survival for NSCLC, many patients might choose to have an operation even if their immediate mortality risk is high. The current National Institute for health and Clinical Excellence (NICE),(14) and

British Thoracic Society (BTS),(25) guidelines do not specifically state an acceptable mortality rate but they do publish the average 30-day mortality for lobectomy (2.3%) and pneumonectomy (5.8%). The previous BTS guidelines from 2001 recommended that surgical mortality should not be greater than 4% for lobectomy and 8% for pneumonectomy.(209) Anecdotal evidence says that most surgeons still use these figures as a guideline and would not want their own surgical mortality figures to be much higher than the national average; most would want to be lower.

Few studies have addressed patients' or clinicians' attitudes towards the risks of treatment, particularly surgical mortality, in patients with lung cancer. The available evidence would suggest that when faced with a guarantee of progressive lung cancer and no alternatives for cure, patients are willing to take relatively high risks of postoperative complications and surgery-related death.(210)

9.1.2 Aims of this study

The aim of this study was to amass qualitative data concerning patient and healthcare professionals' attitudes to the risks associated with treatments, particularly surgery, for lung cancer. Specifically the aims were:

1. *To recruit patients with a recent diagnosis of technically resectable lung cancer for semi-structured in-depth interviews (section 9.1.6)*
2. *To recruit and interview healthcare professionals who are involved with the management of patients with potentially resectable lung cancer (section 9.1.8)*
3. *To analyse interview data using the Framework method (section 9.1.7).*

9.1.3 Ethical approval

The full study protocol was approved by the University of Nottingham, who provided sponsorship and indemnity, and the Nottingham Research Ethics

Committee in August 2012; a summary of the protocol submitted for ethical approval is given in table 9-1. The study was also approved by Nottingham University Hospitals NHS Trust (NUH) Research and Innovation department so that NHS patients could be identified and recruited from the lung cancer multi-disciplinary team (MDT) meetings and clinics. Individual approval from each trust is not required for studies involving healthcare professionals.

Copies of the ethics approval letters and study documents (patient and healthcare professional information sheets and consent forms, case report forms, interview guides and generic letters to participants and health professionals) can be found in Appendix G.

9.1.4 Progress

This study involves recruiting and interviewing patients with resectable lung cancer (who only make up approximately 10% of the lung cancer population) after they receive their diagnosis but (if they are going to have surgery) before their pre-operative assessment date. For logistical regions this is a single centre study and for continuity there is only one interviewer (HP). This has resulted in very slow recruitment of patients and therefore the study is still in progress.

At the time of writing (December 2013) 15 patients have been recruited and interviewed. Saturation of themes may have been reached but detailed analysis of transcripts is needed to confirm this; if new themes are identified in the later interviews it will be necessary to recruit and interview more patients until no new themes emerge. Detailed analysis by framework method is in progress, and is currently at the stage of double coding.

Table 9-1: Attitudes to risk in lung cancer surgery: Summary of study protocol

Objectives	To map patients' and healthcare professionals' beliefs and behaviours concerning treatment with curative intent, particularly mortality following surgery, for lung cancer.
Configuration	This will be a qualitative study consisting of semi-structured in-depth interviews with patients and healthcare professionals (HCP). It will be a multi-centre study with participants recruited for interviews from multiple trusts.
Setting	Secondary care lung cancer service.
Number of participants	The maximum number of participants is expected to be 20 patients and 20 of each of the 3 categories of HCP. Numbers will remain flexible and change depending on emergence, and saturation of, themes in the interviews. The first few participants in each group will be pilots to test the interview guides.
Eligibility criteria	Participants must be over 18 years of age, able to give informed consent and communicate in English. Patients must have clinically diagnosed lung cancer stage I-IIIa and be aware of their diagnosis. HCPs must be employed by the NHS caring for lung cancer patients.
Description of interventions	Qualitative semi-structured in-depth interviews for all participants, following documentation of informed consent. Patient demographics and relevant medical information will be recorded from medical notes.
Duration of study	Interviews are expected to last approximately one hour each. Recruitment and interviews will continue until there are no new themes emerging or until the end of the recruitment period (approximately 2 years). The interview will take place within four weeks of the initial contact for patients and three months for HCPs. No further contact will be required after completion of the interview.
Outcome measures	The primary objective is to amass qualitative data which maps patients' and HCPs' views on the subject of risk associated with treatment with curative intent, predominantly surgery, for lung cancer.
Methods of Analysis	The anonymised transcripts from each interview will be systematically analysed using the Framework method. NVivo software will be used to facilitate analysis of the emergent themes and exploration of data trends and patterns.

9.1.5 Methods: Patient interviews

A schematic diagram for this part of the study is given in Appendix G

Inclusion criteria

Eligible patients met the following criteria:

1. Over 18 years of age (no upper age limit).
2. Able to give informed consent.
3. Patients with a diagnosis of lung cancer stages 1a to 3a (these stages are potentially resectable), who were aware of their diagnosis.

Exclusion criteria

Patients were not eligible for the study if they had any of the following:

1. Cognitive impairment.
2. Unable to communicate in English
3. Metastatic or high stage cancer (>IIIa) which was clearly not amenable to surgery.
4. Patients who had had, or would have had their pre-operative appointment and consented for surgery for lung cancer before an interview could take place.

Recruitment

Recruitment started in March 2013. Eligible patients were identified by HP as they were discussed at lung cancer MDT meetings at NUH. The initial approach was from a member of the patient's clinical care team. Suitable patients who expressed an interest in participating were given an invitation letter and participant information sheet (Appendix G) and informed that they would be contacted by the research team via telephone to answer any further questions. If they gave verbal consent to participate during that telephone call, an appointment for a single semi-structured in-depth interview was arranged.

It was explained to potential participants that entry into the study would be entirely voluntary and that their treatment would not be affected by their decision. It was also explained that they could withdraw at any time but that attempts would be made to avoid this occurrence. In the event of their withdrawal it was explained that their data collected so far could not be erased and consent would be sought to use the data in the final analyses where appropriate.

Duration of participation

Patients participated in the study for up to 4 weeks from first contact to completion.

Informed consent

After ensuring that the patient had at least 24 hours to consider whether or not they wished to participate, a researcher (HP) answered any questions concerning study participation over the telephone prior to arranging an interview.

All participants provided written informed consent on the day of the interview, before the interview commenced. To ensure that I was appropriately trained in obtaining informed consent for clinical research I completed the *Good Clinical Practice* training at the University of Nottingham in February 2013.

Interviews

Interviews took place at the patient's home, or at the University of Nottingham, depending on the patient's preference. Patients were informed that the interview would last approximately 60 minutes. If during the interview it became apparent that it was going to take longer than 60 minutes verbal permission to continue was sought from the patient. Interviews were digitally audio-recorded and transcribed verbatim by an external transcription company who signed a confidentiality agreement. Each transcript was checked for transcription errors to

ensure data quality. All identifiers were removed and a unique identifier assigned to each participant to ensure anonymity. The interviewer kept an interview log of impressions and interpretations which was completed after each interview.

The interviews initially followed the guide shown in Appendix G however this evolved as the interviews progressed. The first few participants were pilots to test the interview guide and this was followed by a pause in recruitment while the research team met to re-discuss the study in light of preliminary data.

Interviewer bias

All interviews were conducted by HP to ensure that any interviewer effects were consistent throughout the study. Patients were not specifically informed that the interviewer was medically trained unless they asked; they were told that I was conducting the interviews in my role as a researcher at the University of Nottingham.

Case report forms

Shortly after the interview the patient's hospital notes were used to complete a case report form (Appendix G) recording demographic information, stage of cancer, histology where known, co-morbidities, performance status and treatment plan. The healthcare professionals with whom the patient had consulted with prior to their interview was also recorded. This was done after the interview to ensure that written consent to access medical records had been obtained and so that the interview was not affected by the interviewer's prior knowledge of the patient's medical history.

Analysis skills

In order to develop some initial skills in qualitative analysis I attended the University of Nottingham *Analysing Interview Transcripts* short course in May 2012.

9.1.6 Current status of recruitment & analysis plan

Patient recruitment

Recruitment up to September 2013 is shown in Figure 9-1. At this point the interviewer felt that no new themes were emerging and therefore recruitment may be complete. Recruitment was therefore stopped and detailed analysis started (see below). If this confirms that saturation of themes no further recruitment will be necessary, if not then recruitment and interviews will continue from January 2014.

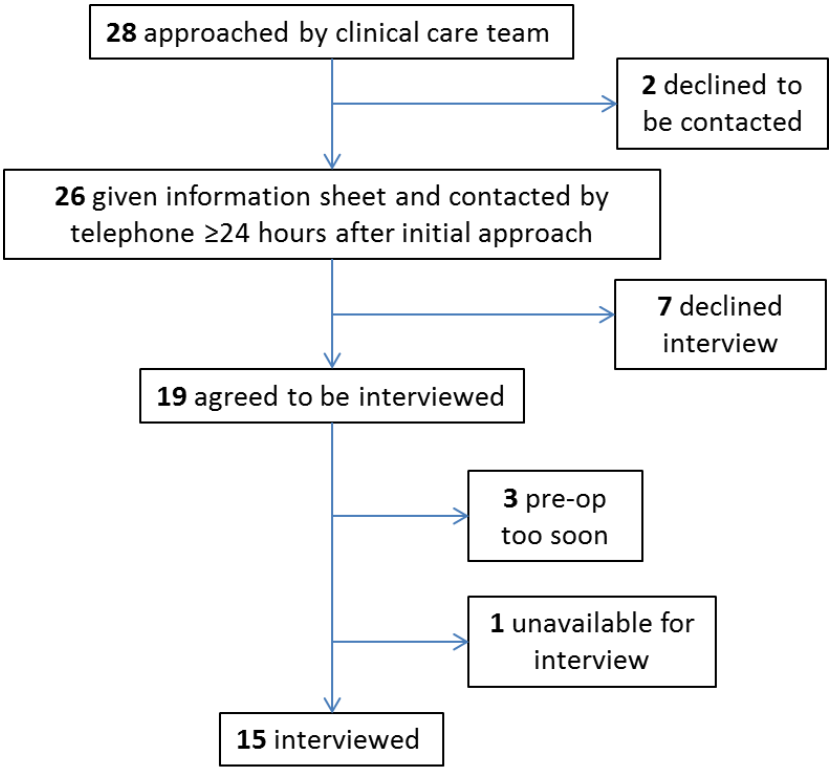


Figure 9-1: Recruitment of patients to qualitative study during March and May-September 2013

The demographics, stage, clinicians seen and treatment plan for the patients interviewed up to September 2013 are shown in Table 9-2. The majority of patients interviewed to date were male; there was a wide age range with median age 76 years. Most patients had stage I lung cancer and for all but two

the treatment plan was surgery. All patients had seen a respiratory physician prior to their interview and all but one had seen a thoracic surgeon (the treatment plan for this patient was surgery). Only 13% of patients (n=2) had seen a clinical oncologist (radiotherapist) prior to the interview and both of these patients went on to have radiotherapy rather than surgery; one of these patients had also seen a surgeon, the other did not have a surgical appointment planned.

Table 9-2: Features of patients interviewed up to September 2013

		N=15	*%*
Age	Median	76	*(Range 58-87)*
Sex	Male	10	*33*
	Female	5	*67*
Stage	1a or 1b	8	*53*
	2a or 2b	4	*27*
	3a	3	*20*
Clinicians seen	Respiratory physician	15	*100*
	Thoracic surgeon	14	*93*
	Clinical oncologist	2	*13*
	Lung cancer nurse specialist	14	*93*
Treatment plan	Surgery	13	*87*
	Radiotherapy	2	*13*

Analysis plan

The anonymised transcripts from each patient interview are currently being systematically analysed using the Framework method.(211) The first stage of this process is open coding of transcripts to facilitate the generation of a code book. The coding system will be validated through 'double coding' by two independent researchers (HP and Dr Manpreet Bains, Lecturer in Qualitative Research Methods, University of Nottingham).

Once the validated code book has been finalised, NVivo software will be used to facilitate systematic analysis of the emergent themes and exploration of data

trends and patterns. The resulting analysis matrix will be used to show the results for discussion with other members of the research team including a thoracic surgeon, lung cancer nurse and lung cancer physician.

9.1.7 Methods: Healthcare professional interviews

I have written the protocol and obtained ethical approval for recruitment and interviews of healthcare professionals, however this aspect of the study will be performed by a lecturer in qualitative research methods and a research assistant at the University of Nottingham and recruitment has not yet started.

A schematic diagram for this part of the study is shown in Appendix G. The methods will be similar to those described for patients, differences are detailed below.

Recruitment

Healthcare professionals will include thoracic surgeons, respiratory physicians and lung cancer specialist nurses, and will be recruited through their professional bodies (the Society of Cardiothoracic Surgeons, British Thoracic Society and the Association of Respiratory Nurse Specialists), via email which will include an invitation letter and participant information sheet (Appendix G). Those who are interested in participating will contact the research team (by telephone or e-mail) who will arrange an appointment for a single semi-structured in-depth interview to be conducted at a place of their choosing. The first 20 responders from each target group will be recruited, if we have more responses than this or if we reach saturation of themes before all 20 have been interviewed we will thank the respondent for their interest but they will not participate in the study.

Healthcare professionals will not be required to complete a questionnaire; the introductory questions within the interview will cover the nature of their employment and how long they have held that post.

Inclusion criteria

Healthcare professionals will be eligible if they are employed by the NHS in a job which involves contributing to treatment decisions for lung cancer patients.

Duration of participation

Healthcare professionals will participate for up to 3 months from first contact. The interviews are expected to last no longer than an hour.

Consent

If telephone interviews are used the healthcare professional will be asked to return a signed copy of the consent form by post prior to the interview taking place.

Interviews

Interviews with healthcare professionals will take place at their place of work or at an agreed conference facility depending on the participant's preference. If necessary these interviews may take place using teleconferencing; however, face-to-face interviews will be conducted wherever possible. The participant will be asked to discuss with their manager in order to decide whether the interview takes place in their own or work time.

Healthcare Professionals will be asked their own general opinions on the subject, developed over the course of their career, rather than those from or of the Trust by whom they are currently employed.

9.1.8 Timescale for completion of study

The results of the patient interviews will be compiled into manuscript format and submitted for publication in a peer reviewed journal before August 2014. Recruitment and interviews of healthcare professionals is expected to begin in August 2014.

9.2 Other ongoing research

9.2.1 Definitions of surgery, chemotherapy and radiotherapy in the NLCA

Research using the NLCA and HES data

The work described in chapters 5 and 7 has informed, and will continue to inform, other studies using the NLCA-HES linked data. I have co-authored two studies which used the definitions of surgery, chemotherapy, and radiotherapy resulting from this thesis. These studies assessed changes in treatment and survival for NSCLC and SCLC over the course of the NLCA. (77) A further study is in progress assessing organisational factors and how these affect rates of lung cancer resection.

NLCA annual reports

Only NLCA surgical data are used to calculate the proportion of patients who receive surgery and chemotherapy for the NLCA annual report. Given that NHS trusts are compared using these results the potential under-reporting of treatment rates has important implications. I presented work in chapters 5 and 7 to the NLCA co-clinical lead (Dr Paul Beckett) and members of the NLCA steering committee and this has triggered further analysis of cases which have surgery recorded in one but not both datasets. Two clinicians are analysing data from their own trusts and will review case notes to try and establish why some patients only have a record of surgery in one database and whether there is a systematic reason for this.

9.2.2 Validation of surgical score

The work described in chapters 5 and 6 produced a model (the NLCA score) which could be used in clinical practice to estimate the risk of early post-operative mortality for patients with NSCLC. It is important that the performance of the score is tested in an independent population before we advocate its use and criticisms of previous scoring systems have included a lack of robust validation in large independent cohorts. A prospective study would be the ideal methodology but in the absence of the resources necessary to conduct such a study a retrospective analysis of prospectively collected data in an unselected population would also provide useful results.

Local audit data

I obtained a local thoracic surgical audit database from Mr John Duffy, consultant thoracic surgeon at NUH, with a view to testing the performance of the NLCA score in these patients. The database contained records for 2,916 patients who underwent thoracic surgery between 14th May 1974 and 5th November 2012. After excluding patients who had undergone thoracic surgery for reasons other than potentially curative lung cancer resection, and those with insufficient data to calculate the score only 315 cases remained (Figure 9-3), of which 21 died within 90 days of surgery.

Further analysis was not performed as it was unlikely to produce useful results with this small number of deaths. In addition, the data were not truly independent as a proportion of the cases (those who had surgery between January 1st 2004 and 31st March 2010) were likely to have been included the NLCA data from which the score was derived.

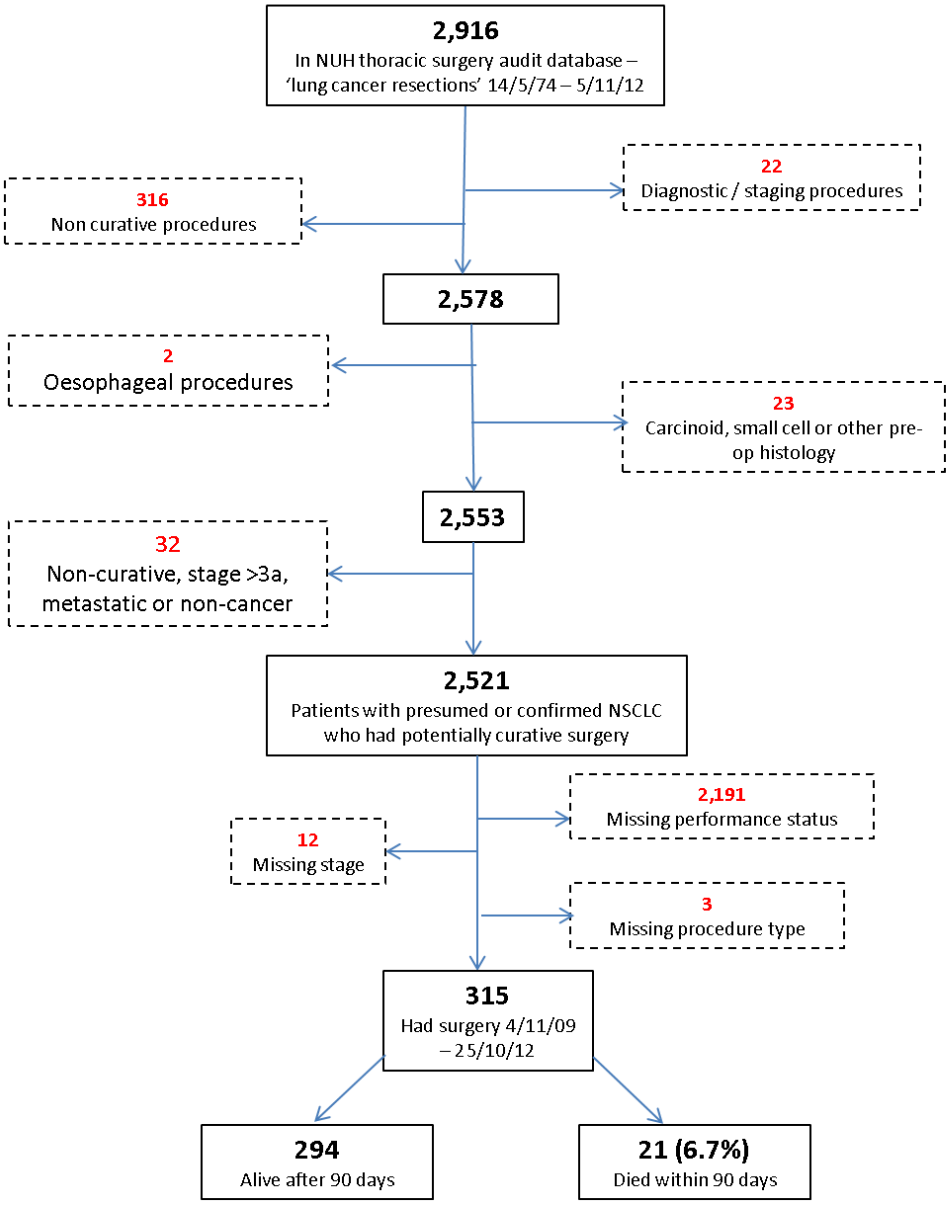


Figure 9-2: Nottingham University Hospitals thoracic surgical audit data: Cases which would have been suitable to use in testing a predictive score

NLCA-HES linked data

I obtained a more recent extract of the linked NLCA-HES dataset in July 2013 (which I used for the work described in chapters 7 and 8) and I am currently assisting another clinical research fellow with a study evaluating the

performance of the score in approximately 6,000 cases of NSCLC who underwent surgery between 1st April 2010 and 31st March 2012; we estimate that approximately 6% (n=360) would have died within 90 days and given improvements in data completeness in the NLCA we expect to be able to test the score in at least 70% of this population.

Danish Lung Cancer Registry

In addition to the above study I have been working with a lecturer in epidemiology at King's College London who has experience of working with a large Danish database of thoracic surgery (part of the Danish Lung Cancer Registry (DLCR)). We have established that this database would be also suitable to test the performance of the NLCA score and we devised a protocol for a validation study (Appendix G) so that approval could be obtained from the DLCR.

This work will involve collaboration between epidemiologists and thoracic surgeons at the University of Nottingham and King's College London, and representatives from the DLCR. The analysis will be performed by Dr Margreet Luchtenborg (Lecturer in Epidemiology, King's College London) and should be underway early in 2014.

9.2.3 Stereotactic radiotherapy

Stereotactic body radiotherapy (SBRT) was described in Chapter 1. It has potential to change the way patients with NSCLC are managed in the future. Trials comparing these methods of radiation treatment with surgical resection, particularly in patients for whom surgery carries a significant risk of mortality and morbidity have begun, and the treatment is becoming more widely available in the UK.

Comparison of outcomes after SBRT with predicted surgical outcomes

I was invited to assist with a study using the NLCA risk prediction model to compare actual short-term outcomes from SBRT with predicted outcomes from

surgery in patients who were considered too high risk for an operation. The estimation of death within 90 days (as opposed to the traditional 30 days used in all other predictive models) is particularly important for this study and others comparing SBRT with surgery because the time taken to complete the course of SBRT is often longer than 30 days. This study is based in Leeds and is being led by Dr Mat Callister; data collection through retrospective analysis of patients' medical notes is underway and I will assist with data analysis and interpretation.

9.3 Chapter summary

In this chapter I have described several ongoing research projects which are closely related to the work in this thesis; in particular a qualitative study exploring the patients' attitudes to the risks associated with lung cancer surgery and three studies which I am continuing to contribute to because they are based the NLCA score derived in Chapter 6. It is expected that when completed all of these studies will be submitted for publication in peer review journals.

In the following chapter I will summarise all of the work in this thesis, draw some conclusions and make suggestions for future research.

CHAPTER 10: SUMMARY OF THESIS AND SUGGESTIONS FOR FURTHER RESEARCH

In this chapter I will summarise the work described in this thesis, discuss some ideas for future research, and draw some conclusions.

10.1 Summary of main findings

The initial stages of this thesis add to existing evidence that women are more susceptible to cigarette smoke that men in terms of risk of lung cancer, and challenge the widely held belief that COPD is an independent risk factor for lung cancer. In carrying out these studies I developed the necessary skills to go on and use the linked NLCA-HES database to examine records of chemotherapy and surgery.

This is the first time that records of surgery and chemotherapy in HES and the NLCA have been explored. By examining patient characteristics and patterns of survival I was able to deduce that patients with a record of chemotherapy in either HES or the NLCA were likely to have received chemotherapy whereas it appeared that a record of surgery in HES was the most accurate means of identifying people who had surgery.

I was able to use the data in the NLCA-HES linked dataset to develop a predictive score which provides an estimate of the risk of death within 90 days of lung cancer surgery. This is currently undergoing evaluation in a more recent extract of the NLCA and, if it proves successful in this assessment, will be extremely useful to clinicians both nationally and internationally in the pre-treatment assessment of patients. I also used the NLCA-HES data to provide information on treatment decisions and real-life survival for people with small cell lung cancer, including an assessment of who completed chemotherapy and how this affected survival.

A qualitative assessment of patients' and healthcare professionals' attitudes to risk is in progress. This is the first time researchers have attempted to interview patients and / or clinicians about their attitudes to risk in lung cancer treatment. This study will provide a wealth of new lines of research into acceptable levels of risk and methods of communication with patients.

10.2 Clinical relevance and suggestions for further research

10.2.1 Early diagnosis of lung cancer and screening

The strong association between COPD and lung cancer is important in identifying people at high risk of lung cancer in general practice and greater awareness of the risk in this population could lead to earlier diagnosis for some patients. Further research in this area is needed and a study implementing a computerised alert system based on the presence of risk factors, including COPD, and coding of symptoms has been proposed.

Screening for lung cancer has now been introduced in the United States and is likely to be introduced in the United Kingdom at some stage. There is, however, controversy over which group to screen to ensure the most effective service. The fact that people with COPD, particularly those recently diagnosed, are at high risk of lung cancer could assist in identifying people who should be screened.

10.2.2 Post-operative mortality

The work in this thesis examined risk factors for early death after lung cancer surgery using the information available in the NLCA and HES. I hope that the validation studies will find this tool to perform well when tested in the independent data, so that it can be used in clinical practice; however there are potential areas for improvement. Performance status was one of the strongest predictors of early death in this study however this is a subjective measure which may be recorded differently by different clinicians. Collection of data such as pre-operative serum albumin, creatinine, and haemoglobin levels, and medication and smoking histories from patient records may enable us to develop a more objective means of assessing mortality risk.

The field of thoracic surgery is constantly evolving and new developments include the use of video assisted thoracic surgery (VATS) which is less invasive and therefore potentially safer in terms of operative mortality. At present VATS

lobectomy is becoming more established in clinical practice but is not offered at all thoracic centres across the UK. The factors which affect early mortality after this type of surgery may differ from those which are important after more traditional thoracic surgical procedures, and therefore repeating this analysis in 5-10 years' time may yield new information.

Postoperative morbidity and quality of life

It can be argued that mortality is the most important post-operative complication; however lung resection also risks leaving patients with long term health problems (morbidity), particularly breathlessness, which can affect their quality of life. Morbidity is more difficult to quantify, and more difficult to predict, than mortality but for some patients it is equally or even more important. A balance must be struck between extending life with surgery and the quality of life, as perceived by the individual patient, during those extra years.

In terms of predicting post-operative breathlessness, this is unlikely to be recorded in a comprehensive manner in routinely collected data. A study of factors associated with post-operative breathlessness specifically would therefore need to be a prospective cohort study; however there may be surrogate markers for morbidity which can be extracted from routinely collected data: A proportion of the patients in the NLCA-HES dataset will have their data linked with data from THIN (their primary care records) in the near future. This dataset could then be used to determine the number of consultations a patient has in a specified time period (such as a year) after their operation, the reasons for these consultations, and the number and nature of prescriptions they receive. The number of days spent in hospital in the year after surgery (which could be extracted from the HES database) could also be used as a surrogate marker of morbidity in these patients.

10.2.3 Communication of risk

Preparation of the interview guides for the qualitative study exploring perceptions of risk highlighted a lack of research into the way the risk and benefits of treatments are communicated to patients by clinicians. Once this study has been completed it will inform further research, which is also likely to require qualitative methodology, into the most effective ways of discussing operative risks with patients.

10.2.4 Chemotherapy in NSCLC and the Systemic Anti-Cancer Therapy database

The factors which influence who gets chemotherapy for NSCLC, and their subsequent survival, are currently unknown, however investigating this would be far more complicated than the analysis of chemotherapy in SCLC presented in this thesis. In contrast to SCLC, there are many different chemotherapy treatment regimens and intents for people with NSCLC, and also several other treatment options which affect survival making it difficult to determine the precise effect that chemotherapy has on long term survival.

The UK Systemic Anti-Cancer Therapy (SACT) database has collected data directly from computerised prescriptions of chemotherapy since April 2012. (212) When these data mature, and when they are linked with resources such as HES and the NLCA, this will be a valuable resource not only to look at chemotherapy in NSCLC but also to repeat the work in this thesis with information on precise chemotherapy regimens, tumour response and reasons for stopping treatment. It may be possible at that stage to provide the oncology community with some evidence on which to base their decision regarding the optimum number of cycles to aim for.

Non-small cell lung cancer is treated with TKIs as well as standard chemotherapy (as described in Chapter 1). These only became part of routine use in the UK in 2010 and therefore the number of patients in the NLCA database who received these drugs is still relatively small. The NLCA does not collect specific drug data for chemotherapy and therefore the link with the SACT

will be necessary before we can start to plan observational studies of the effects of these drugs in the unselected NLCA population (rather than current data which are from the highly selected patients in clinical trials).

10.2.5 Organisational and NHS trust-level factors

In this thesis I described a novel means of defining a chemotherapy trust using the proportion of patients who were referred elsewhere for their chemotherapy as a surrogate marker of a trust's capacity to give chemotherapy on site. Data on the actual number of oncologists with expertise in lung cancer at each trust would further inform the results and these data are needed before recommendations for workforce planning can be made based on patient outcomes.

Work is underway at the University of Nottingham collecting organisational level data on thoracic surgery, including the number of surgeons at each trust, in a study investigating inequalities in access to lung cancer surgery. The NLCA steering committee also appreciate the importance of collecting organisational as well as patient level data and intend to perform a national survey of resources during the next data collection period. This will not only include the number of clinicians with a specialist interest in lung cancer but also information on resources such as whether the trust has a positron emission tomography (PET) scanner on site (personal communication Dr Ian Woolhouse, August 2013). Research using this information combined with the NLCA-HES data will provide further insights into the effects of organisational level features on patient outcomes.

10.3 Conclusion

The studies described in this thesis demonstrate that data collected prospectively as part of routine clinical practice are valuable in answering important clinical questions in lung cancer. It is important to ensure that data such as these are representative of the study population, that outcome measures are accurate, and that any potential bias or systematic anomalies in data entry are identified. Validation studies are therefore essential and the lung cancer cases in both THIN and the NLCA have previously been found to be representative of lung cancer in England. In this thesis I have made progress in the validation of treatment records in HES and the NLCA.

Clearly there are questions that cannot be answered using retrospective analysis of routinely collected data; however numerous research questions remain which can and should be addressed. The linkage of primary care, chemotherapy treatment, and secondary care organisational level data with the NLCA and HES will be key to identifying areas where lung cancer care can be improved with the ultimate aim of improving lung cancer survival.

APPENDICES

Appendix A: Abstracts of thesis work presented at conferences

10th Annual British Thoracic Oncology Group Conference, January 2012

Poster presentation of original research:

Smoking and lung cancer in women

Powell HA,[1] Iyen-Omofoman B,[2] Hubbard RB,[1,2] Baldwin DR,[3] Tata LJ.[1]

[1] Nottingham Respiratory Research Unit, University of Nottingham, Nottingham, UK
[2] Division of Epidemiology and Public Health, University of Nottingham, Nottingham, UK
[3] Nottingham University Hospitals NHS Trust, Nottingham, UK.

Introduction: Women have smaller lungs than men and yet often smoke similar quantities of cigarettes of the same size. Recent studies have shown that for the same quantity of cigarettes smoked women are more likely to develop heart disease. We investigated whether this increased effect of smoking in women is also true for lung cancer.

Methods: Using prospectively collected general practice (GP) data from The Health Improvement Network (a medical research database containing anonymised patient records) we generated a dataset consisting of 12,121 incident cases of lung cancer and 48,216 controls. We classified patients by smoking quantity using the highest smoking quantity ever recorded. The dataset was matched on age, sex, and GP practice so we stratified our population by sex and used conditional logistic regression to calculate odds ratios (OR) for lung cancer. We used a multiplicative test for interaction to see whether the effect of smoking quantity on lung cancer differed between men and women.

Results: The odds of lung cancer were much higher in people who smoked compared to those who had never smoked, the odds increasing with quantity of cigarettes smoked (for the heaviest smokers OR 12.01, 95% confidence interval (CI) 11.16-12.92). The odds of lung cancer in women who had ever smoked heavily compared to those who had never smoked were increased nearly 14-fold (OR 13.85, 95% CI 12.45-15.41) which was more than for men smoking the same quantity (OR 10.66, 95% CI 9.64-11.79). The test for interaction showed strong evidence of a difference in effect of quantity smoked on lung cancer between men and women ($p<0.0001$).

Conclusion: Our findings reinforce the importance of smoking cessation programmes targeted at women. Further research into the effects of cigarette dose per litre lung volume may help to establish reasons for the differences we have observed.

10th Annual British Thoracic Oncology Group Conference, January 2012

Poster presentation of original research:

Is Chronic Obstructive Pulmonary Disease an Independent Risk Factor for Lung Cancer?

Powell HA,[1] Iyen-Omofoman B,[2] Hubbard RB,[1,2] Baldwin DR,[3] Tata LJ.[1]

[1] Nottingham Respiratory Research Unit, University of Nottingham, Nottingham, UK
[2] Division of Epidemiology and Public Health, University of Nottingham, Nottingham, UK
[3] Nottingham University Hospitals NHS Trust, Nottingham, UK.

Introduction: Chronic obstructive pulmonary disease (COPD) and lung cancer are two of the most important smoking related diseases, with a huge combined mortality burden. There is some evidence that COPD may be an independent risk factor for lung cancer, with airway inflammation being the pathophysiological link. This association is heavily confounded by smoking but if confirmed would be important in identifying patients who will benefit from screening, smoking cessation and perhaps chemoprevention. Current evidence comes from predominantly small studies, with few using prospectively collected data and many studying only high risk populations.

Methods: We used prospectively collected general practice (GP) data from The Health Improvement Network (a medical research database containing anonymised patient records) to generate a matched case-control dataset. We assessed the effect of a diagnosis of COPD on lung cancer and when this diagnosis was made in relation to lung cancer diagnosis (within 1 year, 1 to 5 years, 5 to 10 years or more than 10 years before). Using a conditional logistic regression model we adjusted the effect for smoking, socioeconomic status, asthma and previous pneumonia.

Results: We analysed 12,121 cases of lung cancer and 48,216 controls matched on age, sex and GP practice. The odds ratio (OR) for lung cancer was increased over fourteen-fold for patients who had a diagnosis of COPD within 6 months of lung cancer diagnosis (OR 14.39, 95% confidence interval 11.83-17.51). The effect remained when using diagnoses made more than 10 years before (OR 2.7, 95% confidence interval 2.39-3.05), and after adjusting this for smoking (OR 1.66, 95% confidence interval 1.45-1.91).

Conclusion: In this GP population a diagnosis of COPD confers an increased risk of lung cancer. The prospective nature of this study and the use of latency variables are valuable in predicting who is at higher risk of lung cancer.

British Thoracic Society Winter Meeting, December 2012

Oral presentation of original research:

Chronic obstructive pulmonary disease and risk of lung cancer: The importance of smoking and timing of diagnosis of COPD

[1]HA Powell, [2]B Iyen-Omofoman, [3]DR Baldwin, [2]RB Hubbard, [2]LJ Tata.

[1]NottinghamRespiratory Research Unit, University of Nottingham, Nottingham, UK; [2]Division of Epidemiology and Public Health, University of Nottingham, Nottingham, UK; [3]Nottingham University Hospitals NHS Trust, Nottingham, UK

Background: The majority of cases of both lung cancer and COPD are attributable to cigarette smoking. Some consider COPD to be an independent risk factor for lung cancer, even after accounting for smoking, with estimates of increased risk up to 9-fold in previous studies. We undertook a large case-control study using prospectively collected data which allowed us to quantify this association in the UK population, whilst carefully controlling for smoking and the impact of timing of diagnoses.

Methods: We used The Health Improvement Network, a UK general practice database, to identify incident cases of lung cancer and controls matched on age, sex and the practice with which they were registered. Using conditional logistic regression, we assessed the effects of timing of first diagnoses of COPD, pneumonia and asthma on the odds of lung cancer, adjusting for smoking habit and socioeconomic status.

Results: Of 11,888 incident cases of lung cancer, 23% had a prior diagnosis of COPD compared with only 6% of the 37,605 controls. The odds of lung cancer in patients who had COPD diagnosed within 6 months of their cancer diagnosis were eleven-fold those of patients without COPD (Table 1). However, when restricted to earlier COPD diagnoses, with adjustment for smoking, the effect markedly diminished (for COPD diagnoses >10 years before lung cancer diagnosis OR 2.18, 95% CI 1.87-2.54). The pattern was similar for pneumonia (see table). There was some diagnostic overlap between asthma and COPD but analyses which accounted for this produced similar results.

Conclusion: The association between COPD and lung cancer is largely explained by smoking habit, strongly dependent on the timing of COPD diagnosis and not specific to COPD. There is, however, an extremely strong unadjusted relationship of both COPD and pneumonia with lung cancer in the 6 months immediately prior to lung cancer diagnosis. This is useful in a clinical context highlighting the need to consider a diagnosis of lung cancer when making new diagnoses of COPD or pneumonia, and supporting the current NICE recommendation that all patients should have a chest radiograph looking for evidence of lung cancer at the time of COPD diagnosis.

Table 1: Odds ratios for lung cancer (N=49493, 11,888 cases and 37,605 controls)

		Odds ratio (OR)	*95% CI*	Adjusted OR*	*95% CI*
Smoking	Never	1.00		1.00	
Highest ever recorded	Trivial / light	6.00	*5.42-6.65*	5.88	*5.31-6.52*
prior to index date	Moderate	9.67	*8.87-10.54*	9.33	*8.56-10.18*
	Heavy / very heavy	15.58	*14.35-16.91*	14.88	*13.71-16.16*
	Smoker but unknown quantity	3.48	*3.20-3.78*	3.44	*3.17-3.74*
	Missing smoking status	1.79	*1.59-2.02*	1.76	*1.56-1.99*
COPD	No diagnosis prior to index date	1.00		1.00	
Interval between	within 6 months	11.47	*9.38-14.02*	6.81	*5.49-8.45*
first diagnosis &	6 months up to 1 year	4.76	*3.85-5.89*	2.52	*2.00-3.19*
index date	1 year up to 5 years	4.34	*3.95-4.78*	2.48	*2.24-2.75*
	5 years up to 10 years	4.83	*4.29-5.44*	2.68	*2.36-3.05*
	10 years or more	3.74	*3.25-4.31*	2.18	*1.87-2.54*
Pneumonia	No diagnosis prior to index date	1.00		1.00	
Interval between	within 6 months	14.91	*11.75-18.94*	13.33	*10.24-17.35*
first diagnosis &	6 months up to 1 year	3.37	*2.42-4.70*	2.89	*1.99-4.18*
index date	1 year up to 5 years	2.59	*2.22-3.02*	2.16	*1.82-2.57*
	5 years up to 10 years	2.52	*2.04-3.10*	2.11	*1.66-2.67*
	10 years or more	1.68	*1.35-2.09*	1.46	*1.15-1.86*

OR, Odds ratio. CI, confidence interval. COPD, Chronic obstructive pulmonary disease
**Adjusted for smoking & Townsend quintile (a measure of socioeconomic status)*

British Thoracic Society Winter Meeting, December 2012

Poster and short oral presentation of original research:

Early mortality after lung cancer surgery: An Analysis of the UK National Lung cancer Audit

[1]HA Powell, [2]LJ Tata, [3]DR Baldwin, [2]A Khakwani, [4]R Stanley, [2]RB Hubbard.

[1]Nottingham Respiratory Research Unit, University of Nottingham, Nottingham, UK; [2]Division of Epidemiology and Public Health, University of Nottingham, Nottingham, UK; [3]Nottingham University Hospitals NHS Trust, Nottingham, UK; [4]The NHS Information Centre, Leeds, UK

Introduction: Surgical resection is the best chance of cure for most patients with non-small cell lung cancer (NSCLC), for whom 5-year survival is otherwise poor. Selection of patients for surgery should include an estimation of the likely post-operative mortality risk but the tool often used in UK practice is a predictive score that was developed using a French database of thoracic surgical procedures, not specific to lung cancer.

Methods: We used data from the National Lung Cancer Audit linked with Hospital Episode Statistics to estimate the influence of pre-operative patient and tumour factors, and the type of procedure on the odds of death at 30 and 90 days after potentially curative surgery for NSCLC. We used logistic regression to determine which factors were associated with early post-operative mortality and then calculated the percentage of patients who died within 90 days of surgery, stratified by the strongest predictors of early post-operative mortality.

Results: We identified 12,096 patients who had potentially curative surgery for NSCLC in England between January 2004 and March 2010. Three per cent (n=387) and 6% (n=792) of patients died within 30 and 90 days respectively. Of the 12 clinical and socio-demographic factors assessed, age and type of procedure were consistently the most important predictors of early post-operative mortality: Odds ratio (OR) for death at 30 days for pneumonectomy compared with lobectomy 3.03, 95% confidence interval (CI) 2.32-3.94; and for each year increase in age OR 1.06, 95% CI 1.04-1.07. Performance status, co-morbidity score and sex and were also significantly associated with the outcomes. Table 1 shows the percentage of patients who died within 90 days of either lobectomy or pneumonectomy, stratified by age and performance status.

Conclusion: The estimation of post-operative mortality risk is a crucial part of management of patients with NSCLC. Overall mortality following surgery for NSCLC in England is currently 3% at 30-days and 6% at 90-days. We present UK data, stratified by age and performance status, which could be used in clinical practice to assist with the estimation of early post-operative mortality risk.

Table 1: Proportion of patients who died within 90 days of lobectomy or pneumonectomy for NSCLC (italics show total number of patients who underwent the procedure in each category; # no deaths occurred in these groups)

Performance status

Age	0	1	2	3-4	0	1	2
<70	**1%** *1,611*	**4%** *974*	**7%** *160*	**10%** *30*	**8%** *307*	**12%** *205*	**6%** *31*
70-80	**4%** *831*	**7%** *833*	**9%** *128*	**13%** *30*	**19%** *106*	**14%** *94*	**22%** *18*
>80	**7%** *151*	**6%** *209*	**24%** *29*	*#* *4*	**22%** *9*	**19%** *16*	*#* *0*
	LOBECTOMY				***PNEUMONECTOMY***		

11th Annual British Thoracic Oncology Group Conference, January 2013

Oral and poster presentation of original research.

Awarded 1st prize for abstract.

90-day mortality after surgery for lung cancer – An analysis of the National Lung Cancer Audit

Powell HA, Tata LJ, Baldwin DR, Stanley RA, Khakwani A & Hubbard RB.

Introduction: Almost all previous estimates of early mortality after lung cancer surgery are based on deaths within 30-days of operation: in the UK this has been estimated at 2.3% and 5.8% for lobectomy and pneumonectomy respectively. An estimate of the chances of surviving longer than 30-days may be more important to patients, especially if there is a substantial difference from 30-day mortality.

Methods: We used data from the National Lung Cancer Audit, linked with Hospital Episodes Statistics, to determine the proportion of patients who died within 30- and 90-days of potentially curative surgery for NSCLC. We then compared demographic, co-morbid, tumour and procedure related factors of patients who died between 0 and 30 days of surgery with those who died between 31 and 90 days.

Results: Of 10,991 patients who underwent surgery with curative intent for NSCLC between 2004 and 2010, 3% (334) died within 30 days of surgery and a further 2.9% (313) between 31 and 90 days. There were no significant differences in age, performance status, lung function or co-morbidity (measured by the Charlson index) between these two groups. Stage, laterality and histology also showed similar distributions within the groups. A higher proportion of those who died within 30-days had a pneumonectomy or bi-lobectomy compared with those who died between 31 and 90 days (31% vs. 20%) (see table).

Conclusion: It is important to recognise that a similar number of patients die between 31 and 90 days after lung cancer surgery as die within the first 30 days, and the features of patients who die within both of these early post-operative time periods are similar. Given that full post-operative recovery usually takes at least 2 months, we would suggest that 90-day mortality risk is a more appropriate outcome to discuss with patients prior to surgery.

Table: Comparison of features of patients who died within 30 days of surgery and between 31 and 90 days

		Overall N=10,991		Died within 30 days n=334	Died 31 – 90 days n=313
		n	*% of total*	n	n
Sex	Female	4,824	*43.9*	107	103
	Male	6,167	*56.1*	227	210
Age group	<55	1,008	*9.2*	12	23
	55-59	1,090	*9.9*	21	24
	60-64	1,847	*16.8*	31	32
	65-69	2,128	*19.4*	56	49
	70-74	2,226	*20.3*	84	78
	75-79	1,828	*16.6*	88	62
	80-84	730	*6.6*	34	35
	85+	134	*1.2*	12	10
Performance status	0	3,422	*31.1*	72	60
	1	2,815	*25.6*	84	93
	2	465	*4.2*	23	28
	3-4	108	*1.0*	11	9
	Missing	4,181	*38.0*	144	123
Per cent predicted FEV1	>80%	1,891	*17.2*	34	39
	60-79%	1,499	*13.6*	42	45
	40-59%	726	*6.6*	23	27
	<40%	141	*1.3*	5	7
	Missing	6,734	*61.3*	230	195
Stage (pre-op)	IA	2,249	*20.5*	37	41
	IB	3,064	*27.9*	87	82
	IIA	334	*3.0*	6	3
	IIB	1,494	*13.6*	55	47
	IIIA	933	*8.5*	41	39
	missing	2,857	*26.0*	108	101
Procedure	Segmentectomy/sleeve/wedge	1,686	*15.3*	35	35
	Lobectomy	7,036	*64.0*	160	165
	Bi-lobectomy	431	*3.9*	25	13
	Pneumonectomy	1,121	*10.2*	78	51
	Other	717	*6.5*	36	49

FEV1 Forced expiratory volume in 1 second

British Thoracic Society Winter Meeting, December 2013

Poster and short oral presentation of original research:

Identifying patients who receive Chemotherapy for small cell lung cancer from large Datasets

Powell HA, Tata LJ, Stanley RA, Baldwin DR, Hubbard RB

Introduction: The National Lung Cancer Audit (NLCA) has collected data on primary lung cancer in England since 2004, and has now been linked with Hospital Episodes Statistics (HES) for research into inequalities in access to treatment. How well these two large datasets capture chemotherapy for small cell lung cancer (SCLC) is not known.

Methods: We identified all cases of SCLC in the NLCA diagnosed between January 2004 and March 2012. We calculated the proportion of patients with a HES code for chemotherapy, and the proportion with a start date for chemotherapy in the NLCA, within 6 months of diagnosis. We inspected survival curves for people with a chemotherapy record in HES only or the NLCA only, people who had records of chemotherapy in both databases (who we could be reasonably sure had chemotherapy), and those with no record of chemotherapy. We assessed whether the results changed over time as case ascertainment in the NLCA increased from 19% to 98% between 2004 and 2009.

Results: We identified 18,398 cases of histologically confirmed SCLC; 9,484 (52%) had chemotherapy records in both databases and 5,100 (28%) had no record of chemotherapy in either. 737 patients (4%) had chemotherapy recorded only in HES and 2,539 (14%) only in the NLCA. For people with a record of chemotherapy in a single database (NLCA only or HES only) survival was similar to that of people with records of chemotherapy in both datasets (figure 1); the average age, stage and performance status was also very similar for people in these three groups. Survival patterns were the same when we analysed the data by year of diagnosis however the proportions with chemotherapy records in HES only or the NLCA only decreased to 3% and 12% respectively in 2011.

Conclusion: Our results suggest that it is best to identify people who received chemotherapy using data in the NLCA and HES combined. A record of chemotherapy in either database appears to be a valid means of determining who received chemotherapy but if a single database is used the proportion treated is likely to be an under-estimate.

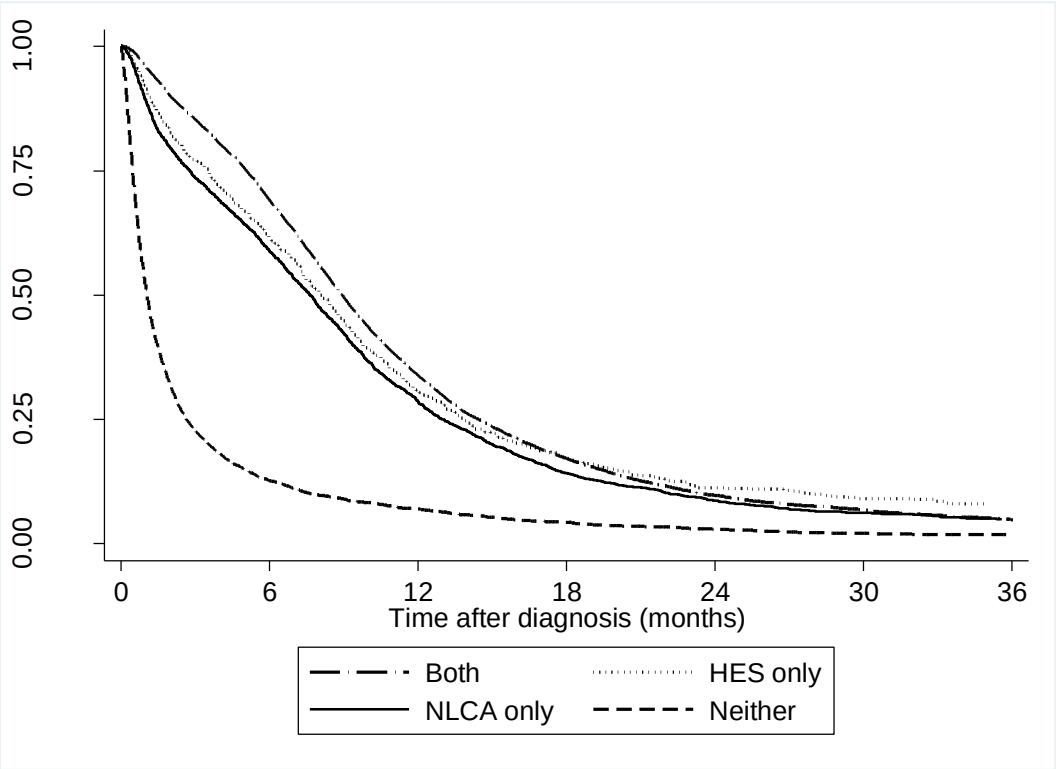


Figure 1: Survival after diagnosis for people with SCLC according to records of chemotherapy

British Thoracic Society Winter Meeting, December 2013

Poster and short oral presentation of original research:

Identifying patients who had surgical resection for Non-small cell lung cancer in large datasets

Powell HA, Tata LJ, Stanley RA, Baldwin DR & Hubbard RB

Introduction: Surgical resection rates have become an important indicator of NHS Trust performance and efforts to increase them are on-going with the aim of improving overall survival. The National Lung Cancer Audit (NLCA) has collected data on primary lung cancer since 2004 and has now been linked with Hospital Episode Statistics (HES) for research into inequalities in access to treatment. How well these two large datasets capture surgical data is not known.

Methods: We used the NLCA to identify all cases of NSCLC, excluding stage IIIB or IV, diagnosed between January 2004 and March 2010. We calculated the proportion of cases with a procedure date in the NLCA, and the proportion with a code in HES, for potentially curative surgery less than 6 months after or 3 months before diagnosis. We looked at the age, lung function, performance status, stage and survival according to where surgery was recorded. Given the increase in NLCA case ascertainment from approximately 19% in 2004 to 98% in 2009 we also looked for changes in our results over time.

Results: There were 60,196 people in the NLCA who met the inclusion criteria; 8,535 (14%) had a record of surgery in both databases. An additional 2,568 (4%) had a record of surgery in HES and 795 (1%) in the NLCA. The features of people who had surgery in HES only or the NLCA only were similar, however median survival was shorter, and the proportion that died soon after surgery was higher, in the NLCA only group compared with those with surgery records in both databases (table 1). The proportion with HES only records of surgery decreased from 6% (n=215) in 2004 to 3% (n=367) in 2009; the patterns of survival each year were similar to the overall results.

Conclusion: The proportion of people who had potentially curative surgery differed according to the database used to identify surgical procedures. There are many possible explanations for our results; however use of either database alone is likely to under-estimate the proportion of people who had surgery and this should be taken into account in studies investigating access to surgery.

Table: Features and survival of people according to the database in which records of surgery were present

N=60,196	Record of surgical procedure			
	Both n=8,535 (14%)	HES only n=2,568 (4%)	NLCA only n=795 (1%)	Neither n=48,298 (80%)
Mean age (years)	67.4	66.8	67.8	72.6
Mean % predicted FEV1	77.1	74.7	74.2	63.8
Missing FEV1 (% of total)	54.6	77.8	68.7	81.8
Stage (% of non-missing) 1a or 1b	67.2	56.4	58.4	36.2
2a or 2b	21.9	23.0	21.7	19.6
3a	10.9	20.6	19.9	44.2
Missing stage (% of total)	14.5	60.6	52.0	72.9
Performance status (% of non-missing) 0-1	92.3	86.2	85.5	47.9
2	6.4	10.2	9.0	24.1
3-4	1.2	3.6	5.5	27.9
Missing performance status (% of total)	28.2	58.9	38.2	50.4
Median survival (months)*	62	41	18	7
****Died within 30-days of surgery (%)**	2.6	4.4	5.8	
Died within 90-days of surgery (%)	5.3	8.6	16.7	

Survival is calculated from date of diagnosis not date of procedure;FEV1 Forced expiratory Volume in 1 second; **HES date of procedure unless NLCA only

12th Annual British Thoracic Oncology Group Conference, January 2014

Poster presentation of original research:

Small-cell lung cancer: Chemotherapy cycles and survival

Powell HA, Tata LJ, Baldwin DR, Potter VA, Stanley RA, Khakwani A, Hubbard RB.

Introduction
Chemotherapy is the mainstay of treatment for small cell lung cancer but is associated with side effects and toxicity that can limit the number of cycles given. We used the English National Lung Cancer Audit (NLCA) and Hospital Episodes Statistics (HES) to investigate how many cycles of chemotherapy were given to patients with SCLC, and the associated differences in survival.

Methods
We identified people in the NLCA with histologically confirmed SCLC diagnosed between January 2006 and September 2011. We used HES data to identify those who received at least one cycle of chemotherapy, and to determine the number of chemotherapy cycles each patient received, in the first 6 months after diagnosis. We calculated survival from the end of the last cycle of chemotherapy (to minimise immortal time bias), according to disease stage and the number of cycles received.

Results
Of 7,866 patients who had evidence in HES of having started chemotherapy, 63% received four or more cycles and 26% only received 1 or 2 cycles. Survival according to number of cycles received, for limited and extensive stage disease, is shown in Figure 1. People who received 1 or 2 cycles are grouped together, as are those who received 4 or 5 cycles because their survival curves were almost identical. Median survival for people who received four or more cycles was 4.9 months for extensive stage and 10.9 months for limited stage disease.

Conclusion
Patients who received more cycles of chemotherapy survived longer, even after taking into account the time during which they were undergoing treatment. We are not, however, able to recommend that patients should receive more cycles of chemotherapy from these data as we do not know the degree of tumour response to chemotherapy or the reasons for stopping treatment.

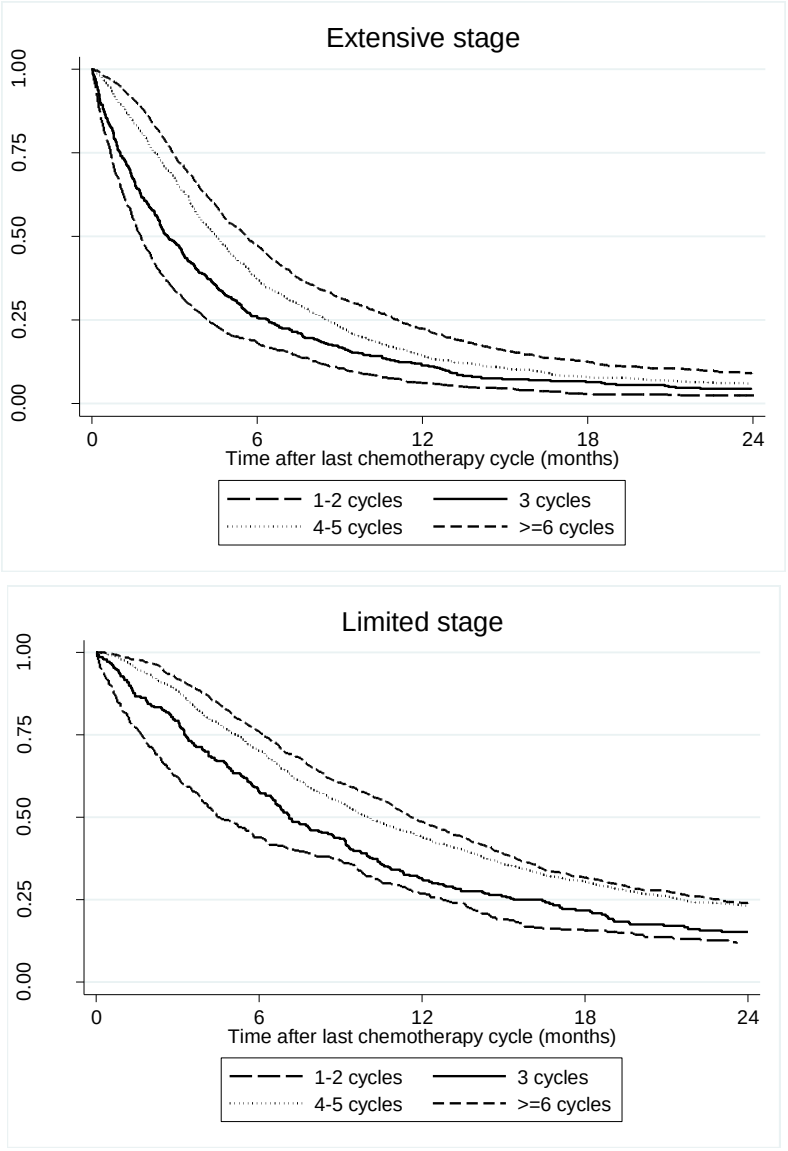


Figure 1: Kaplan Meier curves showing survival after last chemotherapy cycle according to number of cycles received

12th Annual British Thoracic Oncology Group Conference, January 2014

Poster presentation of original research:

Less chemotherapy and Poor outcomes for people with small-cell lung cancer diagnosed through emergency admission

Powell HA, Tata LJ, Baldwin DR, Potter VA, Stanley RA, Khakwani A, Hubbard RB.

Introduction
People with lung cancer in England are often diagnosed as the result of an emergency admission rather than through referral from their General Practitioner (GP), despite the existence of referral guidelines. We used the English National Lung Cancer Audit (NLCA) to examine the routes to diagnosis for people with small cell lung cancer (SCLC) and to determine how this was associated with treatment rates and survival.

Methods
Cases of SCLC diagnosed between 2006 and 2011 were identified. Linked data from Hospital Episodes Statistics (HES) were used combined with NLCA treatment records to identify patients who were treated with chemotherapy. Office for National Statistics death records were used to measure survival.
Emergency presentations were defined as referral to the lung cancer team from an emergency department or as the result of an emergency hospital admission. We used logistic regression to estimate odds of receiving chemotherapy, and Cox regression to assess survival after diagnosis.

Results
Of 15,091 cases of histologically confirmed SCLC, 48% were referred by their GP and 23% presented as an emergency. After adjustment for age, sex, performance status, co-morbidity and stage, those who presented as an emergency were less likely to have been treated with chemotherapy compared with those who were referred by their GP (table). They were also less likely to survive (HR for death 1.30 (95% confidence interval 1.23-1.37) compared with GP referrals). Median survival for emergency admissions and GP referrals was 2.6 and 7.8 moths respectively.

Conclusion
A substantial number of patients with SCLC were diagnosed via the emergency route. These patients were significantly less likely to receive chemotherapy and, perhaps consequently, less likely to survive. Further research is needed to determine how much of this effect is due to residual confounding or whether there are organisational factors which could be modified to improve outcomes.

Table: Routes of referral and receipt of chemotherapy for people with small-cell lung cancer

Route of referral	Total (N=15,091)	%	Had chemotherapy	%	Adjusted* odds ratio for receiving chemotherapy	95% confidence interval	
Emergency admission	2,323	*15.4*	1,355	*58.3*	**0.68**	*0.60*	*0.77*
General Practitioner	7,267	*48.2*	5,624	*77.4*	**1.00**		
Consultant referral	2,729	*18.1*	1,869	*68.5*	**0.91**	*0.81*	*1.02*
Other (includes private)	887	*5.9*	589	*66.4*	**0.73**	*0.61*	*0.87*
Emergency department	1,120	*7.4*	630	*56.3*	**0.60**	*0.51*	*0.70*
Missing	765	*5.1*	515	*67.3*	**0.79**	*0.66*	*0.96*

**Adjusted for stage, co-morbidity, performance status, socio-economic status, age and sex.*

Appendix B: Clinical training

Clinics

Combined lung oncology clinic

Lung cancer nurses

Oncology - chemotherapy

Oncology - radiotherapy

Surgical (pre and post-op)

Meetings

Multi-disciplinary team meetings

Network / local lung cancer policy planning meetings

National Cancer Intelligence Network lung cancer leads workshops

Procedures

Electrocautery, brachytherapy catheter placement, TBNA & EBUS

Papworth EBUS course 14-15th October 2012

Thoracoscopy

Observed thoracic surgery (open and VATS)

Observed radiotherapy sessions (palliative and SBRT)

Tutorials

The National Agenda for lung Cancer and Mesothelioma

Essential Documents in Lung Cancer and Mesothelioma

The Cancer Reform Strategy

Running an MDT meeting

Professional relationships and the MDT

An effective lung cancer service

Clinical aspects 1 – selection for radical treatment

Clinical aspects 2 – palliative chemotherapy

Clinical aspects 3 – palliative radiotherapy

Clinical aspects 4 – endo-bronchial therapy

Clinical aspects 5 – Specialist palliative Care

Clinical aspects 6 – keeping patients informed

Clinical aspects 7 – The lung cancer nurse specialist

Change management

Managing Conflict

Appendix C: Code lists for studies using the THIN database

Lung cancer Read codes

Read code	Description
B22..00	Malignant neoplasm of trachea, bronchus and lung
B220.00	Malignant neoplasm of trachea
B220z00	Malignant neoplasm of trachea NOS
B221.00	Malignant neoplasm of main bronchus
B221000	Malignant neoplasm of carina of bronchus
B221100	Malignant neoplasm of hilus of lung
B221z00	Malignant neoplasm of main bronchus NOS
B222.00	Malignant neoplasm of upper lobe, bronchus or lung
B222.11	Pancoast's syndrome
B222000	Malignant neoplasm of upper lobe bronchus
B222100	Malignant neoplasm of upper lobe of lung
B222z00	Malignant neoplasm of upper lobe, bronchus or lung NOS
B223.00	Malignant neoplasm of middle lobe, bronchus or lung
B223000	Malignant neoplasm of middle lobe bronchus
B223100	Malignant neoplasm of middle lobe of lung
B223z00	Malignant neoplasm of middle lobe, bronchus or lung NOS
B224.00	Malignant neoplasm of lower lobe, bronchus or lung
B224000	Malignant neoplasm of lower lobe bronchus
B224100	Malignant neoplasm of lower lobe of lung
B224z00	Malignant neoplasm of lower lobe, bronchus or lung NOS
B225.00	Malignant neoplasm of overlapping lesion of bronchus & lung
B22y.00	Malignant neoplasm of other sites of bronchus or lung
B22z.00	Malignant neoplasm of bronchus or lung NOS
B22z.11	Lung cancer
B26..00	Malignant neoplasm, overlap lesion of resp & intrathor orgs
B2zz.00	Malignant neoplasm of respiratory tract NOS
B551100	Malignant neoplasm of chest wall NOS
B551z00	Malignant neoplasm of thorax NOS
Byu2.00	Malignant neoplasm of respiratory and intrathoracic orga
Byu2000	Malignant neoplasm of bronchus or lung, unspecified
Byu2400	Malignant neoplasm/ill-defined sites within resp system

Smoking status Read codes

Read code	Description	Status
137..00	Tobacco consumption	see AHD
137..11	Smoker - amount smoked	Current
1371.00	Never smoked tobacco	Never
1371.11	Non-smoker	see AHD
1372.00	Trivial smoker - < 1 cig/day	Current
1372.11	Occasional smoker	Current
1373.00	Light smoker - 1-9 cigs/day	Current
1374.00	Moderate smoker - 10-19 cigs/d	Current
1375.00	Heavy smoker - 20-39 cigs/day	Current

1376.00	Very heavy smoker - 40+cigs/d	Current
1377.00	Ex-trivial smoker (<1/day)	Ex
1378.00	Ex-light smoker (1-9/day)	Ex
1379.00	Ex-moderate smoker (10-19/day)	Ex
137A.00	Ex-heavy smoker (20-39/day)	Ex
137B.00	Ex-very heavy smoker (40+/day)	Ex
137C.00	Keeps trying to stop smoking	Current
137D.00	Admitted tobacco cons untrue ?	Unknown
137E.00	Tobacco consumption unknown	Unknown
137F.00	Ex-smoker - amount unknown	Ex
137G.00	Trying to give up smoking	Current
137H.00	Pipe smoker	Current
137J.00	Cigar smoker	Current
137K.00	Stopped smoking	Ex
137L.00	Current non-smoker	see AHD
137M.00	Rolls own cigarettes	Current
137N.00	Ex pipe smoker	Ex
137O.00	Ex cigar smoker	Ex
137P.00	Cigarette smoker	Current
137P.11	Smoker	Current
137Q.00	Smoking started	Current
137Q.11	Smoking restarted	Current
137R.00	Current smoker	Current
137S.00	Ex smoker	Ex
137T.00	Date ceased smoking	Ex
137V.00	Smoking reduced	Current
137X.00	Cigarette consumption	see AHD
137Y.00	Cigar consumption	see AHD
137Z.00	Tobacco consumption NOS	see AHD
137a.00	Pipe tobacco consumption	see AHD
137b.00	Ready to stop smoking	Current
137c.00	Thinking about stopping smoking	Current
137d.00	Not interested in stopping smoking	Current
137e.00	Smoking restarted	Current
137f.00	Reason for restarting smoking	Current
137g.00	Cigarette pack-years	Unknown
137h.00	Minutes from waking to first tobacco consumption	Current
13p..00	Smoking cessation milestones	Unknown
13p0.00	Negotiated date for cessation of smoking	Current
13p1.00	Smoking status at 4 weeks	Unknown
13p2.00	Smoking status between 4 and 52 weeks	Unknown
13p3.00	Smoking status at 52 weeks	Unknown
13p4.00	Smoking free weeks	Unknown
13p5.00	Smoking cessation programme start date	Current
13p6.00	Carbon monoxide reading at 4 weeks	Unknown
4I90.00	Expired carbon monoxide concentration	Unknown
6791.00	Health ed. - smoking	Current
67A3.00	Pregnancy smoking advice	Current
67H1.00	Lifestyle advice regarding smoking	Current
6893.00	Tobacco usage screen	see AHD
68T..00	Tobacco usage screen	see AHD
745H.00	Smoking cessation therapy	Unknown
745H000	Nicotine replacement therapy using nicotine patches	Current
745H100	Nicotine replacement therapy using nicotine gum	Current
745H200	Nicotine replacement therapy using nicotine inhalator	Current

745H300	Nicotine replacement therapy using nicotine lozenges	Current
745H400	Smoking cessation drug therapy	Current
745Hy00	Other specified smoking cessation therapy	Current
745Hz00	Smoking cessation therapy NOS	Unknown
8B2B.00	Nicotine replacement therapy	Current
8B3Y.00	Over the counter nicotine replacement therapy	Current
8B3f.00	Nicotine replacement therapy provided free	Current
8BP3.00	Nicotine replacement therapy provided by community pharmacist	Current
8CAL.00	Smoking cessation advice	Current
8CAg.00	Smoking cessation advice provided by community pharmacist	Current
8H7i.00	Referral to smoking cessation advisor	Current
8HTK.00	Referral to stop-smoking clinic	Current
8I2I.00	Nicotine replacement therapy contraindicated	Current
8I39.00	Nicotine replacement therapy refused	Current
9N2k.00	Seen by smoking cessation advisor	Unknown
9N4M.00	DNA - Did not attend smoking cessation clinic	Unknown
9OO..00	Anti-smoking monitoring admin.	Unknown
9OO..11	Stop smoking clinic admin.	Unknown
9OO..12	Stop smoking monitoring admin.	Unknown
9OO1.00	Attends stop smoking monitor.	Unknown
9OO2.00	Refuses stop smoking monitor	Unknown
9OO3.00	Stop smoking monitor default	Unknown
9OO4.00	Stop smoking monitor 1st letter	Unknown
9OO5.00	Stop smoking monitor 2nd letter	Unknown
9OO6.00	Stop smoking monitor 3rd letter	Unknown
9OO7.00	Stop smoking monitor verb.inv.	Current
9OO8.00	Stop smoking monitor phone inv	Current
9OO9.00	Stop smoking monitoring delete	Unknown
9OOA.00	Stop smoking monitor. check done	Unknown
9OOZ.00	Stop smoking monitor admin.NOS	Unknown
9hG..00	Exception reporting: smoking quality indicators	Exception
9hG0.00	Excepted from smoking quality indicators: Patient unsuitable	Exception
9hG1.00	Excepted from smoking quality indicators: Informed dissent	Exception
E023.00	Nicotine withdrawal	Unknown
E251.00	Tobacco dependence	Current
E251100	Tobacco dependence, continuous	Current
E251300	Tobacco dependence in remission	Ex
E251z00	Tobacco dependence NOS	Current
ZG23300	Advice on smoking	Current
ZRBm200	Fagerstrom test for nicotine dependence	Current
ZRBm211	FTND - Fagerstrom test for nicotine dependence	Current
ZRaM.00	Motives for smoking scale	Current
ZRaM.11	MFS - Motives for smoking scale	Current
ZRao.00	Occasions for smoking scale	Current
ZRh4.00	Reasons for smoking scale	Current
ZRh4.11	RFS - Reasons for smoking scale	Current
ZV11600	Personal history of tobacco abuse	Unknown
ZV4K000	Tobacco use	see AHD
ZV6D800	Tobacco abuse counselling	Current
137j.00	Ex-cigarette smoker	Ex

Quantity smoked Read codes

Read code	Description	Quantity
1374.00	Moderate smoker - 10-19 cigs/d	current/moderate
1373.00	Light smoker - 1-9 cigs/day	current/light
1375.00	Heavy smoker - 20-39 cigs/day	current/heavy
1372.00	Trivial smoker - < 1 cig/day	current/trivial
1376.00	Very heavy smoker - 40+cigs/d	current/very heavy
1379.00	Ex-moderate smoker (10-19/day)	Ex/moderate
1378.00	Ex-light smoker (1-9/day)	Ex/light
137A.00	Ex-heavy smoker (20-39/day)	Ex/heavy
1377.00	Ex-trivial smoker (<1/day)	Ex/trivial
137B.00	Ex-very heavy smoker (40+/day)	Ex/very heavy
137..00	Tobacco consumption	see AHD
137Z.00	Tobacco consumption NOS	see AHD
137a.00	Pipe tobacco consumption	see AHD
137Y.00	Cigar consumption	see AHD
137X.00	Cigarette consumption	see AHD
ZV4K000	Tobacco use	see AHD

Quantity smoked Additional Health Data (AHD) codes

AHD code	Description	Value 1	Value 2
1003040000	Smoking	No. of cigarettes smoked/ day	Smoking status
1003040001	Smoking type	No. of cigars smoked/ day	Ounces of tobacco/ day
1003040002	Smoking dates	Date started smoking	Date stopped smoking

Smoking status: Y=Current; N=Never; D=Ex.

COPD Read codes

Read code	Description
66YI.00	COPD self-management plan given
66YL.00	Chronic obstructive pulmonary disease follow-up
66YL.11	COPD follow-up
66YL.12	COAD follow-up
66YM.00	Chronic obstructive pulmonary disease annual review
8H2R.00	Admit COPD emergency
14B3.00	History of COPD
H3...00	Chronic obstructive pulmonary disease
H3...11	Chronic obstructive airways disease
H31..00	Chronic bronchitis
H310.00	Simple chronic bronchitis
H310000	Chronic catarrhal bronchitis
H310z00	Simple chronic bronchitis NOS
H311.00	Mucopurulent chronic bronchitis
H311000	Purulent chronic bronchitis
H311100	Fetid chronic bronchitis
H311z00	Mucopurulent chronic bronchitis NOS
H312.00	Obstructive chronic bronchitis
H312100	Emphysematous bronchitis

H312200	Acute exacerbation of chronic obstructive airways disease
H312z00	Obstructive chronic bronchitis NOS
H313.00	Mixed simple and mucopurulent chronic bronchitis
H31y.00	Other chronic bronchitis
H31y100	Chronic tracheobronchitis
H31yz00	Other chronic bronchitis NOS
H31z.00	Chronic bronchitis NOS
H32..00	Emphysema
H320.00	Chronic bullous emphysema
H320000	Segmental bullous emphysema
H320100	Zonal bullous emphysema
H320200	Giant bullous emphysema
H320300	Bullous emphysema with collapse
H320z00	Chronic bullous emphysema NOS
H321.00	Panlobular emphysema
H322.00	Centrilobular emphysema
H32y.00	Other emphysema
H32y000	Acute vesicular emphysema
H32y100	Atrophic (senile) emphysema
H32y111	Acute interstitial emphysema
H32y200	MacLeod's unilateral emphysema
H32yz00	Other emphysema NOS
H32z.00	Emphysema NOS
H36..00	Mild chronic obstructive pulmonary disease
H37..00	Moderate chronic obstructive pulmonary disease
H38..00	Severe chronic obstructive pulmonary disease
H3y..00	Other specified chronic obstructive airways disease
H3y..11	Other specified chronic obstructive pulmonary disease
H3z..00	Chronic obstructive airways disease NOS
H3z..11	Chronic obstructive pulmonary disease NOS
Hyu3000	Other emphysema
Hyu3100	Other specified chronic obstructive pulmonary disease
H312000	Chronic asthmatic bronchitis
H312011	Chronic wheezy bronchitis
H312300	Bronchiolitis obliterans
H320311	Tension pneumatocoele
H32yz11	Sawyer - Jones syndrome
H3y0.00	Chronic obstruct pulmonary disease with acute lower resp infection
H3y1.00	Chronic obstruct pulmonary dis wth acute exacerbation, unspecified

Asthma Read codes

Read code	Description
173A.00	Exercise induced asthma
173c.00	Occupational asthma
173d.00	Work aggravated asthma
178..00	Asthma trigger
1780.00	Aspirin induced asthma
1J70.00	Suspected asthma
1O2..00	Asthma confirmed
2126200	Asthma resolved
212G.00	Asthma resolved

663..11	Asthma monitoring
663d.00	Emergency asthma admission since last appointment
663e.00	Asthma restricts exercise
663e000	Asthma sometimes restricts exercise
663e100	Asthma severely restricts exercise
663f.00	Asthma never restricts exercise
663h.00	Asthma - currently dormant
663j.00	Asthma - currently active
663m.00	Asthma accident and emergency attendance since last visit
663N.00	Asthma disturbing sleep
663n.00	Asthma treatment compliance satisfactory
663N000	Asthma causing night waking
663N100	Asthma disturbs sleep weekly
663N200	Asthma disturbs sleep frequently
663O.00	Asthma not disturbing sleep
663O000	Asthma never disturbs sleep
663P.00	Asthma limiting activities
663p.00	Asthma treatment compliance unsatisfactory
663q.00	Asthma daytime symptoms
663Q.00	Asthma not limiting activities
663r.00	Asthma causes night symptoms 1 to 2 times per month
663s.00	Asthma never causes daytime symptoms
663t.00	Asthma causes daytime symptoms 1 to 2 times per month
663u.00	Asthma causes daytime symptoms 1 to 2 times per week
663U.00	Asthma management plan given
663v.00	Asthma causes daytime symptoms most days
663V.00	Asthma severity
663V000	Occasional asthma
663V100	Mild asthma
663V200	Moderate asthma
663V300	Severe asthma
663w.00	Asthma limits walking up hills or stairs
663W.00	Asthma prophylactic medication used
663x.00	Asthma limits walking on the flat
663y.00	Number of asthma exacerbations in past year
66Y5.00	Change in asthma management plan
66Y9.00	Step up change in asthma management plan
66YA.00	Step down change in asthma management plan
66YC.00	Absent from work or school due to asthma
66YE.00	Asthma monitoring due
66YJ.00	Asthma annual review
66YK.00	Asthma follow-up
66YP.00	Asthma night-time symptoms
66YQ.00	Asthma monitoring by nurse
66YR.00	Asthma monitoring by doctor
66YZ.00	Does not have asthma management plan
679J.00	Health education - asthma
8791.00	Further asthma - drug prevent.
8793.00	Asthma control step 0
8794.00	Asthma control step 1
8795.00	Asthma control step 2
8796.00	Asthma control step 3
8797.00	Asthma control step 4
8798.00	Asthma control step 5
8B3j.00	Asthma medication review

8CR0.00	Asthma clinical management plan
8H2P.00	Emergency admission, asthma
8HTT.00	Referral to asthma clinic
9hA..00	Exception reporting: asthma quality indicators
9hA1.00	Excepted from asthma quality indicators: Patient unsuitable
9hA2.00	Excepted from asthma quality indicators: Informed dissent
9Q21.00	Patient in asthma study
G581.11	Asthma - cardiac
H312000	Chronic asthmatic bronchitis
H33..00	Asthma
H33..11	Bronchial asthma
H330.00	Extrinsic (atopic) asthma
H330.11	Allergic asthma
H330.12	Childhood asthma
H330.13	Hay fever with asthma
H330.14	Pollen asthma
H330000	Extrinsic asthma without status asthmaticus
H330011	Hay fever with asthma
H330100	Extrinsic asthma with status asthmaticus
H330111	Extrinsic asthma with asthma attack
H330z00	Extrinsic asthma NOS
H331.00	Intrinsic asthma
H331.11	Late onset asthma
H331000	Intrinsic asthma without status asthmaticus
H331100	Intrinsic asthma with status asthmaticus
H331111	Intrinsic asthma with asthma attack
H331z00	Intrinsic asthma NOS
H332.00	Mixed asthma
H333.00	Acute exacerbation of asthma
H334.00	Brittle asthma
H33z.00	Asthma unspecified
H33z000	Status asthmaticus NOS
H33z011	Severe asthma attack
H33z100	Asthma attack
H33z111	Asthma attack NOS
H33z200	Late-onset asthma
H33zz00	Asthma NOS
H33zz11	Exercise induced asthma
H33zz12	Allergic asthma NEC
H35y600	Sequoiosis (red-cedar asthma)
H35y700	Wood asthma
H47y000	Detergent asthma
SLF7.00	Antiasthmatic poisoning
SLF7z00	Antiasthmatic poisoning NOS
TJF7.00	Adverse reaction to antiasthmatics
TJF7300	Adverse reaction to theophylline (asthma)
TJF7z00	Adverse reaction to antiasthmatic NOS
U60F600	Antiasthmats caus adverse effects in therapeut use, NEC
U60F611	Adverse reaction to antiasthmatics
U60F615	Adverse reaction to theophylline - asthma
U60F61A	Adverse reaction to antiasthmatic NOS

Pneumonia Read codes

Read code	Description
H26..00	Pneumonia due to unspecified organism
H25..00	Bronchopneumonia due to unspecified organism
H062.00	Acute lower respiratory tract infection
H21..00	Lobar (pneumococcal pneumonia)
H2z..00	Pneumonia or Influenza NOS
H2...00	Pneumonia and influenza
H261.00	Basal pneumonia due to unspecified organism
H28..00	Atypical pneumonia
H260.00	Lobar pneumonia due to unspecified organism
H231.00	Pneumonia due to mycoplasma pneumoniae
H20..00	Viral pneumonia
H20z.00	Viral pneumonia NOS
H22z.00	Bacterial pneumonia NOS
H540000	Hypostatic pneumonia
SP13100	Other aspiration pneumonia as a complication of care
H56y100	Interstitial pneumonia
H540100	Hypostatic bronchopneumonia
H22..00	Other bacterial pneumonia
H23..00	Pneumonia due to other specified organisms
H470312	Aspiration pneumonia due to vomit
H223.00	Pneumonia due to streptococcus
H201.00	Pneumonia due to respiratory syncitial virus
H2y..00	Other specified pneumonia or influenza
H22..11	Chest infection- other bacterial pneumonia
H22y200	Pneumonia-legionella
H25..11	Chest infection- unspecified bronchopneumonia
H270000	Influenza with bronchopneumonia
A3BXA00	Mycoplasma pneumoniae (PPLO) cause/dis classifd/oth
H224.00	Pneumonia due to staphylococcus
H270.00	Influenza with pneumonia
H20..11	Chest infection- viral pneumonia
H220.00	Pneumonia due to klebsiella pneumoniae
H23z.00	Pneumonia due to specified organism NOS
A3BXB00	Klebsiella pneumoniae/cause/disease classifd/oth chapt
A789300	HIV disease resulting in Pneumocystis carinii pneumonia
H060A00	Acute bronchitis due to mycoplasma pneumoniae
H221.00	Pneumonia due to pseudomonas
H233.00	Chlamydial pneumonia
H24y200	Pneumonia with pneumocystis carinii
H262.00	Postoperative pneumonia
H270.11	Chest infection- influenza with pneumonia
A116.00	Tuberculous pneumonia
A380300	Septicaemia due to streptococcus pneumoniae
AB24.11	Pneumonia- candidal
H06z112	Acute lower respiratory tract infection
H20y.00	Viral pneumonia NEC
H222.00	Pneumonia due to haemophilus influenzae
H22y000	Pneumonia due to Escherichia coli
H22yz00	Pneumonia due to bacteria NOS
H24..00	Pneumonia with infectious disease EC
H24y700	Pneumonia with varicella

Appendix D: NLCA data entry form

Demographics

NHS Number		Organisation Code	
Forenames		Surname	
Sex		Date of Birth	
Postcode			

Referral Information

Source of referral	☐ Following an emergency admission	☐ Following a domiciliary visit
	☐ Referral from consultant other than in A&E	☐ Referral from GP
	☐ Following A&E attendance	☐ General Dental Practitioner
	☐ Community Dental Service	☐ Other source of referral
	☐ Not known	
Date of decision to refer (2-week patients only)		
Lung cancer specialist referral date (non-2-week patients)		
Date first seen		
Place first seen		

Investigations

Had a CT scan?	☐ No	☐ Yes.....Date:
Had a PET scan?	☐ No	☐ Yes.....Date:
Had a bronchoscopy?	☐ No	☐ Yes.....Date:
Had a CT-guided biopsy?	☐ No	☐ Yes.....Date:
Had other diagnostic biopsy?	☐ No	☐ Yes.....Date:

Diagnosis

Date of diagnosis		
Place of diagnosis		
Pre-treatment histology		
Primary site diagnosis	☐ Bronchus or lung, unspecified	☐ Malignant neoplasm of bronchus or lung
	☐ Main bronchus, Carina, Hilus of lung	☐ Upper lobe, bronchus or lung (incl. pancoast)
	☐ Middle lobe/lingular, bronchus or lung	☐ Lower lobe, bronchus or lung
	☐ Trachea	☐ Overlapping lesion of bronchus and lung
	☐ Mediastinum, part unspecified	☐ Pleura
	☐ Malignant neoplasm of heart, mediastinum and pleura	☐ Overlapping lesion of heart, mediastinum and pleura
	☐ Mesothelioma	☐ Mesothelioma of pleura

Laterality	☐ Left	☐ Midline	☐ Right
	☐ Bilateral	☐ Unknown	☐ Not applicable

Basis of diagnosis	☐ Death certificate	☐ Clinical
	☐ Clinical investigation	☐ Specific tumour markers
	☐ Cytology	☐ Histology of a metastasis
	☐ Histology of a primary tumour	☐ Unknown

Staging

Staging procedure performed?			
Mediastinoscopy/Mediastinotomy			
FNA staging procedure performed?			
Other staging procedure performed?			
Unknown staging procedure performed?			
Pre-treatment Stage	T	N	M
NSCLC Stage	Will be calculated based on TNM above		
SCLC Stage	☐ Limited	☐ Extensive	☐ Unknown

Care Plan/MDT

Discussed at MDT?	☐ Yes…….Date:	☐ No	☐ Unknown
Treatment intent	☐ Curative	☐ Palliative	
	☐ Palliative (supportive care only)	☐ Unknown	
	☐ No specific anti-cancer treatment		
Treatment modalities	☐ Single modality	☐ Multiple modality	☐ Unknown
Suggested plan	☐ Surgery	☐ Radiotherapy	
	☐ Chemotherapy	☐ Brachytherapy	
	☐ Palliative care	☐ Active monitoring	
	☐ Sequential chemotherapy and radiotherapy		
	☐ Concurrent chemotherapy and radiotherapy		
	☐ Induction chemo to downstage before surgery		
	☐ Neo-adjuvant chemotherapy and surgery		
	☐ Surgery followed by chemotherapy		

Co-Morbidities

Was there any reason why the patient did not receive the first choice of treatment?	☐ Died	☐ COPD	☐ Refused		
	☐ Co-morbidity precluding treatment				
Co-morbidities	☐ Dementia/Cerebrovascular disease		☐ Cardiovascular disease		
	☐ Renal failure		☐ Other malignancy		
	☐ Severe weight loss		☐ Other		
FEV1 Absolute					
FEV1 percentage					
Performance Status	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ Not recorded

Treatment - Surgery

Hospital code				
Date of decision to operate				
Date of surgery				
Main surgical procedure	☐ Wedge resection of lesion of lung		☐ Multiple wedges resected	
	☐ Segmental resection		☐ Sleeve resection	
	☐ Lung resection with resection of chest wall (not identifying which lobe resection)			
	☐ Carinal resection			
	☐ Lobectomy		☐ Pneumonectomy	
	☐ Open operation on lung (open and close)		☐ Bilobectomy	
	☐ Other open operation on lung		☐ Extrapleural pneumonectomy	
	☐ Debulking pleurectomy		☐ Pleurodesis	
Completeness of resection	☐ Presence of residual tumour cannot be assessed	☐ No residual tumour	☐ Microscopic residual tumour	☐ Macroscopic residual tumour
Surgical histology				
Date of surgical histology				
Pathological stage	pT	pN	pM	
Pathological NSCLC Stage				
Pathological SCLC Stage				

Treatment - Chemotherapy

Hospital code	

Treatment - Chemotherapy

Date of decision to treat	
Date of start of treatment	
Chemotherapy intent	☐ Chemotherapy alone
	☐ Neo-adjuvant chemotherapy before surgery
	☐ Part of a chemotherapy / radiotherapy treatment plan
	☐ Adjuvant chemotherapy post surgery
	☐ Induction chemotherapy to down stage before surgery

Treatment - Radiotherapy

Hospital code		
Date of decision to treat		
Date of start of treatment		
Radiotherapy site	☐ Trachea	☐ Lung
	☐ Mediastinum	☐ Skin
	☐ Chest wall	☐ Bone
	☐ Mesothelioma drain site	☐ Other Region of Body
	☐ Brain	
Radiotherapy intent	☐ Curative (radical) radiotherapy	
	☐ Curative (CHART / CHARTWEL)	
	☐ Part of a chemotherapy / radiotherapy treatment plan	
	☐ Adjuvent following surgical treatment	
	☐ Palliative Radiotherapy	

Treatment - Brachytherapy

Hospital code	
Date of decision to treat	
Date of start of treatment	

Treatment – Palliative Care

Hospital code		
Date of decision to treat		
Date of start of treatment		
Palliative Care Provider Type	☐ Hospital	☐ Community
Palliative Care Community Provider	☐ Hospice	☐ Nursing Home
	☐ Home care	☐ Other
	☐ Unknown	
Palliative Care Intervention Given	☐ No ☐ Yes.....Date:	

Treatment – Active Monitoring

Hospital code	
Date of decision to treat	

Outcomes

Trial status	☐ Patient eligible, consented to and entered trial		
	☐ Patient not entered into clinical trial		
	☐ Clinical trial status unknown		
Date of death			
Was death treatment-related?	☐ Yes	☐ No	☐ Unknown
Morbidity type	☐ Surgery	☐ Chemotherapy	
	☐ Radiotherapy	☐ Combination	
Was PCI given	☐ Yes	☐ No	☐ Unknown
Was the original plan carried out?	☐ Yes	☐ No	☐ Unknown
Reason for failure of original plan	☐ Cancer progressed through treatment such that a new treatment plan required		
	☐ Patient choice		
	☐ Patient died		
	☐ Treatment toxicity		
	☐ Disease progression		

Lung Cancer Nurse Specialist

Was patient assessed by LCNS?	☐ Yes	☐ No	☐ Unknown
Date of first assessment by LCNS			
How was patient first assessed by LCNS?	☐ In clinic		☐ Home visit
	☐ Ward Visit		☐ Telephone
	☐ Other		☐ Unknown
	☐ Not recorded		
At what stage was the patient assessed by LCNS?	☐ Before diagnosis		☐ After diagnosis
	☐ Before and after diagnosis		☐ At diagnosis only
	☐ Unknown		☐ Not recorded
Was LCNS present when patient received their diagnosis?	☐ Yes	☐ No	☐ Unknown

Appendix E: Code lists for surgery studies

Surgical procedure codes: OPCS-4 codes for potentially curative surgery for NSCLC

Categories in order of priority (most complicated first): Pneumonectomy (P), bi-lobectomy (B), lobectomy (L), segmentectomy / wedge resection (S), Other (O).

E391 Open excision of lesion of trachea (O)
E398 Other specified partial excision of trachea (O)
E399 Unspecified partial excision of trachea (O)
E438 Other specified other open operations on trachea (O)
E439 Unspecified other open operations on trachea (O)
E461 Sleeve resection of bronchus and anastomosis HFQ (L)
E463 Excision of lesion of bronchus NEC (O)
E468 Other specified partial extirpation of bronchus (O)
E478 Other specified other open operations on bronchus (O)
E528 Other specified other operations on bronchus (O)
E529 Unspecified other operations on bronchus (O)
E541 Total pneumonectomy (P)
E542 Bi-lobectomy of lung (B)
E543 Lobectomy of lung (L)
E544 Excision of segment of lung (S)
E545 Partial lobectomy of lung NEC (S)
E548 Other specified excision of lung (O)
E549 Unspecified excision of lung (O)
E552 Open excision of lesion of lung (O)
E558 Other specified open extirpation of lesion of lung (O)
E559 Unspecified open extirpation of lesion of lung (O)
E578 Other specified other open operations on lung (O)
E598 Other specified other operations on lung (O)
E599 Unspecified other operations on lung (O)
T013 Excision of lesion of chest wall (O)
T038 Other specified opening of chest (O)
T039 Unspecified opening of chest (O)
T058 Other specified other operations on chest wall (O)
T059 Unspecified other operations on chest wall (O)

NLCA surgical procedure codes

Categories in order of priority (most complicated first): Pneumonectomy (P), bi-lobectomy (B), lobectomy (L), segmentectomy / wedge resection (S), Other (O).

E54.4A	Wedge resection of lesion of lung (segment)	(S)
E54.8A	Multiple wedges resected	(S)
E54.4B	Segmental resection	(S)
E54.8B	Sleeve resection	(S)
E54.8 + T01	Lung resection with resection of chest wall	(O)
E44.1	Carinal resection	(S)

E54.3	Lobectomy	(L)
E54.1	Pneumonectomy	(P)
E54.2	Bi-lobectomy	(B)
E57.4	Open operation on lung (Incision of lung NEC)	(O)
E57.8	Other open operation on lung	(O)
01	Extrapleural pneumonectomy	(Excluded)
02	Debulking pleurectomy	(Excluded)
03	Pleurodesis	(Excluded)

Charlson index ICD 10 codes

Myocardial Infarction

I210 I211 I212 I213 I214 I219 I220 I221 I228 I229 I252

Congestive Heart Failure

I110 I130 I132 I500 I501 I509 I420 I425 I426 I427 I428 I429 I430 I431 I432 I438 I099

Peripheral Vascular disease

I700 I701 I702 I708 I709 I710 I711 I712 I713 I714 I715 I716 I718 I719 I731 I738 I739 I771 I790 I792 K558 K559 K551 Z958 Z959

Cerebrovascular disease

I600 I601 I602 I603 I604 I605 I606 I607 I608 I609 I610 I611 I612 I613 I614 I615 I616 I618 I619 I620 I621 I629 I630 I631 I632 I633 I634 I635 I636 I638 I639 I640 I650 I651 I652 I653 I658 I659 I660 I661 I662 I663 I664 I668 I669 I670 I671 I672 I673 I674 I675 I676 I677 I678 I679 I680 I681 I682 I688 I690 I691 I692 I693 I694 I698 G450 G451 G452 G453 G454 G458 G459 G460 G461 G462 G463 G464 G465 G466 G467 G468 H340

Dementia

F000 F001 F002 F009 F010 F011 F012 F013 F018 F019 F020 F021 F022 F023 F024 F028 F030 F051 G300 G301 G308 G309 G311

Chronic Pulmonary disease

I278 I279 J400 J410 J411 J418 J420 J430 J431 J432 J438 J439 J440 J441 J448 J449 J450 J451 J458 J459 J460 J470 J600 J610 J620 J628 J630 J631 J632 J633 J634 J635 J638 J640 J650 J660 J661 J662 J668 J670 J671 J672 J673 J674 J675 J676 J677 J678 J679 J684J701 J703

Connective Tissue disease

M050 M051 M052 M053 M058 M059 M060 M061 M062 M063 M064 M068 M069 M315 M320 M321 M328 M329 M330 M331 M332 M339 M340 M341 M342 M348 M349 M351 M353 M360

Ulcer disease

K250 K251 K252 K253 K254 K255 K256 K257 K258 K259K260 K261 K262 K263 K264 K265 K266 K267 K268 K269K270 K271 K272 K273 K274 K275 K276 K277 K278 K279K280 K281 K282 K283 K284 K285 K286 K287 K288 K289

Diabetes Mellitus

E100 E101 E109 E110 E111 E119 E120 E121 E129 E130 E131 E139 E140 E141 E149

Diabetes Mellitus with Chronic Complication

E102 E103 E104 E105 E106 E107 E108 E112 E113 E114 E115 E116 E117 E118 E122 E123 E124 E125 E126 E127 E128 E132 E133 E134 E135 E136 E137 E138 E142 E143 E144 E145 E146 E147 E148

Hemiplegia

G041 G114 G801 G802 G830 G831 G832 G833 G839 G834 G810 G811 G819G820 G821 G822 G823 G824 G825

Moderate/Severe Renal Failure

I120 N032 N033 N034 N035 N036 N037 N052 N053 N054 N055 N056 N057 N181 N182 N183 N184 N185 N189 N190 N250 Z490 Z491 Z492 Z940 Z992

Mild Liver disease

B180 B181 B182 B188 B189 K702 K703 K709 K713 K714 K715 K717 K730 K731 K732 K738 K739 K743 K744 K745 K746 Z944 K760 K700 K701 K740 K741 K742

Moderate/Severe Liver disease

K766 I850 I859 I864 I982 K711 K704 K721 K729 K765 K767

AIDS

B200 B201 B202 B203 B204 B205 B206 B207 B208 B209 B210 B211 B212 B213 B217 B218 B219 B220 B221 B222 B227 B240

Any Tumour

Excluded: C340 C341 C342 C343 C348 C349 (lung cancer)

C000 C001 C002 C003 C004 C005 C006 C008 C009 C010 C020 C021 C022 C023 C024 C028 C029 C030 C031 C039 C040 C041 C048 C049 C050 C051 C052 C058 C059 C060 C061 C062 C068 C069 C070 C080 C081 C088 C089 C090 C091 C098 C099 C100 C101 C102 C103 C104 C108 C109 C110 C111 C112 C113 C118 C119 C120 C131 C132 C138 C139 C140 C142 C148 C150 C151 C152 C153 C154 C155 C158 C159 C160 C161 C162 C163 C164 C165 C166 C168 C169 C170 C171 C172 C173 C178 C179 C180 C181 C182 C183 C184 C185 C186 C187 C188 C189 C190 C200 C210 C211 C212 C218 C220 C221 C222 C223 C224 C227 C229 C230 C240 C241 C248 C249 C250 C251 C252 C253 C254 C257 C258 C259 C260 C261 C268 C269 C300 C301 C310 C311 C312 C313 C318 C319 C320 C321 C322 C323 C328 C329 C330 C370 C380 C381 C382 C383 C384 C388 C390 C398 C399 C400 C401 C402 C403 C408 C409 C410 C411 C412 C413 C414 C418 C419 C431 C432 C433 C434 C435 C436 C437 C438 C439 C450 C451 C452 C457 C459 C460 C461 C462 C463 C467 C468 C469 C470 C471 C472 C473 C474 C475 C476 C478 C479 C480 C481 C482 C488 C490 C491 C492 C493 C494 C495 C496 C498 C499 C500 C501 C502 C503 C504 C505 C506 C508 C509 C510 C511 C512 C518 C519 C520 C530 C531 C538 C539 C540 C541 C542 C543 C548 C549 C550 C560 C570 C571 C572 C573 C574 C577 C578 C579 C580 C600 C601 C602 C608 C609 C610 C620 C621 C629 C630 C631 C632 C637 C638 C639 C640 C650 C660 C670 C671 C672 C673 C674 C675 C676 C677 C678 C679 C680 C681 C688 C689 C690 C691 C692 C693 C694 C695 C696 C698 C699 C700 C701 C709 C710 C711 C712 C713 C714 C715 C716 C717 C718 C719 C720 C721 C722 C723 C724 C725 C728 C729 C730 C740 C741 C749 C750 C751 C752 C753 C754 C755 C758 C759 C760 C761 C762 C763 C764 C765 C767 C768

Metastatic Solid Tumor

C770 C771 C772 C773 C774 C775 C778 C779 C780 C781 C782 C783 C784 C785 C786 C787 C788 C790 C791 C792 C793 C794 C795 C796 C797 C798 C799 C800 C809

Leukemia

C910 C911 C913 C914 C915 C916 C917 C918 C919 C920 C921 C922 C923 C924 C925 C926 C927 C928 C929 C930 C931 C933 C937 C939 C940 C942 C943 C944 C946 C947 C950 C951 C957 C959 D450

Lymphoma

C810 C811 C812 C813 C814 C817 C819 C820 C821 C822 C823 C824 C825 C826 C827 C829 C830 C831 C833 C835 C837 C838 C839 C840 C841 C844 C845 C846 C847 C848 C849 C851 C852 C857 C859 C880 C900 C901 C902 C903 C960 C962 C964 C965 C966 C967 C968 C969

Appendix F: Code lists for Chemotherapy studies

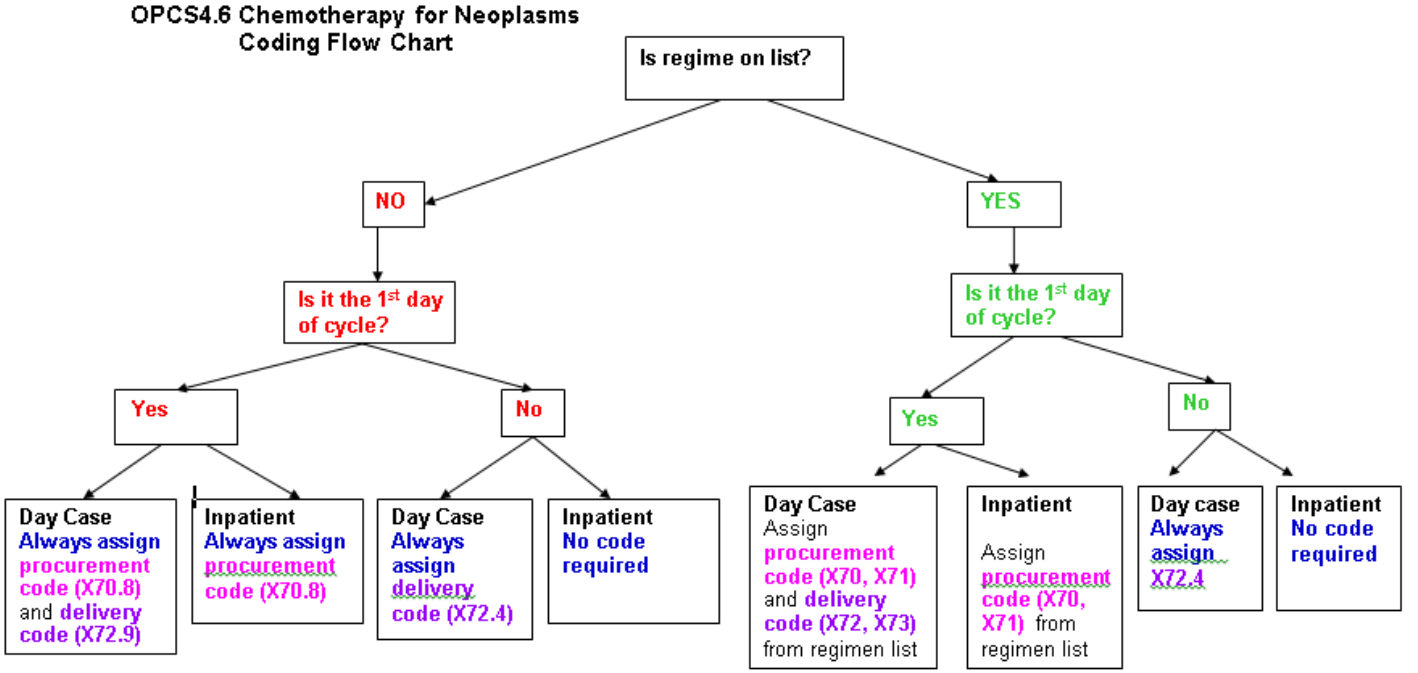
OPCS-4 chemotherapy codes

Code	Definition
X35.2	Intravenous chemotherapy –only chemotherapy code available until 1/4/2006
X72.1	Delivery of complex chemotherapy for neoplasm including prolonged infusional treatment at first attendance
X72.2	Delivery of complex parenteral chemotherapy for neoplasm at first attendance
X72.3	Delivery of simple parenteral chemotherapy for neoplasm at first attendance
X72.4	Delivery of subsequent element of cycle of chemotherapy for neoplasm
X72.8	Other specified delivery of chemotherapy for neoplasm
X72.9	Unspecified delivery of chemotherapy for neoplasm
X70.1	Procurement of drugs for chemotherapy for neoplasm for regimens in Band 1
X70.2	Procurement of drugs for chemotherapy for neoplasm for regimens in Band 2
X70.3	Procurement of drugs for chemotherapy for neoplasm for regimens in Band 3
X70.4	Procurement of drugs for chemotherapy for neoplasm for regimens in Band 4
X70.5	Procurement of drugs for chemotherapy for neoplasm for regimens in Band 5
X70.8	Other specified procurement of drugs for chemotherapy for neoplasm Bands 1-5
X70.9	Unspecified procurement of drugs for chemotherapy for neoplasm in Bands 1-5
X71.1	Procurement of drugs for chemotherapy for neoplasm for regimens in Band 6
X71.2	Procurement of drugs for chemotherapy for neoplasm for regimens in Band 7
X71.3	Procurement of drugs for chemotherapy for neoplasm for regimens in Band 8
X71.4	Procurement of drugs for chemotherapy for neoplasm for regimens in Band 9
X71.5	Procurement of drugs for chemotherapy for neoplasm for regimens in Band 10
X71.8	Other specified procurement of drugs for chemotherapy for neoplasm Bands 6-10
X71.9	Unspecified procurement of drugs for chemotherapy for neoplasm in Bands 6-10

OPCS-4 radiotherapy codes

Code	Description
X651	Delivery of a fraction of total body irradiation
X652	Delivery of a fraction of intracavitary radiotherapy
X653	Delivery of a fraction of interstitial radiotherapy
X654	Delivery of a fraction of external beam radiotherapy NEC
X656	Delivery of a fraction of intraluminal brachytherapy
X658	Other specified radiotherapy delivery
X659	Unspecified radiotherapy delivery
Y918	Other specified Delivery of Radiotherapy
Y919	Unspecified Delivery of Radiotherapy
X671	Preparation for intensity modulated radiation therapy
X672	Preparation for total body irradiation
X673	Preparation for hemi body irradiation
X674	Preparation for simple radiotherapy with imaging and dosimetry
X675	Preparation for simple radiotherapy with imaging and simple calculation
X676	Preparation for superficial radiotherapy with simple calculation
X677	Preparation for complex conformal radiotherapy
X678	Other specified preparation for external beam radiotherapy
X679	Unspecified preparation for external beam radiotherapy
X681	Preparation for intraluminal brachytherapy
X682	Preparation for intracavitary brachytherapy
X683	Preparation for interstitial brachytherapy
X688	Other specified preparation for brachytherapy
Y921	Technical support for preparation for radiotherapy

Flow chart for clinical coding at Nottingham University Hospitals



- There is no change in the assignment of ICD10 code Z51.1
- Inpatient attendances on 1st day of cycle require an OPCS4.6 **procurement code (X70, X71)** only
- Day case attendances on 1st day of cycle require an OPCS4.6 **procurement code (X70, X71)** and a **delivery code (X72, X73)**
- OPCS4.6 code **X72.4** is a generic code for all subsequent deliveries within a cycle for regimes which appear on the list **(Day cases only).**
- All cycles should be coded.
- Where a regime does not follow a cycle (very unusual) but continues for a long time the **procurement code** should be assigned once a month (inpatient).

Appendix G: Study protocols and documents

Proposal for analysis of Danish Lung Cancer Registry (DCLR) surgical data

Introduction

In pre-operative assessment of mortality risk the commonly used Thoracoscore was not developed using data solely on people with lung cancer, nor has it been validated in such a population.(170) There is concern in the lung cancer community that Thoracoscore may under (or over-) estimate risks when used for people with lung cancer, and there is enthusiasm for a more sophisticated score

Using English National Lung Cancer Audit (NLCA) data, supplemented with data from inpatient hospital episodes the Nottingham group has produced a score (Table 1) which is designed to estimate the risk of death within 90 days of surgery for lung cancer. (200) This has not been tested in an independent dataset.

Study population

The score was based on all patients in the NLCA with confirmed or presumed NSCLC (in the NLCA if histology data are not entered the patient is presumed to have NSCLC) who had a surgical procedure which, in a patient with lung cancer, could reasonably represent an attempt at cure.

Procedures which took place between 1st January 2004 and 31st March 2010 were included. A list of procedure codes is attached as an appendix. People with stage 3b or 4 disease were excluded as surgery for these people would not be curative. People with missing FEV1, stage or performance status were also excluded.

Variables

The score comprises the following patient, tumour and procedure related variables:

Age	At diagnosis as surrogate for age at surgery
Sex	M/F
Performance status	As defined by ECOG
FEV_1	Percentage of predicted
Procedure type	Pneumonectomy; (bi-)lobectomy, wedge or segmentectomy; other (see attached code list).
Charlson co-morbidity index	See below
Stage	UICC TNM version 6 lung cancer stage, see text

The method of calculating co-morbidity score may require some discussion. The NLCA score was developed using the original Charlson co-morbidities and weighting, identified through coding from inpatient hospital episodes which took place any time before the procedure date.(90) The only exception was that lung cancer was not included in 'tumour'.

It is not possible to convert between UICC TNM staging versions 6 and 7 with the information available in the NLCA database (during 2010 clinicians in the UK started to use version 7 rather than version 6). If the same is true for the DCLR it would probably be necessary to accept this as a limitation to the validation study and use stage regardless of TNM system accepting the minor differences.

Outcome variables

Death within 90 days of surgery requires the date of procedure and date of death, or a censor date at least 90 days after the latest procedure date.

Table 1: NLCA predictive score for 90 day mortality after lung cancer surgery

		Coefficient
Age (years)	<55	0
	55-65	0.31
	66-75	0.97
	>75	1.40
Sex	Female	0
	Male	0.23
Performance status	0	0
	1-2	0.68
	≥3	0.21‡
% predicted	>80%	0
FEV_1	61-80%	0.20
	40-60%	0.69
	<40%	0.95
Procedure type	Pneumonectomy	1.16
	(Bi-)lobectomy, wedge, or segmentectomy	0
	Other [b]	0.07
Charlson	0-1	0
index	≥2	0.33
Stage	1a	0
	1b	0.42
	2a or 2b	0.51
	3a	0.84
	Constant	**-5.28**

FEV1 forced expiratory volume in 1 second; ǂ Only 40 patients and 2 deaths in this group; b Other includes procedures listed in Appendix. See text for method of calculating percentage risk of death within 90 days.

Plan for statistical analysis

The NLCA score was developed using a multivariate logistic regression model. The risk of death within 90 days of surgery, as a percentage, for an individual patient is estimated as:

odds / (1+ odds) where odds = exp (total of coefficients + constant)

This is the same methodology as used in Thoracoscore.(170)

We would use ROC modelling to test the performance of the score in the DCLR (which from 2005 onwards contains >90% of people diagnosed with lung cancer in Denmark) complemented with data from the Central Population Register, National Pathology Registry and National Hospital Register. We would restrict the study population to those with complete data for all components of the score and would use all available data from 2005 onwards.

We would also use multivariate logistic regression to produce a score using the DCLR data. This would include all variables with significant univariate associations with 90 day mortality, which remain significant in a multivariate model. We would include additional variables such as height and weight (which would be used to calculate BMI) if they had significant associations in the multivariate model. We could then test the performance of this score in the English data using ROC modelling.

Power

In order to achieve 90% power at the 0.05 significance level, based on the age variable (< or >70 years in which groups the 90-day post-operative mortality was 4% and 7% respectively) we would require data on 2,556 people who had undergone surgery for lung cancer.

The DLCR data from 2005 to 2010 (inclusive) contains at least 3,152 people with NSCLC who underwent surgical resection. (213) Assuming the 90-day mortality is approximately the same as that in England, even if some of these people are excluded due to incomplete data, we will have sufficient power to assess whether the actual outcomes were significantly different to those predicted by the score.

H Powell, M Luchtenbörg & R Hubbard – August 2013

Qualitative study to map attitudes to risks surrounding treatment for lung cancer – study documents

Schematic diagram of study design – patient interviews

1. Members of MDTs (lung cancer nurse specialists, surgeons, respiratory physicians and oncologists) asked to assist with recruitment.

2. Study, including inclusion and exclusion criteria, explained to members of lung cancer MDTs.

3. Suitable patients invited to participate by a member of the MDT (who is also part of their clinical care team), given an invitation letter and participant information sheet and asked if they would be happy to be contacted by a researcher.

4. Patients who agree to be contacted receive a telephone call from a researcher who answers questions about the study and, if the patient gives verbal consent to participate, arranges a time and place for interview.

5. Participant allocated unique study identifier

6. Patient signs consent form in presence of investigator (at least 24 hours after step 3)

7. INTERVIEW takes place, audio recorded. No further patient participation required from this point.

8. Researcher uses patient's medical notes to complete case report form.

9. Data entered into database and stored with audio recordings on password protected UoN computer. All identifiers apart from unique study identifier removed.

10. Audio-recordings transcribed by external transcription company

11. Transcripts analysed as per study protocol for analysis. Analysis of transcripts looking for saturation of themes will determine when recruitment stops.

Letter to MDT members asking for assistance with recruitment

(Final version 1.1 08/05/12)

Study title: A qualitative study to map attitudes to risks surrounding treatment for lung cancer.
Research team: The University of Nottingham, Department of Public Health and Epidemiology
Names of Researchers: Professor Richard Hubbard, Dr Helen Powell, Dr Laura Jones, Dr Manpreet Bains, Dr Laila Tata and Dr David Baldwin.

Dear MDT member,

We are writing to ask for your assistance in recruiting patients for our study. We would like to talk to patients who have recently been diagnosed with lung cancer, to explore their opinions about the risks associated with treatment for lung cancer, and in particular with surgery.

The enclosed information sheet and letter of invitation explain the study in full. We would be grateful if you would inform your patients about the study if they meet the inclusion criteria and you feel may agree to participate. If they express an interest please give them the enclosed information pack, confirm that they agree to be contacted by a researcher and give us their details so that we can contact them to discuss the study further.

Inclusion criteria for the study are: any patient recently diagnosed with lung cancer stage 1a to 3a (inclusive), who is aware of their diagnosis, and has not yet had or is not going to have surgery. They must be over 18 years of age, able to give informed consent and able to communicate (hear, speak and understand) in English without an interpreter. Patients will be interviewed in their own home by one of the researchers after we have obtained informed consent. After the interview we will collect some information on tumour stage, co-morbidity and treatment plan from the patient's hospital notes to assist with interpreting our data.

If you would like any further information, please contact Helen Powell, who is a member of the research team on 0115 8231378 or at helen.powell@nottingham.ac.uk.

Many thanks for your assistance.
Yours faithfully,

Helen Powell
Clinical Research Fellow, University of Nottingham

Invitation letter to patients

(Final version 1.1 08/05/12)

Study title: A qualitative study to map attitudes to risks surrounding treatment for lung cancer.
Research team: The University of Nottingham, Department of Public Health and Epidemiology
Names of Researchers: Professor Richard Hubbard, Dr Helen Powell, Dr Laura Jones, Dr Manpreet Bains, Dr Laila Tata and Dr David Baldwin.

Dear Patient,

Thank you for thinking about taking part in our study. We would like to talk to people, like you, who have recently been diagnosed with lung cancer, to find out how patients feel about the treatments for lung cancer and in particular the risks associated with surgery.

The enclosed information explains the study in full and you should read it carefully before deciding if you would like to take part.

If you decide that you would like to be involved in the study, you will be invited to take part in an individual interview with a researcher at a time convenient to you. The interview will take about an hour and you can choose whether the researcher comes to your house or whether you are interviewed at Nottingham City hospital. This informal one to one discussion will focus on lung cancer and your feelings about treatment.

The interview will be audio recorded to allow the researcher to pay full attention to what you are saying. Recording the interview will also allow the research team to do further analysis at a later date. In addition, we will ask for your permission to collect a few pieces of information from your medical notes. This information and the audio-recordings will be kept strictly confidential, stored securely within our department and only used for the purposes of the study.

Thank you for taking the time to read this letter. If you would like any further information, you can talk to your lung cancer nurse specialist or Helen Powell, who is a member of the research team on 0115 8231378 or at helen.powell@nottingham.ac.uk.

Yours faithfully,

Helen Powell
Clinical Research Fellow, University of Nottingham

Participant information sheet – patients

The University of Nottingham

Nottingham University Hospitals NHS
NHS Trust

Participant Information Sheet for Patients
(Final version 1.0 13/02/12)

Title of Study: A qualitative study to map attitudes to risks surrounding treatment for lung cancer.

Researchers: Professor Richard Hubbard, Dr Helen Powell, Dr Laura Jones, Dr Manpreet Bains, Dr Laila Tata and Dr David Baldwin.

We would like to invite you to take part in our research study. Before you decide whether to participate we would like you to understand why the research is being done and what it would involve for you. One of our team will go through the information sheet with you and answer any questions you have. Talk to others about the study if you wish. Ask us if there is anything that is not clear.

What is the purpose of the study?

We are conducting this study to gather information about the views and attitudes of patients with lung cancer towards the risks associated with their treatment. Scientists have researched which sorts of treatment work best for lung cancer and many teams are working on developing new treatments. Treatments, particularly surgery, do however have a risk associated with them and we feel it is important for lung cancer specialists to know how patients feel about these risks. This information will help healthcare professionals when they are discussing treatment with their patients in the future and may improve access to treatment for some patients.

Why have I been invited?

You are being invited to take part because you have recently been diagnosed with lung cancer. We are initially inviting 20 people like you to take part but depending on the results may need up to 60 participants.

Do I have to take part?

Participation in this study is entirely voluntary. Once you have read this information sheet, and asked any questions that you might have, if you decide to take part, you will be asked to sign a consent form. You may withdraw from the study at any time and without giving a reason. This would not affect your legal rights. If you decide not to take part, or to withdraw, this will not affect your medical care or treatment in any way.

What will happen to me if I take part?

A researcher will contact you to arrange a time and place which is convenient for you for an interview. A researcher will usually come to your home to conduct the interview but if you would prefer to come to Nottingham City Hospital (address below) we can make arrangements to hold the interview there. We aim to have completed the interview within four weeks of your doctor or nurse first telling you about the study.

The interview is informal and will focus on how you feel about lung cancer and the treatments you are going to receive or other treatments which you may have discussed with your doctor.

We expect the interview to last about one hour and after this your involvement in the study will be complete.

All interviews will be digitally audio-recorded, which will allow the researcher to pay full attention to what you are saying. Recordings will be transcribed by an external company prior to analysis by the research team at the University of Nottingham.

After the interview, one of our researchers will review your medical notes and record your age, sex, type of lung cancer, any other medical problems that you have, what treatment (if any) is planned for you and which doctors you have spoken to about your diagnosis and treatment.

Expenses and payments

Participants will not be paid to participate in the study.

What are the possible disadvantages and risks of taking part?

We know that this is a difficult time for you having recently been told that you have lung cancer, and we are asking you to discuss your feelings about aspects of this disease and your treatment. The interviewer has experience of working with patients who have cancer and will try to ensure that questions are asked in a sensitive manner. If you are unsure whether to participate for this reason may we suggest that you discuss it with a relative, friend or your lung cancer nurse specialist.

What are the possible benefits of taking part?

If you agree to take part, you will assist us in identifying ways to help future patients with lung cancer to get better information from their doctors and, perhaps, better access to treatment.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. The researchers contact details are given at the end of this information sheet. If you remain unhappy and wish to complain formally, you can do this by contacting the Patient Advice and Liaison Service (PALS), details of which can be obtained from your hospital.

You may find that the interview makes you think of questions about lung cancer or your treatment that you have not asked before. The interviewer will be unable to comment or advise you on your care, but if you do have questions or wish to discuss your care

further, they can direct you to the most appropriate person (this will usually be your lung cancer nurse specialist at the hospital where you were diagnosed).

Will my taking part in the study be kept confidential?

We will inform your lung cancer nurse specialist that you are taking part in the study. We will not inform your general practitioner.

We will follow ethical and legal practice and all information about you will be handled in confidence. If you participate in the study, the data collected for the study (including information copied from your medical records) will be looked at by the authorised persons from the University of Nottingham who are organising the research. They may also be looked at by authorised people to check that the study is being carried out correctly. All will have a duty of confidentiality to you as a research participant and we will do our best to meet this duty.

All information which is collected about you during the course of the research will be kept **strictly confidential**, stored in a secure and locked office, and on a password protected database. Any information about you which leaves the hospital will have your name and address removed (anonymised) and a unique code will be used so that you cannot be recognised from it.

Your personal data (name, address, telephone number) will be kept until all interviews are complete and the data has been analysed. After this we will destroy these records. All other research data will be kept securely for seven years then disposed of securely. During this time all precautions will be taken by all those involved to maintain your confidentiality, only members of the research team will have access to your personal data.

During interviews participants very occasionally tell the researcher (interviewer) something which may be sensitive. Depending on the nature of the disclosure, the researcher may feel it is appropriate to report this to the appropriate authorities or medical team.

What will happen if I don't want to carry on with the study?

Your participation is voluntary and you are free to withdraw at any time, without giving any reason, and without your legal rights being affected. If you withdraw, then the information collected so far cannot be erased and this information may still be used in the project analysis. Your medical treatment will not be affected in any way if you withdraw from the study.

What will happen to the results of the research study?

The results of the study will be collated and conclusions drawn. We expect to have a summary of the results available about 18 months after the final interviews are completed. Individual participants will not be identified in the published results but anonymous quotes may be included to demonstrate particular points. Please let us know

if you would like to receive a summary of these results, bearing in mind that it may be sometime after your participation.

The study is also part of a PhD project being undertaken by Dr Helen Powell and will be included in her PhD thesis which is due to be completed in August 2014.

Who is organising and funding the research?

This research is being organised by the University of Nottingham and is being funded by the National Institute for Health Research (NIHR) through the Nottingham Respiratory Biomedical Research Unit (BRU).

Who has reviewed the study?

All research in the NHS is looked at by independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by Nottingham Research Ethics Committee.

Further information and contact details

For further information please contact:

Dr Helen Powell
Clinical Sciences Building
Nottingham City Hospital
Hucknall Road
Nottingham
NG5 1PB
0115 8231378
helen.powell@nottingham.ac.uk

The Chief Investigator for this study is:

Professor Richard Hubbard.
Clinical Sciences Building
Nottingham City Hospital
Hucknall Road
Nottingham
NG5 1PB
0115 8231385
richard.hubbard@nottingham.ac.uk

Consent form for patients

The University of Nottingham

Nottingham University Hospitals NHS
NHS Trust

CONSENT FORM FOR PATIENTS
(Final Version 1.0: 13/02/2012)

Title of Study: A qualitative study to map attitudes to risks surrounding treatment for lung cancer.

REC ref: 12/EM/0123

Name of Researcher: _____

Name of Participant: _____ **Please initial box**

Study Identifier: _____

1. I confirm that I have read and understand the information sheet version number 1.0 dated 13/02/12 for the above study and have had the opportunity to ask questions. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without my medical care or legal rights being affected. I understand that should I withdraw then the information collected so far cannot be erased and that this information may still be used in the project analysis. ☐

3. I understand that relevant sections of my medical notes and data collected in the study may be looked at by authorised individuals from the University of Nottingham, the research group and regulatory authorities where it is relevant to my taking part in this study. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential. ☐

4. I understand that the interview will be recorded and that anonymous direct quotes from the interview may be used in the study reports. ☐

5. I agree to take part in the above study and give the researchers permission to inform my lung cancer nurse specialist that I am doing so. ☐

Name of Participant	Date	Signature
Name of Person taking consent	Date	Signature

2 copies: 1 for participant, 1 for the project notes

Letter to lung cancer clinical nurse specialists

(Final version 1.0 13/02/12)

Study title: A qualitative study to map attitudes to risks surrounding treatment for lung cancer.
Research team: The University of Nottingham, Department of Public Health and Epidemiology
Names of Researchers: Professor Richard Hubbard, Dr Helen Powell, Dr Laura Jones, Dr Manpreet Bains, Dr Laila Tata and Dr David Baldwin.

Dear colleague,

Re: Patient name:

Date of Birth:

Hospital number:

We are writing to inform you that your patient has agreed to take part in the above study.

The enclosed participant information sheet explains the study in full. We do not need you to do anything in response to this letter, however it is possible that as a result of taking part in an interview your patient may wish to discuss aspects of their diagnosis or treatment with you or a member of your team and we are very grateful to you for facilitating this.

If you would like any further information, please contact Helen Powell, who is a member of the research team on 0115 8231378 or at helen.powell@nottingham.ac.uk.

Many thanks for your assistance.

Yours faithfully,

Helen Powell
Clinical Research Fellow, University of Nottingham

Initial interview guide for patients

Introduction

- Thank them for coming and taking part.
- Check consent form has been signed. Keep a copy and give participant a copy.
- Statement on confidentiality, right to withdraw consent, recording of the interview.
- Explain the purpose of the study in general:
- To explore attitudes towards treatment in lung cancer
- Their opinions on risks of treatment
- Ask if they have any further questions before starting interview
- The interview will last between 30 and 60 minutes.

Background

- Tell me a bit about yourself
    - Do you have family nearby?
    - Do you work?
- You are taking part in this study because you have recently been diagnosed with lung cancer, can you tell me about what happened when you were diagnosed?
    - When?
    - What tests did you have?
    - Who broke the news?
    - Were you expecting it?
    - How did you feel?
    - How did your family feel?

- What is going to happen now?
    - What is the plan for treatment or more investigations?
    - What do you want to happen?
    - When is your next consultation?
    - Are you happy with the plan?

Knowledge

- What do you know about lung cancer and the possible treatments?
    - What did you know at the time you were diagnosed?
    - How have you found out what you know?
    - What have the hospital / your GP told you?
    - What treatments are you aware of?
    - Were you offered a choice of treatments?
    - Did you understand what they told you?
    - Did you ask many questions?
    - Did they talk about prognosis? Did you want them to?
    - Who was most helpful?
    - Do you have enough information now?

Risks & communication of risks

- What sort of problems do you know about that can arise from lung cancer treatments?
- What sort of risks did your doctors and nurses tell you about?
- How do you feel about the possibility that something might go wrong?
- Discuss risk in the context of the treatment they are going to have, or were offered.

- If the risk had been quoted as X instead of Y what would your thoughts be?
- Present and ask them to discuss different scenarios including different degrees of mortality risk, survival and post-op disability.

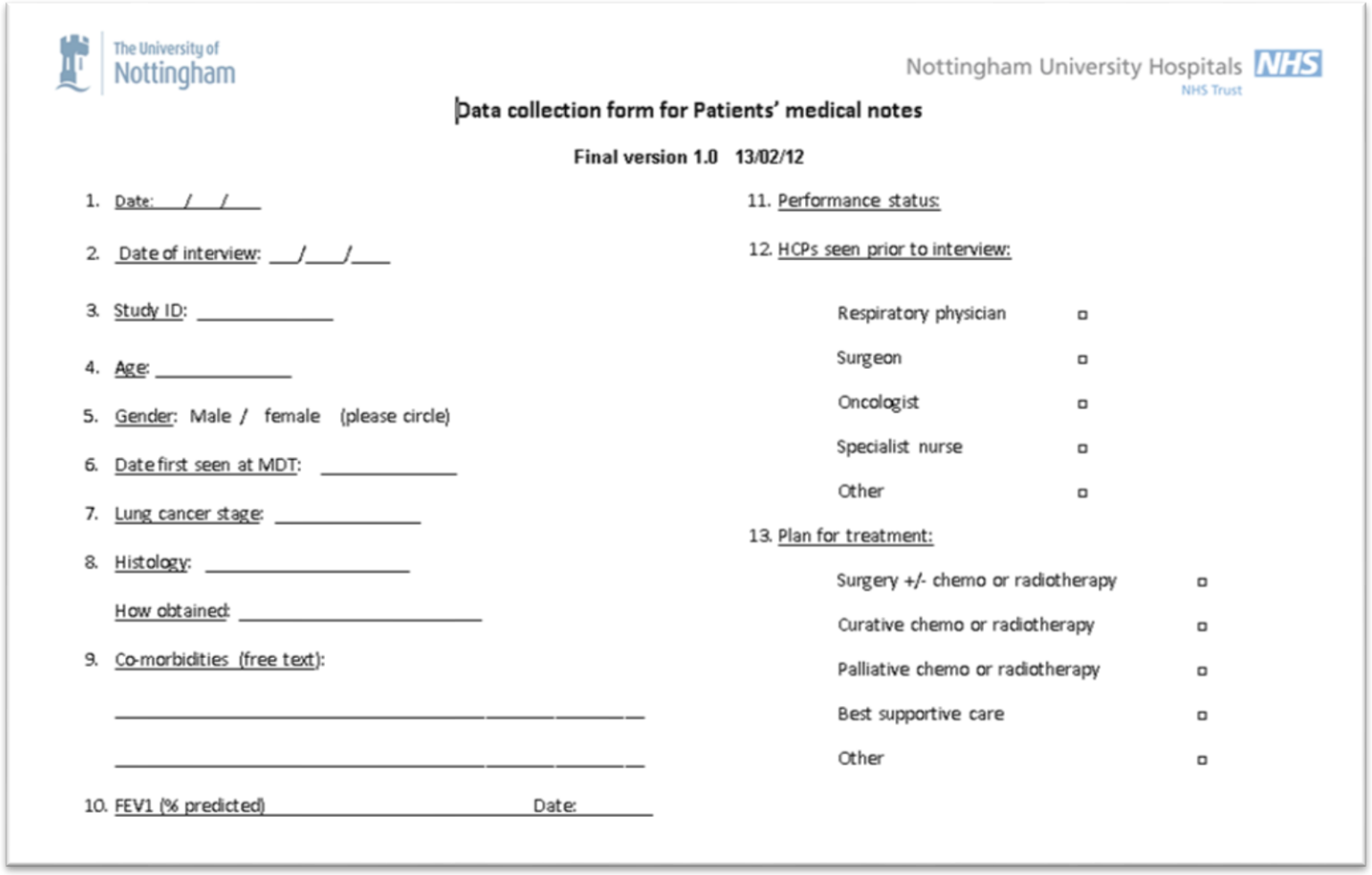
Closing questions

- Has this interview raised issues which you haven't considered before?
  - What are they?
  - Will you want more information?
- Advise them that if they think of any further questions about their diagnosis or treatment they can contact their lung cancer specialist nurse.

Conclusion

- Tell the patient that they have reached the end of the interview
- Do they have any questions in return?
- Remind them about confidentiality.
- Thank them for their time.

Case report form



The University of Nottingham

Nottingham University Hospitals NHS
NHS Trust

Data collection form for Patients' medical notes

Final version 1.0 13/02/12

1. Date: ___/___/___
2. Date of interview: ___/___/___
3. Study ID: _____
4. Age: _____
5. Gender: Male / female (please circle)
6. Date first seen at MDT: _____
7. Lung cancer stage: _____
8. Histology: _____

   How obtained: _____
9. Co-morbidities (free text):

   _____

   _____
10. FEV1 (% predicted) _____ Date: _____
11. Performance status:
12. HCPs seen prior to interview:

    Respiratory physician ☐
    Surgeon ☐
    Oncologist ☐
    Specialist nurse ☐
    Other ☐
13. Plan for treatment:

    Surgery +/- chemo or radiotherapy ☐
    Curative chemo or radiotherapy ☐
    Palliative chemo or radiotherapy ☐
    Best supportive care ☐
    Other ☐

Schematic diagram of study design – healthcare professionals

1. Representatives of professional bodies (SCTS, ARNS and BTS) asked to assist with recruitment

2. Study inclusion and exclusion criteria explained to representatives.

3. HCPs invited to participate by a member of their professional body and given a participant information sheet.

4. HCPs who wish to participate contact the named investigator who answers any additional questions and arranges a time and place for interview.

5. Participant allocated unique study identifier

6. Participant signs consent form in presence of investigator to confirm informed consent to participate (at least 24 hours after step 3)*

7. Participant completes questionnaire (on day of interview)*

8. INTERVIEW takes place, audio recorded. No further participation from individual HCP required after this point.

9. Data entered into database and stored with audio recordings on password protected UoN computer. All but unique study identifier removed.

10. Audio-recordings transcribed by external transcription company

11. Transcripts analysed as per study protocol for analysis. Analysis of transcripts looking for saturation of themes will determine when recruitment stops.

* If the interview is not face-to-face the participant will be sent a consent form and questionnaire which they will sign and return by post prior to the interview.

Abbreviations: ARNS Association of Respiratory Nurse Specialists; BTS British Thoracic Society; HCP Healthcare Professional; MDT Multi-Diciplinary Team; SCTS Society of Cardiothoracic Surgeons; UoN University of Nottingham;

Letter to representatives of professional bodies asking for assistance with recruitment

(Final version 1.0 13/02/12)

Study title: A qualitative study to map attitudes to risks surrounding treatment for lung cancer.
Research team: The University of Nottingham, Department of Public Health and Epidemiology
Names of Researchers: Professor Richard Hubbard, Dr Helen Powell, Dr Laura Jones, Dr Manpreet Bains, Dr Laila Tata and Dr David Baldwin.

Dear Colleague,

We are writing to ask for your assistance in recruiting participants to our study. We would like to interview healthcare professionals who are involved in treatment decisions, particularly those involving surgery, for patients with lung cancer, in order to investigate how and why opinions and practice vary.

Inclusion criteria for the study are that the participant is aged over 18 years old, able to give informed consent and can communicate in English. Healthcare professionals must be employed by the NHS and involved in caring for patients with lung cancer, in particular in contributing to the decision whether or not the patient will be offered surgery.

The enclosed information sheet and letter of invitation explain the study in full and we would be grateful if you would distribute this to members of your professional society who you feel meet our inclusion criteria and may agree to participate.

If you would like any further information please contact Helen Powell, who is a member of the research team, on 0115 8231378 or at helen.powell@nottingham.ac.uk.

Yours faithfully,

Helen Powell
Clinical Research Fellow, University of Nottingham.

Invitation letter to healthcare professionals

(Final version 1.0 13/02/12)

Study title: A qualitative study to map attitudes to risks surrounding treatment for lung cancer.
Research team: The University of Nottingham, Department of Public Health and Epidemiology
Names of Researchers: Professor Richard Hubbard, Dr Helen Powell, Dr Laura Jones, Dr Manpreet Bains, Dr Laila Tata and Dr David Baldwin.

Dear Colleague,

Thank you for considering taking part in our study. We would like to interview healthcare professionals who are involved in treatment decisions, particularly those involving surgery, for patients with lung cancer, in order to investigate how and why peoples' opinions and practice vary.

The enclosed information sheet explains the study in full and you should read it carefully before deciding if you would like to take part.

If you decide that you would like to be involved in the study, you will be invited to take part in an individual interview with a researcher at a time convenient to you. The interview will take about an hour. We will endeavour to find a location which is convenient for you, but if this is not possible and you would be willing to take part in a telephone or video interview this would also be an option. The interview will be an informal one-to-one discussion focusing on treatment for lung cancer and your opinions about the risks.

The interview will be audio recorded then transcribed for analysis. The audio-recordings and transcripts will be kept strictly confidential, stored securely within our department and only used for the purposes of the study.

Thank you for taking the time to read this letter. If you would like any further information please contact Helen Powell, who is a member of the research team, on 0115 8231378 or at helen.powell@nottingham.ac.uk.

Yours faithfully,

Helen Powell
Clinical Research Fellow, University of Nottingham.

Participant Information Sheet for Healthcare Professionals
(Final version 1.0 13/02/12)

Title of Study: A qualitative study to map attitudes to risks surrounding treatment for lung cancer.

Researchers: Professor Richard Hubbard, Dr Helen Powell, Dr Laura Jones, Dr Manpreet Bains, Dr Laila Tata and Dr David Baldwin.

We would like to invite you to take part in our research study. Before you decide, we would like you to understand why the research is being done and what it would involve for you. Please ask us if there is anything that is not clear.

What is the purpose of the study?

The purpose of this study is to accumulate evidence about the views and attitudes of patients and healthcare professionals towards the risks associated with treatment, particularly surgery, for lung cancer. This study will involve interviewing thoracic surgeons, lung cancer nurse specialists and respiratory physicians; we are also carrying out interviews with lung cancer patients.

Why have I been invited?

You are being invited to take part because you are involved in deciding whether patients are offered treatment, particularly surgery, for lung cancer. We are initially inviting 20 National Health Service (NHS) healthcare professionals to take part, but due to the qualitative nature of the study may require more participants depending on our results.

Do I have to take part?

Participation in this study is entirely voluntary. Once you have read this information sheet, and asked any questions that you might have, if you do decide to take part you will be asked to sign a consent form. You may withdraw from the study at any time and without giving a reason. This would not affect your employment or your legal rights.

What will happen to me if I take part?

You will be contacted by one of our researchers who will arrange a time and place for an interview. You can decide whether the interview takes place in your own or work time; if in work time please agree this with your line manager. The interview can be carried out at your place of work, at a conference facility, at the University of Nottingham or by telephone or video-call.

The interview will be informal and will focus on your attitudes towards the risks associated with surgery for lung cancer patients. Interviews will be digitally audio-recorded, which will allow the researcher to pay full attention to what you are saying. Recordings will be transcribed by an external company prior to analysis by the research

team at the University of Nottingham. We expect interviews to last approximately one hour and after the interview your participation in the study will be complete.

Expenses and payments

Participants will not be paid to participate in the study.

What are the possible disadvantages and risks of taking part?

We do not envisage any disadvantage or risk for you due to participating in this study, although we will require approximately 90 minutes of your time to complete the consent form and the interview.

What are the possible benefits of taking part?

The information we collate is expected to give clinicians a better understanding of what their patients see as acceptable level of risk in terms of treatment for lung cancer. We hope that this will lead to better communication about risk between patients and healthcare professionals and that it will promote more uniform access to treatment for lung cancer.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. The researchers' contact details are given at the end of this information sheet.

Will my taking part in the study be kept confidential?

We will follow ethical and legal practice and all information about you will be handled in confidence. If you join the study, some parts of the data collected for the study will be looked at by authorised persons from the University of Nottingham who are organising the research. They may also be looked at by authorised people in the process of auditing. All will have a duty of confidentiality to you as a research participant and we will do our best to meet this duty.

All information which is collected about you during the course of the research will be kept **strictly confidential**, stored in a secure and locked office, and on a password protected database. Any information about you will have your name and address removed (anonymised) and a unique code will be used so that you cannot be recognised from it.

Your personal data (name and contact details) will be kept securely until data analysis is complete. After this these records will be destroyed. All other anonymised research data will be kept for seven years then disposed of securely. During this time, all precautions will be taken by all those involved to maintain your confidentiality, only members of the research team will have access to your personal data.

During interviews participants very occasionally tell the researcher (interviewer) something which may be sensitive. Depending on the nature of the disclosure, the researcher may feel it is appropriate to report this to the appropriate authorities or medical team.

What will happen if I don't want to carry on with the study?

Your participation is voluntary and you are free to withdraw at any time, without giving any reason, and without your legal rights being affected. If you withdraw then the information collected so far cannot be erased and this information may still be used in the project analysis.

What will happen to the results of the research study?

The results of the study will be collated and conclusions drawn. We expect to have a summary of the results available about 18 months after the final interviews are completed. Individual participants will not be identified in the published results but anonymous quotes may be included to demonstrate particular points. Please let us know if you would like to receive a summary of these results.

The study is also part of a PhD project being undertaken by Dr Helen Powell and will be included in her PhD thesis which is due to be completed in August 2014.

Who is organising and funding the research?

This research is being organised by the University of Nottingham and is being funded by the National Institute for Health Research (NIHR) through the Nottingham Respiratory Biomedical Research Unit (BRU).

Who has reviewed the study?

All research in the NHS is looked at by independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by Nottingham Research Ethics Committee.

Further information and contact details

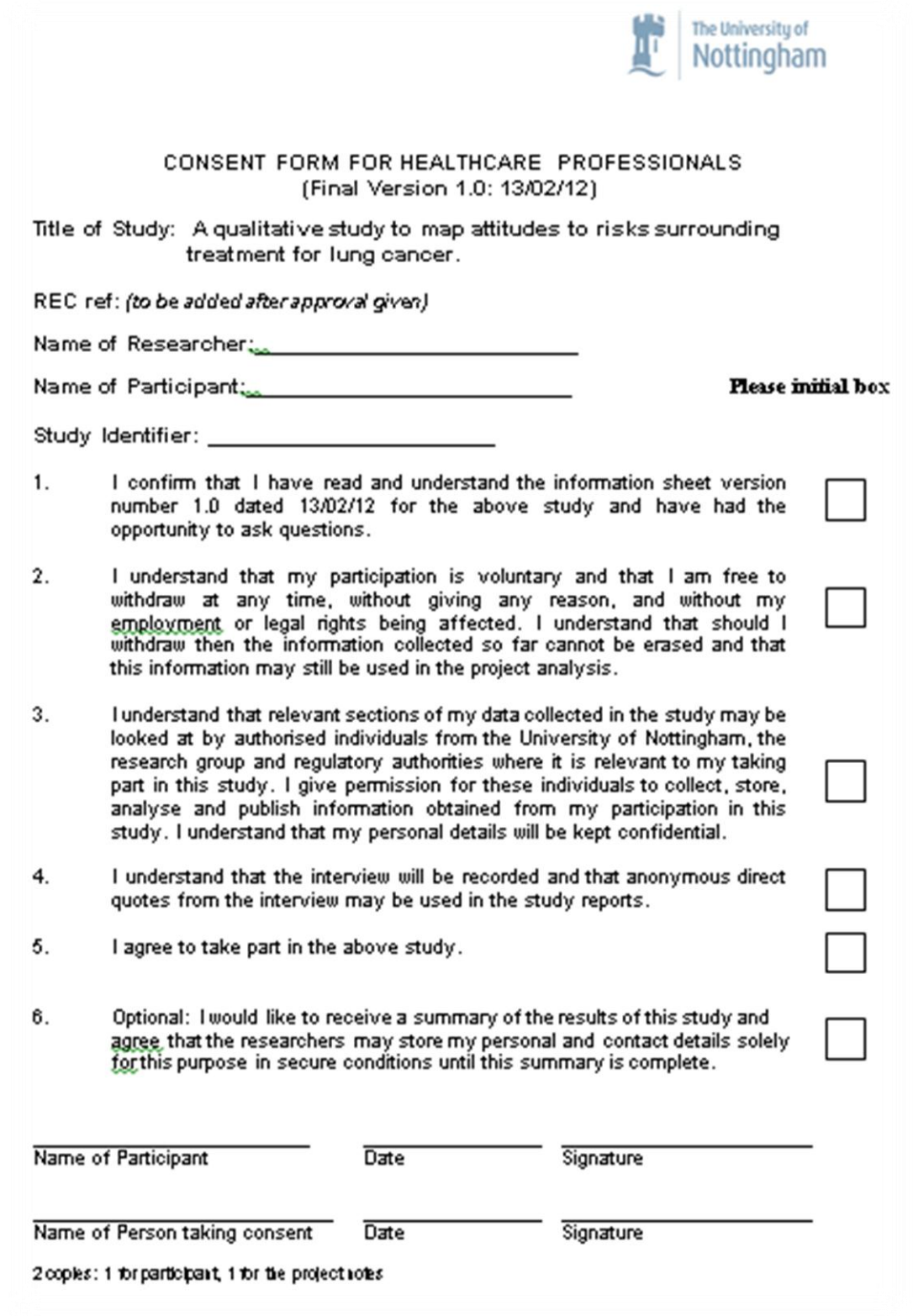
For further information please contact:

Dr Helen Powell
Clinical Sciences Building
Nottingham City Hospital
Hucknall Road
Nottingham
NG5 1PB
0115 8231378
helen.powell@nottingham.ac.uk

The Chief Investigator for this study is:

Professor Richard Hubbard.
Clinical Sciences Building
Nottingham City Hospital
Hucknall Road
Nottingham
NG5 1PB
0115 8231385
richard.hubbard@nottingham.ac.uk

Consent form for healthcare professionals



The University of Nottingham

CONSENT FORM FOR HEALTHCARE PROFESSIONALS
(Final Version 1.0: 13/02/12)

Title of Study: A qualitative study to map attitudes to risks surrounding treatment for lung cancer.

REC ref: *(to be added after approval given)*

Name of Researcher: _____

Name of Participant: _____ **Please initial box**

Study Identifier: _____

1. I confirm that I have read and understand the information sheet version number 1.0 dated 13/02/12 for the above study and have had the opportunity to ask questions. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without my employment or legal rights being affected. I understand that should I withdraw then the information collected so far cannot be erased and that this information may still be used in the project analysis. ☐

3. I understand that relevant sections of my data collected in the study may be looked at by authorised individuals from the University of Nottingham, the research group and regulatory authorities where it is relevant to my taking part in this study. I give permission for these individuals to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential. ☐

4. I understand that the interview will be recorded and that anonymous direct quotes from the interview may be used in the study reports. ☐

5. I agree to take part in the above study. ☐

6. Optional: I would like to receive a summary of the results of this study and agree that the researchers may store my personal and contact details solely for this purpose in secure conditions until this summary is complete. ☐

Name of Participant	Date	Signature
Name of Person taking consent	Date	Signature

2 copies: 1 for participant, 1 for the project notes

Initial interview guide for healthcare professionals

Introduction

- Thank them for coming and taking part.
- Check consent form has been signed. Keep a copy and give participant a copy.
- Statement on confidentiality, right to withdraw consent, recording of the interview.
- Explain the purpose of the study in general:
    - To explore attitudes towards treatment in lung cancer
    - Their opinions on risks of surgery
- Ask if they have any further questions before starting interview
- Interviews will last between 30 and 60 minutes

Background

- Tell me briefly about your job.
    - How long have you held that role?
    - Has it changed over time?
- How do you see your role in relation to lung cancer patients?
- What role do you take in deciding what sort of treatment they should have?

Risks

- What are the risks involved with treatment for lung cancer?
- Do you have a particular figure in your mind of an acceptable mortality:
    - For your patients?
    - If you were a patient?

- Do you feel there is enough evidence available to help you predict risk?

Treatment decisions

- Who do you think should contribute to the decision of which treatment a patient **is offered**?
- In your practice who makes this decision?
    - Do MDT members always agree?

Communication of risks

- Who should tell the patients about the risks involved with treatments?
- Talk me through a typical consultation with a patient with lung cancer regarding treatment.
    - Are there certain things you tell everyone?
    - What risks do you discuss?
    - How do you express the risks?
    - What would you say specifically about cure?
    - Do you talk about treatments which you are not offering them?
- How important do you think it is that the patient fully understands the risks?
    - Do you think this is usually the case?

Conclusion

- Tell the participants that they have reached the end of the interview
- Do they have any questions in return?
- Remind them about confidentiality.
- Thank them for their time.

The University of Nottingham

Our reference:
RGS 12008
Your reference:
12/EM/0123

Research and Graduate Services
University of Nottingham
King's Meadow Campus
Lenton Lane Nottingham
NG7 2NR

0115 8467906
angela.shone@nottingham.ac.uk

Professor Richard Hubbard, Chief Investigator
Professor of Respiratory Epidemiology
Clinical Sciences Building
Nottingham City Hospital
Hucknall Road
Nottingham
NG5 1PB

Nottingham 1 REC
East Midlands REC centre
The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

1st March 2012

Dear Sir or Madam,

Sponsorship Statement

Re: A qualitative study to map attitudes to risks surrounding treatment for lung cancer.

I can confirm that this research proposal has been discussed with the Chief Investigator and agreement to sponsor the research is in place.

An appropriate process of scientific critique has demonstrated that this research proposal is worthwhile and of high scientific quality.*

Any necessary indemnity or insurance arrangements will be in place before this research starts. Arrangements will be in place before the study starts for the research team to access resources and support to deliver the research as proposed.

Arrangements to allocate responsibilities for the management, monitoring and reporting of the research will be in place before the research starts.

The duties of sponsors set out in the NHS Research Governance Framework for Health and Social Care will be undertaken in relation to this research.**

* Not applicable to student research (except doctoral research).
** Not applicable to research outside the scope of the Research Governance Framework.

Yours faithfully

Angela Shone
Research Governance Manager
University of Nottingham

world-changing research
from The University of Nottingham

NRES Committee East Midlands - Nottingham 1
The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

Telephone: 0115 8839436
Facsimile: 0115 8839924

01 June 2012

Helen Powell
Clinical Research Fellow
C100 Clinical Services Building
Nottingham City Hospital
NG5 1PB

Dear Helen

Full title of study:	A qualitative study to map attitudes to risks surrounding treatment for lung cancer.
REC reference number:	12/EM/0123

Thank you for your e-mail of 29th May 2012. I can confirm the REC has received the documents listed below as evidence of compliance with the approval conditions detailed in our letter dated 10 April 2012. Please note these documents are for information only and have not been reviewed by the committee.

Documents received

The documents received were as follows:

Document	*Version*	*Date*
Other: Letter to MDT members	1.1	08 May 2012

You should ensure that the sponsor has a copy of the final documentation for the study. It is the sponsor's responsibility to ensure that the documentation is made available to R&D offices at all participating sites.

12/EM/0123	Please quote this number on all correspondence

Yours sincerely

A.

Miss Andrea Graham
Committee Co-ordinator

E-mail: Andrea.Graham@nottspct.nhs.uk

Nottingham University Hospitals NHS
NHS Trust

research & innovation

Research & Innovation
Nottingham Integrated Clinical Research Centre
C Floor, South Block
Queen's Medical Centre Campus
Derby Road, Nottingham NG7 2UH
Direct Dial: 0115 849 3320
R&I Dept: 0115 970 9049
Tel: 0115 924 9924
Fax: 0115 849 3295
www.nuh.nhs.uk
nuhrise.org

09 JUL 2012

Dr David R Baldwin
Respiratory Medicine Unit
David Evans Research Centre, Nottingham City Hospital
Nottingham City Hospital
Hucknall Road
Nottingham
NG5 1PB

Dear Dr David R Baldwin

Re: 12RM006

CSP# 92660

A qualitative study to map attitudes to risks surrounding treatment for lung cancer.

The R&I Department has considered the following documents:

REC favourable opinion with conditions 18 April 2012
REC favourable opinion conditions met 1 June 2012
Protocol version 1.0 13 February 2012
Schematic diagram for patients version 1.0 13 February 2012
Schematic diagram for healthcare professionals version 1.0 13 February 2012
Letter to MDT members version 1.1 08 May 2012
Letter of invitation to participant: Patient version 1.0 13 February 2012
Participant information sheet: Patients version 1.0 13 February 2012
Participant consent form: Patients version 1.0 13 February 2012
Letter to Lung Cancer Nurse Specialists version 1.0 13 February 2012
Interview guide for Patients version 1.0 13 February 2012
Data collection form for patients' medical notes version 1.0 13 February 2012
Letter to representatives of Professional Bodies version 1.0 13 February 2012
Invitation letter to healthcare professionals version 1.0 13 February 2012

Participant information sheet: Healthcare Professionals version 1.0 13 February 2012
Participant consent form: Healthcare Professionals version 1.0 13 February 2012
Interview guide for Healthcare Professionals version 1.0 13 February 2012

Your study now has R&I approval, on the understanding and provision that you will follow the conditions set out below.

Conditions of Approval

That you:

1. Comply with all relevant laws, regulations and codes of practice applicable to the trial including but not limited to, the UK Clinical Trials Regulations, Medicines for Human Use (Clinical Trial) Regulations 2004, principles of Good Clinical Practice, the World Medical Association Declaration of Helsinki entitled 'Ethical Principles for Medical Research Involving Human Subjects' (1996 version), the Human Rights Act 1998, the Data Protection Act 1998 the Medicines Act 1968, the NHS Research Governance Framework for Health and Social Care (version 2 April 2005). Should any of these be revised and reissued the latest version of the relevant laws and regulations will apply. Copies of the regulations are available from the R&I Office or via the R&I website http://nuhrise.org

2. For NUH sponsored studies ;
   - Accept the responsibilities as outlined in the "Clinical Trial Delegation of Sponsorship responsibilities to Chief Investigator" agreement.
   - Request R&I review of any proposed amendments or changes to study documentation prior to submission to the regulatory authorities by submitting documents to rdappl@nuh.nhs.
     *Submissions may not be made to REC and MHRA (as appropriate) without NUH sponsor authorisation.

3. For studies ***not*** sponsored by NUH request written approval from the R&I department, Ethics Committee and MHRA (as appropriate) for any Protocol Amendments, changes to study documentation or changes to study team.

   For studies adopted onto the NIHR portfolio please submit documents to NUHNT.TRENTCLRN@nhs.net

For all other studies please use rdamend@nuh.nhs.uk

4. Ensure all study personnel, not employed by the Nottingham University Hospitals NHS Trust hold either honorary contracts/letters of access with this Trust, before they have access to any patients or staff, their data, tissue or organs or any NUH facilities.

5. According to R&D SOP 11 - "Adverse Event Monitoring, Recording and Reporting for investigators" report any Serious Adverse Events to the R&D department. For NUH sponsored studies, report any Serious Adverse Events to the R&I department according to R&I SOP 11 - "Adverse Event Monitoring, Recording and Reporting for investigators

6. According to R&D SOP 12 - "Protocol Violations and Serious Breach Reporting" report any Serious Breach of the UK Clinical Trial regulations in connection with the trial or Serious Breach of the protocol, immediately after becoming aware of the breach to R&I.

7. Complete Annual Safety, Progress reports and End of Study reports as required by R&I, Ethics Committee and the MHRA.

8. Notify R&I within 7 calendar days of the first participant recruited onto the study, as well as the detail of the specific recruitment date. Please email the recruitment notification to rdmon@nuh.nhs.uk

9. For GTAC-approved studies, the R&D approval letter should be forwarded to GTAC via the sponsor. GTAC should then issue a site authorisation letter which must be received by each site prior to recruitment commencing. A copy of this letter must be forwarded to R&I.

This approval letter constitutes a favourable Site Specific Assessment (SSA) for this site.

Please note that the R&D department has a database containing study related information, and personal information about individual investigators e.g. name, address, contact details etc. This information will be managed according to the principles established in the Data Protection Act.

Yours sincerely,

Dr Brian Thomson / Dr Maria Koufali
Director of R&D / Deputy Director Research and Innovation

cc Nottingham Research Ethics Committee

REFERENCES

1. Cancer Research UK. CancerStats key facts: Lung cancer and smoking. 2013.

2. Ferlay J, Shin H-R, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International Journal of Cancer. 2010;127(12):2893-917.

3. Doll R, Hill AB. Lung cancer and other causes of death in relation to smoking - A 2nd report on the mortality of British doctors. British Medical Journal. 1956;2(NOV10):1071-81.

4. Centres for Disease Control and prevention. Centers for Disease Control and Prevention. Achievements in public health, United States, 1900-1999: tobacco use. . 1999 Contract No.: 48.

5. Cancer Research UK. Lung Cancer - UK Incidence Statistics. 2011 [updated 3rd October 2013]; Available from: http://info.cancerresearchuk.org/cancerstats/types/lung/incidence/.

6. Hal Weinstein - Vice President and Creative Director Leo Burnett Company. How an agency builds a brand - The Virginia Slims Story. 1970.

7. Graham H. Smoking prevalence among women in the European community 1950-1990. Social Science & Medicine. 1996;43(2):243-54.

8. Peto R, Darby S, Deo H, Silcocks P, Whitley E, Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. BMJ. 2000;321(7257):323-9.

9. Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. The Lancet Oncology. 2013;14(12):1165-74.

10. Haddad R, Massaro D. Idiopathic diffuse interstitial pulmonary fibrosis (fibrosing alveolitis), atypical epithelial proliferation and lung cancer. The American Journal of Medicine. 1968;45(2):211-9.

11. Brenner DR, McLaughlin JR, Hung RJ. Previous lung diseases and lung cancer risk: a systematic review and meta-analysis. PLoS ONE [Electronic Resource]. 2011;6(3):e17479.

12. Nitadori J-i, Inoue M, Iwasaki M, Otani T, Sasazuki S, Nagai K, et al. Association between lung cancer incidence and family history of lung

cancer*: Data from a large-scale population-based cohort study, the jphc study. CHEST Journal. 2006;130(4):968-75.

13. Cassidy A, Myles JP, Duffy SW, Liloglou T, Field JK. Family history and risk of lung cancer: age-at-diagnosis in cases and first-degree relatives. Br J Cancer. 2006;95(9):1288-90.

14. National Institute for Health and Clinical Excellence. The diagnosis and treatment of lung cancer (update) NICE guideline. 2011.

15. Elias AD. Small cell lung cancer : State-of-the-art therapy in 1996. CHEST Journal. 1997;112(4_Supplement):251S-8S.

16. Union for International Cancer Control. TNM Classification of Malignant Tumours, 7th Edition. Leslie H. Sobin, Mary K. Gospodarowicz, Christian Wittekind, editors: Wiley-Blackwell; 2009.

17. Zelen M. Keynote address on biostatistics and data retrival. Cancer Chemotherapy Reports Part 3 Program Information-Supplement. 1973;4(2):31-42.

18. Shepherd FA, Crowley J, Van Houtte P, Postmus PE, Carney D, Chansky K, et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: Proposals Regarding the Clinical Staging of Small Cell Lung Cancer in the Forthcoming (Seventh) Edition of the Tumor, Node, Metastasis Classification for Lung Cancer. Journal of Thoracic Oncology. 2007;2(12):1067-77 10.97/JTO.0b013e31815bdc0d.

19. Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, et al. The IASLC Lung Cancer Staging Project: Proposals for the Revision of the TNM Stage Groupings in the Forthcoming (Seventh) Edition of the TNM Classification of Malignant Tumours. Journal of Thoracic Oncology. 2007;2(8):706-14 10.1097/JTO.0b013e31812f3c1a.

20. Saunders M, Dische S, Barrett A, Harvey A, Griffiths G, Parmar M, et al. Continuous, hyperfractionated, accelerated radiotherapy (CHART) versus conventional radiotherapy in non-small cell lung cancer: mature data from the randomised multicentre trial. Radiotherapy and Oncology. 1999;52(2):137-48.

21. Senan S, Palma DA, Lagerwaard FJ. Stereotactic ablative radiotherapy for stage I NSCLC: Recent advances and controversies. Journal of Thoracic Disease. 2011;3(3):189-96.

22. National Institute for Health and Clinical Excellence. Pemetrexed for the first line treatment of non-small-cell lung cancer 2009.

23. Non-Small Cell Lung Cancer Collaborative G. Chemotherapy and supportive care versus supportive care alone for advanced non-small cell

lung cancer. The Cochrane database of systematic reviews. 2010(5):CD007309-CD.

24. Arriagada R, Bergman B, Dunant A, Le Chevalier T, Pignon J-P, Vansteenkiste J, et al. Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. New England Journal of Medicine. 2004;350(4):351-60.

25. Lim E, Baldwin D, Beckles M, Duffy J, Entwisle J, Faivre-Finn C, et al. Guidelines on the radical management of patients with lung cancer. Thorax. 2010;65:1-27.

26. Non-small Cell Lung Cancer Collaborative Group. Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. BMJ. 1995;311(7010):899-909.

27. National institute for Health and Clinical Excellence. Erlotinib for the treatment of non-small cell lung cancer. 2008.

28. National Institute for Health and Clinical Excellence. Gefitinib for the first-line treatment of locally advanced or metastatic non-small cell lung cancer. 2010.

29. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, Tan EH, Hirsh V, Thongprasert S, et al. Erlotinib in Previously Treated Non–Small-Cell Lung Cancer. New England Journal of Medicine. 2005;353(2):123-32.

30. Mok TS, Wu Y-L, Thongprasert S, Yang C-H, Chu D-T, Saijo N, et al. Gefitinib or Carboplatin–Paclitaxel in Pulmonary Adenocarcinoma. New England Journal of Medicine. 2009;361(10):947-57.

31. Green RA, Humphrey E, Close H, Patno ME. Alkylating agents in bronchogenic carcinoma. The American Journal of Medicine. 1969;46(4):516-25.

32. Lowenbraun S, Bartolucci A, Smalley RV, Lynn M, Krauss S, Durant JR. Superiority of combination chemotherapy over single agent chemotherapy in small cell lung-carcinoma. Cancer. 1979;44(2):406-13.

33. Alberto P, Brunner KW, Martz G, Obrecht JP, Sonntag RW. Treatment of bronchogenic carcinoma with simultaneous or sequential combination chemotherapy, including methotrexate, cyclophosphamide, procarbazine and vincristine. Cancer. 1976;38(6):2208-16.

34. Pujol JL, Carestia, Daures JP. Is there a case for cisplatin in the treatment of small-cell lung cancer? A meta-analysis of randomized trials of a cisplatin-containing regimen versus a regimen without this alkylating agent. Br J Cancer. 2000;83(1):8-15.

35. Sundstrøm S, Bremnes RM, Kaasa S, Aasebø U, Hatlevoll R, Dahle R, et al. Cisplatin and Etoposide Regimen Is Superior to Cyclophosphamide, Epirubicin, and Vincristine Regimen in Small-Cell Lung Cancer: Results From a Randomized Phase III Trial With 5 Years' Follow-Up. Journal of Clinical Oncology. 2002;20(24):4665-72.

36. Roth BJ, Johnson DH, Einhorn LH, Schacter LP, Cherng NC, Cohen HJ, et al. Randomized study of cyclophosphamide, doxorubicin, and vincristine versus etoposide and cisplatin versus alternation of these two regimens in extensive small-cell lung cancer: a phase III trial of the Southeastern Cancer Study Group. Journal of Clinical Oncology. 1992;10(2):282-91.

37. Fukuoka M, Furuse K, Saijo N, Nishiwaki Y, Ikegami H, Tamura T, et al. Randomized Trial of Cyclophosphamide, Doxorubicin, and Vincristine Versus Cisplatin and Etoposide Versus Alternation of These Regimens in Small-Cell Lung Cancer. J Natl Cancer Inst. 1991;83(12):855-61.

38. Takada M, Fukuoka M, Kawahara M, Sugiura T, Yokoyama A, Yokota S, et al. Phase III Study of Concurrent Versus Sequential Thoracic Radiotherapy in Combination With Cisplatin and Etoposide for Limited-Stage Small-Cell Lung Cancer: Results of the Japan Clinical Oncology Group Study 9104. Journal of Clinical Oncology. 2002;20(14):3054-60.

39. Turrisi AT, Kim K, Blum R, Sause WT, Livingston RB, Komaki R, et al. Twice-Daily Compared with Once-Daily Thoracic Radiotherapy in Limited Small-Cell Lung Cancer Treated Concurrently with Cisplatin and Etoposide. New England Journal of Medicine. 1999;340(4):265-71.

40. Pignon J-P, Arriagada R, Ihde DC, Johnson DH, Perry MC, Souhami RL, et al. A Meta-Analysis of Thoracic Radiotherapy for Small-Cell Lung Cancer. New England Journal of Medicine. 1992;327(23):1618-24.

41. Lassen U, Osterlind K, Hansen M, Dombernowsky P, Bergman B, Hansen HH. Long-term survival in small-cell lung cancer: posttreatment characteristics in patients surviving 5 to 18+ years--an analysis of 1,714 consecutive patients. Journal of Clinical Oncology. 1995;13(5):1215-20.

42. Greco FA, Richardson RL, Snell JD, Stroup SL, Oldham RK. Small cell lung cancer - complete remission and improved survival. American Journal of Medicine. 1979;66(4):625-30.

43. Sørensen M, Pijls-Johannesma M, Felip E, Group ObotEGW. Small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2010;21(suppl 5):v120-v5.

44. Simon GR, Turrisi A. Management of Small Cell Lung Cancer*ACCP Evidence-Based Clinical Practice Guidelines (2nd Edition). CHEST Journal. 2007;132(3_suppl):324S-39S.

45. Fried DB, Morris DE, Poole C, Rosenman JG, Halle JS, Detterbeck FC, et al. Systematic Review Evaluating the Timing of Thoracic Radiation Therapy in Combined Modality Therapy for Limited-Stage Small-Cell Lung Cancer. Journal of Clinical Oncology. 2004;22(23):4837-45.

46. Spiro SG, James LE, Rudd RM, Trask CW, Tobias JS, Snee M, et al. Early Compared With Late Radiotherapy in Combined Modality Treatment for Limited Disease Small-Cell Lung Cancer: A London Lung Cancer Group Multicenter Randomized Clinical Trial and Meta-Analysis. Journal of Clinical Oncology. 2006;24(24):3823-30.

47. Slotman B, Faivre-Finn C, Kramer G, Rankin E, Snee M, Hatton M, et al. Prophylactic Cranial Irradiation in Extensive Small-Cell Lung Cancer. New England Journal of Medicine. 2007;357(7):664-72.

48. Cao KJ, Huang FY, Tu MC, Pan GY. Long-term results of prophylactic cranial irradiation for limitedstage small-cell lung cancer in complete remission. Chin Med J. 2005;118(15):1258-62.

49. Temel JS, Greer JA, Muzikansky A, Gallagher ER, Admane S, Jackson VA, et al. Early Palliative Care for Patients with Metastatic Non-Small-Cell Lung Cancer. New England Journal of Medicine. 2010;363(8):733-42.

50. Department of Health. The NHS Cancer plan: A plan for investment, a plan for reform. 2000.

51. Health and Social Care Information Centre. National Lung Cancer Audit Report - Report for the audit period 2011. 2012; Available from: http://www.hscic.gov.uk/lung.

52. Cancer Research UK. Cancer survival statistics for common cancers. 2013 [22nd Nov 2013]; Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/survival/common-cancers/.

53. Coleman MP, Forman D, Bryant H, Butler J, Rachet B, Maringe C, et al. Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data. Lancet. 2011;377(9760):127-38.

54. Janssen-Heijnen MLG, Gatta G, Forman D, Capocaccia R, Coebergh JWW. Variation in survival of patients with lung cancer in Europe, 1985–1989. Eur J Cancer. 1998;34(14):2191-6.

55. Khakwani A, Rich AL, Tata LJ, Powell HA, Stanley RA, Baldwin DR, et al. The pathological confirmation rate of lung cancer in England using the NLCA database. Lung Cancer. 2013;79(2):125-31.

56. NHS Information Centre. National Lung Cancer Audit 2011 - Report for the Audit period 2010. 2011.

57. Damhuis R, Schutte P. Resection rates and postoperative mortality in 7,899 patients with lung cancer. European Respiratory Journal. 1996;9(1):7-10.

58. Fry WA, Phillips JL, Menck HR. Ten-year survey of lung cancer treatment and survival in hospitals in the United States - A National Cancer Data Base report. Cancer. 1999;86(9):1867-76.

59. Gatta G, Sant M (on behalf of the EUROCARE working group). Patterns of treatment of lung cancer patients in Europe (abstract). Annual meeting of the International Association of Cancer Registries. 1993.

60. Rich AL, Tata LJ, Free CM, Stanley RA, Peake MD, Baldwin DR, et al. Inequalities in outcomes for non-small cell lung cancer: the influence of clinical characteristics and features of the local lung cancer service. Thorax. 2011;66(12):1078-84.

61. Brown JS, Eraut D, Trask C, Davison AG. Age and the treatment of lung cancer. Thorax. 1996;51(6):564-8.

62. Levy DT, Chaloupka F, Gitchell J. The Effects of Tobacco Control Policies on Smoking Rates: A Tobacco Control Scorecard. Journal of Public Health Management and Practice. 2004;10(4):338-53.

63. Wilson JMG, Jungner G. Principles and practice of screening for disease. Boletin De La Oficina Sanitaria Panamericana. 1968;65(4):281-&.

64. Chambers SK, Dunn J, Occhipinti S, Hughes S, Baade P, Sinclair S, et al. A systematic review of the impact of stigma and nihilism on lung cancer outcomes. Bmc Cancer. 2012;12.

65. Iyen-Omofoman B, Hubbard R, Smith C, Sparks E, Bradley E, Bourke A, et al. The distribution of lung cancer across sectors of society in the United Kingdom: A study using national primary care data. BMC Public Health. 2011;11(1):857.

66. Gatsonis CA, Natl Lung Screening Trial Res T. The National Lung Screening Trial: Overview and Study Design. Radiology. 2011;258(1):243-53.

67. The National Lung Screening Trial Research Team. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. New England Journal of Medicine. 2011;365(5):395-409.

68. UKLS. UK Lung cancer Screening Trial (UKLS). 2012 [cited 2012 14/6/12]; Available from: http://public.ukcrn.org.uk/search/StudyDetail.aspx?StudyID=10297.

69. van Iersel CA, de Koning HJ, Draisma G, Mali WPTM, Scholten ET, Nackaerts K, et al. Risk-based selection from the general population in a screening trial: Selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON). International Journal of Cancer. 2007;120(4):868-74.

70. Department of Health NCAT, and Cancer Research UK,. National Awareness & Early Diagnosis Initiative. 2012 [14/6/12]; Available from: http://info.cancerresearchuk.org/spotcancerearly/naedi/AboutNAEDI/.

71. National Cancer Action Team. Lung cancer awareness campaign evaluation - Impact on patients attending GP appointments. February 2012.

72. Walters S, Maringe C, Coleman MP, Peake MD, Butler J, Young N, et al. Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK: a population-based study, 2004-2007. Thorax. 2013;68(6):551-64.

73. Hamilton W, Peters TJ, Round A, Sharp D. What are the clinical features of lung cancer before the diagnosis is made? A population based case-control study. Thorax. 2005;60(12):1059-65.

74. Iyen Omofoman B. Lung cancer in United Kingdom general practice and the possibility of developing an early-warning score: University of Nottingham; 2012.

75. The European Initiative for Quality Management in Lung Cancer Care. ERS taskforce report. European Respiratory Journal. 2013;In press.

76. Riaz SP, Linklater KM, Page R, Peake MD, Møller H, Lüchtenborg M. Recent trends in resection rates among non-small cell lung cancer patients in England. Thorax. 2012.

77. Khakwani A, Rich AL, Powell HA, Tata LJ, Stanley RA, Baldwin DR, et al. Lung cancer survival in England: trends in non-small-cell lung cancer survival over the duration of the National Lung Cancer Audit. British Journal of Cancer. 2013;109(8):2058-65.

78. National Cancer Action Team. National Cancer Peer Review Programme Report 2010/2011: An overview of the findings from the 2010/2011 National Cancer Peer Review of Cancer Services in England. 2011.

79. Aveling E-L, Martin G, Jiménez García S, Martin L, Herbert G, Armstrong N, et al. Reciprocal peer review for quality improvement: an ethnographic

case study of the Improving Lung Cancer Outcomes Project. Bmj Quality & Safety. 2012;21(12):1034-41.

80. The Health Foundation. Improving Lung Cancer Outcomes Project (ILCOP) - Final Report for award holders. 2013.

81. Rich A. Validation of the National Lung Cancer Audit database and analysis of the information it contains: University of Nottingham; 2011.

82. Rich AL, Tata LJ, Stanley RA, Free CM, Peake MD, Baldwin DR, et al. Lung cancer in England: Information from the National Lung Cancer Audit (LUCADA). Lung cancer (Amsterdam, Netherlands). 2011;72(1):16-22.

83. Rich AL, Tata LJ, Free CM, Stanley RA, Peake MD, Baldwin DR, et al. How do patient and hospital features influence outcomes in small-cell lung cancer in England? British Journal of Cancer. 2011;105(6):746-52.

84. Cegidim Strategic Data (CSD). THIN data from EPIC: A guide for researchers. 2009.

85. Townsend P, Phillimore P, Beattie A. Health and Deprivation: Inequality and the North. London: Croom Helm; 1988.

86. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. American Journal of Clinical Oncology. 1982;5(6):649-56.

87. Berrino F, Verdecchia A, Capocaccia R, Hakulinen T, Esteve J. Survival of cancer patients in Europe: The EUROCARE Study. In: IARC Scientific Publications, editor.1995. p. xiii+463p.

88. National Cancer Institute. Surveillance Epidemiology and End Results. 2013; Available from: http://seer.cancer.gov/.

89. The Health and Social Care Information Centre. HESonline. 2005-2012 [17/11/2012].

90. Charlson ME, Pompei P, Ales KL, Mackenzie CR. A new method of classifying prognostic co-morbidity in longitudinal-studies - development and validation. Journal of Chronic Diseases. 1987;40(5):373-83.

91. Audit Commission. Improving data quality in the NHS Annual report on the PbR assurance programme. Audit Commission Publishing Team, 2010.

92. American Cancer Society Report. 2010 Cancer facts and figures. Atlanta, GA: 2010.

93. Thun MJ, Henley SJ, Burns D, Jemal A, Shanks TG, Calle EE. Lung cancer death rates in lifelong nonsmokers. J Natl Cancer Inst. 2006;98(10):691-9.

94. US Department of Health and Human Services. The health consequences of smoking: a report of the surgeon general. Atlanta,GA: National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2004.

95. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ. 2004;328(7455):1519.

96. Huxley RR, Woodward M. Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. Lancet. 2011;378(9799):1297-305.

97. Zang EA, Wynder EL. Differences in Lung Cancer Risk Between Men and Women: Examination of the Evidence. J Natl Cancer Inst. 1996;88(3-4):183-92.

98. Risch HA, Howe GR, Jain M, Burch JD, Holowaty EJ, Miller AB. Are female smokers at higher risk for lung-cancer than male smokers - A case control analysis by histologic type. American Journal of Epidemiology. 1993;138(5):281-93.

99. Freedman ND, Leitzmann MF, Hollenbeck AR, Schatzkin A, Abnet CC. Cigarette smoking and subsequent risk of lung cancer in men and women: analysis of a prospective cohort study. The Lancet Oncology. 2008;9(7):649-56.

100. Bain C, Feskanich D, Speizer FE, Thun M, Hertzmark E, Rosner BA, et al. Lung Cancer Rates in Men and Women With Comparable Histories of Smoking. J Natl Cancer Inst. 2004;96(11):826-34.

101. Brownson RC, Chang JC, Davis JR. Gender and histologic type variations in smoking-related risk of lung cancer. Epidemiology. 1992;3(1):61-4.

102. Kreuzer M, Boffetta P, Whitley E, Ahrens W, Gaborieau V, Heinrich J, et al. Gender differences in lung cancer risk by smoking: a multicentre case-control study in Germany and Italy. Br J Cancer. 1999;82(1):227-33.

103. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung-volumes and forced ventilatory flows. Report working party "Standardization of lung-function tests". European Community for Steel and Coal and the European Respiratory Society. European Respiratory Journal. 1993;6:5-40.

104. McDuffie HH, Klaassen DJ, Dosman JA. Female-male differences in patients with primary lung-cancer. Cancer. 1987;59(10):1825-30.

105. Osann KE, Antonculver H, Kurosaki T, Taylor T. Sex-differences in lung-cancer risk associated with cigarette-smoking. International Journal of Cancer. 1993;54(1):44-8.

106. Harris RE, Zang EA, Anderson JI, Wynder EL. Race and sex-differences in lung-cancer risk associated with cigarette-smoking. International Journal of Epidemiology. 1993;22(4):592-9.

107. Morabia A, Wynder EL. Cigarette smoking and lung cancer cell types. Cancer. 1991;68(9):2074-8.

108. Engeland A. Trends in the incidence of smoking-associated cancers in Norway, 1954–93. International Journal of Cancer. 1996;68(1):39-46.

109. Prescott E, Osler M, Hein HO, Borch-Johnsen K, Lange P, Schnohr P, et al. Gender and smoking-related risk of lung cancer. The Copenhagen Center for Prospective Population Studies. Epidemiology. 1998;9(1):79-83.

110. Tulinius H, Sigfusson N, Sigvaldason H, Bjarnadottir K, Tryggvadottir L. Risk factors for malignant diseases: A cohort study on a population of 22,946 Icelanders. Cancer Epidemiology Biomarkers & Prevention. 1997;6(11):863-73.

111. Henschke CI, Sone S, Hanaoka T, Markowitz S, Miller A, Klingler K, et al. Women's susceptibility to tobacco carcinogens and survival after diagnosis of lung cancer. JAMA - Journal of the American Medical Association. 2006;296(2):180-4.

112. Frueh M, Cerny D, Ess S, Cerny T. Lung cancer in women in St. Gallen, Eastern Switzerland: An analysis of sex-associated differences in smoking habits, disease presentation and survival. Lung Cancer. 2009;64:S32-S.

113. Ryu J-S, Jeon S-H, Kim J-S, Lee JH, Kim SH, Hong JT, et al. Gender Differences in Susceptibility to Smoking among Patients with Lung Cancer. Korean Journal of Internal Medicine. 2011;26(4):427-31.

114. Huxley R, Jamrozik K, Lam T, Barzi F, Ansary-Moghaddam A, Jiang C, et al. Impact of Smoking and Smoking Cessation on Lung Cancer Mortality in the Asia-Pacific Region. American Journal of Epidemiology. 2007;165(11):1280-6.

115. Office for National Statistics. Smoking and drinking among adults, 2009 - A report on the 2009 General Lifestyle Survey. 2009.

116. International Agency for Research on Cancer. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Tobacco Smoke and Involuntary Smoking. 2004.

117. Wells AJ, English PB, Posner SF, Wagenknecht LE, Perez-Stable EJ. Misclassification rates for current smokers misclassified as nonsmokers. American Journal of Public Health. 1998;88(10):1503-9.

118. Green J, Cairns BJ, Casabonne D, Wright FL, Reeves G, Beral V. Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk. The Lancet Oncology. 2011;12(8):785-94.

119. Gunnell D, Okasha M, Smith GD, Oliver SE, Sandhu J, Holly JMP. Height, leg length, and cancer risk: A systematic review. Epidemiologic Reviews. 2001;23(2):313-42.

120. Lee J, Kolonel LN. Body height and lung-cancer risk. Lancet. 1983;1(8329):877-.

121. Ragnarsdottir M, Kristinsdottir EK. Breathing movements and breathing patterns among healthy men and women 20-69 years of age - Reference values. Respiration. 2006;73(1):48-54.

122. Muhammad-Kah R, Liang Q, Frost-Pineda K, Mendes PE, Roethig HJ, Sarkar M. Factors affecting exposure to nicotine and carbon monoxide in adult cigarette smokers. Regulatory Toxicology and Pharmacology. 2011;61(1):129-36.

123. Eissenberg T, Adams C, Riggins EC, 3rd, Likness M. Smokers' sex and the effects of tobacco cigarettes: subject-rated and physiological measures. Nicotine & Tobacco Research. 1999;1(4):317-24.

124. Nakamura H, Ando K, Shinmyo T, Morita K, Mochizuki A, Kurimoto N, et al. Female Gender Is an Independent Prognostic Factor in Non-small-cell Lung Cancer: A Meta-analysis. Annals of Thoracic and Cardiovascular Surgery. 2011;17(5):469-80.

125. Steliga MA, Dresler CM. Smoking cessation: crucial to target women as well as men. Lancet. 2011;378(9799):1278-9.

126. Powell HA, Iyen-Omofoman B, Hubbard RB, Baldwin DR, Tata LJ. The Association Between Smoking Quantity and Lung Cancer in Men and Women. CHEST Journal. 2013;143(1):123-9.

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

128. Mathers CD, Loncar D. Projections of Global Mortality and Burden of Disease from 2002 to 2030. PLoS Med. 2006;3(11):e442.

129. Kishi K, Gurney JW, Schroeder DR, Scanlon PD, Swensen SJ, Jett JR. The correlation of emphysema or airway obstruction with the risk of lung cancer: a matched case-controlled study. European Respiratory Journal. 2002;19(6):1093-8.

130. Schwartz AG, Cote ML, Wenzlaff AS, Van Dyke A, Chen W, Ruckdeschel JC, et al. Chronic obstructive lung diseases and risk of non-small cell lung cancer in women. Journal of Thoracic Oncology: Official Publication of the International Association for the Study of Lung Cancer. 2009;4(3):291-9.

131. Adcock IM, Caramori G, Barnes PJ. Chronic obstructive pulmonary disease and lung cancer: new molecular insights. Respiration. 2011;81(4):265-84.

132. Santillan AA, Camargo CA, Colditz GA. A meta-analysis of asthma and risk of lung cancer (United States). Cancer Causes & Control. 2003;14(4):327-34.

133. Littman AJ, Thornquist MD, White E, Jackson LA, Goodman GE, Vaughan TL. Prior lung disease and risk of lung cancer in a large prospective study. Cancer Causes & Control. 2004;15(8):819-27.

134. Wang X-R, Yu ITS, Chiu YL, Qiu H, Fu Z, Goggins W, et al. Previous pulmonary disease and family cancer history increase the risk of lung cancer among Hong Kong women. Cancer Causes & Control. 2009;20(5):757-63.

135. Khurana V, Bejjanki HR, Caldito G, Owens MW. Statins Reduce the Risk of Lung Cancer in Humans*. Chest. 2007;131(5):1282-8.

136. National Clinical Guideline Centre. Chronic obstructive pulmonary disease: management of chronic obstructive pulmonary disease in adults in primary and secondary care. 2010.

137. Manser R, Wright G, Hart D, Byrnes G, Campbell D, Wainer Z, et al. Surgery for local and locally advanced non-small cell lung cancer. Cochrane Database of Systematic Reviews 2010 2011;Art. No.: CD004699(Issue 10).

138. Smith CJP, Gribbin J, Challen KB, Hubbard RB. The impact of the 2004 NICE guideline and 2003 General Medical Services contract on COPD in primary care in the UK. QJM. 2008;101(2):145-53.

139. Szatkowski L, Lewis S, McNeill A, Huang Y, Coleman T. Can data from primary care medical records be used to monitor national smoking prevalence? Journal of Epidemiology and Community Health. 2011.

140. Ekberg-Aronsson M, Nilsson PM, Nilsson J-Å, Pehrsson K, Löfdahl C-G. Socio-economic status and lung cancer risk including histologic subtyping—A longitudinal study. Lung Cancer. 2006;51(1):21-9.

141. Mao Y, Hu J, Ugnat A-M, Semenciw R, Fincham S, Group atCCRER. Socioeconomic status and lung cancer risk in Canada. International Journal of Epidemiology. 2001;30(4):809-17.

142. Turner MC, Chen Y, Krewski D, Calle EE, Thun MJ. Chronic obstructive pulmonary disease is associated with lung cancer mortality in a prospective study of never smokers. American Journal of Respiratory & Critical Care Medicine. 2007;176(3):285-90.

143. Leung CC, Lam TH, Yew WW, Law WS, Tam CM, Chang KC, et al. Obstructive lung disease does not increase lung cancer mortality among female never-smokers in Hong Kong. The International Journal of Tuberculosis and Lung Disease. 2012;16(4):546-52.

144. Brenner AV, Wang ZY, Kleinerman RA, Wang LD, Zhang SZ, Metayer C, et al. Previous pulmonary diseases and risk of lung cancer in Gansu Province, China. International Journal of Epidemiology. 2001;30(1):118-24.

145. Purdue MP, Gold L, Järvholm B, Alavanja MCR, Ward MH, Vermeulen R. Impaired lung function and lung cancer incidence in a cohort of Swedish construction workers. Thorax. 2007;62(1):51-6.

146. Schabath MB, Delclos GL, Martynowicz MM, Greisinger AJ, Lu C, Wu XF, et al. Opposing effects of emphysema, hay fever, and select genetic variants on lung cancer risk. American Journal of Epidemiology. 2005;161(5):412-22.

147. Tockman MS, Anthonisen NR, Wright EC, Donithan MG. Airways obstruction and the risk for lung cancer. Annals of Internal Medicine. 1987;106(4):512-8.

148. Wu-Williams A, Dai X, Blot W, Xu Z, Sun X, Xiow H, et al. Lung cancer among women in north-east China. British Journal of Cancer. 1990;62:982-7.

149. Islam SS, Schottenfeld D. Declining FEV1 and chronic productive cough in cigarette smokers - A 25-year prospective-study of lung cancer incidence in Tecumseh, Michigan. Cancer Epidemiology Biomarkers & Prevention. 1994;3(4):289-98.

150. Wu AH, Fontham ETH, Reynolds P, Greenberg RS, Buffler P, Liff J, et al. Previous Lung Disease and Risk of Lung Cancer among Lifetime Nonsmoking Women in the United States. American Journal of Epidemiology. 1995;141(11):1023-32.

151. Brownson RC, Alavanja MC. Previous lung disease and lung cancer risk among women (United States). Cancer Causes & Control. 2000;11(9):853-8.

152. Mannino DM, Aguayo SM, Petty TL, Redd SC. Low lung function and incident lung cancer in the United States - Data from the First National Health and Nutrition Examination Survey follow-up. Archives of Internal Medicine. 2003;163(12):1475-80.

153. Wasswa-Kintu S, Gan WQ, Man SFP, Pare PD, Sin DD. Relationship between reduced forced expiratory volume in one second and the risk of lung cancer: a systematic review and meta-analysis. Thorax. 2005;60(7):570-5.

154. de Torres JP, Bastarrika G, Wisnivesky JP, Alcaide AB, Campo A, Seijo LM, et al. Assessing the relationship between lung cancer risk and emphysema detected on low-dose CT of the chest. Chest. 2007;132(6):1932-8.

155. Wilson DO, Weissfeld JL, Balkan A, Schragin JG, Fuhrman CR, Fisher SN, et al. Association of radiographic emphysema and airflow obstruction with lung cancer. American Journal of Respiratory & Critical Care Medicine. 2008;178(7):738-44.

156. Yang PMDP, Sun ZMD, Krowka MJMD, Aubry M-CMD, Bamlet WRMS, Wampfler JABS, et al. Alpha1-Antitrypsin Deficiency Carriers, Tobacco Smoke, Chronic Obstructive Pulmonary Disease, and Lung Cancer Risk. Archives of Internal Medicine. 2008;168(10):1097-103.

157. Kiri VA, Soriano J, Visick G, Fabbri L. Recent trends in lung cancer and its association with COPD: an analysis using the UK GP Research Database. Primary Care Respiratory Journal. 2010;19(1):57-61.

158. Maldonado F, Bartholmai BJ, Swensen SJ, Midthun DE, Decker PA, Jett JR. Are Airflow Obstruction and Radiographic Evidence of Emphysema Risk Factors for Lung Cancer? :A Nested Case-Control Study Using Quantitative Emphysema Analysis. Chest. 2010;138:1295-302. Epub March 26, 2010.

159. Gonzalez-Perez A, Fernandez-Vidaurre C, Rueda A, Rivero E, Garcia Rodriguez LA. Cancer incidence in a general population of asthma patients. Pharmacoepidemiology & Drug Safety. 2006;15(2):131-8.

160. Houghton AM. Mechanistic links between COPD and lung cancer. Nat Rev Cancer. 2013;13(4):233-45.

161. Powell HA, Iyen-Omofoman B, Baldwin DR, Hubbard RB, Tata LJ. Chronic Obstructive Pulmonary Disease and Risk of Lung Cancer: The Importance

of Smoking and Timing of Diagnosis. Journal of Thoracic Oncology. 2013;8(1):6-11 0.1097/JTO.0b013e318274a7dc.

162. Molins L, Agusti À. Chronic Obstructive Pulmonary Disease and Risk of Lung Cancer: The Importance of Smoking and Timing of Diagnosis. Journal of Thoracic Oncology. 2013;8(4):e34 10.1097/JTO.0b013e318286c1c1.

163. Chansky K, Sculier J-P, Crowley JJ, Giroux D, Van Meerbeeck J, Goldstraw P, et al. The International Association for the Study of Lung Cancer Staging Project: Prognostic Factors and Pathologic TNM Stage in Surgically Managed Non-small Cell Lung Cancer. Journal of Thoracic Oncology. 2009;4(7):792-801 10.1097/JTO.0b013e3181a7716e.

164. Roth K, Nilsen TIL, Hatlen E, Sørensen KS, Hole T, Haaverstad R. Predictors of long time survival after lung cancer surgery: A retrospective cohort study. BMC Pulmonary Medicine. 2008;8(22).

165. van der Pijl LLR, Birim O, van Gameren M, Kappetein AP, Maat APWM, Steyerberg EW, et al. Validation of a prognostic model to predict survival after non-small-cell lung cancer surgery. European Journal of Cardio-Thoracic Surgery. 2010;38(5):615-9.

166. Birim O, Kappetein AP, Waleboer M, Puvimanasinghe JPA, Eijkemans MJC, Steyerberg EW, et al. Long-term survival after non-small cell lung cancer surgery: Development and validation of a prognostic model with a preoperative and postoperative mode. J Thorac Cardiovasc Surg. 2006;132(3):491-8.

167. al-Kattan K, Sepsas E, Townsend ER, Fountain SW. Factors affecting long term survival following resection for lung cancer. Thorax. 1996;51(12):1266-9.

168. Powell HA, Tata LJ, Baldwin DR, Stanley RA, Hubbard RB. P13 Identifying patients who had sugical resection for non-small cell lung cancer using large datasets. Thorax. 2013;68(Suppl 3):A80.

169. Riaz SP, Lüchtenborg M, Jack RH, Coupland VH, Linklater KM, Peake MD, et al. Variation in surgical resection for lung cancer in relation to survival: Population-based study in England 2004–2006. Eur J Cancer. 2011(0).

170. Falcoz PE, Conti M, Brouchet L, Chocron S, Puyraveau M, Mercier M, et al. The Thoracic Surgery Scoring System (Thoracoscore): Risk model for in-hospital death in 15,183 patients requiring thoracic surgery. J Thorac Cardiovasc Surg. 2007;133(2):325-32.

171. Robinson LA, Ruckdeschel JC, Wagner H, Stevens CW. Treatment of Non-small Cell Lung Cancer-Stage IIIA*. Chest. 2007;132(3 suppl):243S-65S.

172. Saklad M. Grading of patients for surgical procedures. Anesthesiology. 1941;2:281-4.

173. Prause G, Ratzenhofer-Comenda B, Pierer G, Smolle-Ju¨ttner F, Glanzer H, Smolle J. Can ASA grade or Goldman's cardiac risk index predict peri-operative mortality? A study of 16 227 patients. Anaesthesia. 1997;52(3):203-6.

174. Daabiss M. American Society of Anaesthesiologists physical status classification. Indian journal of anaesthesia. 2011;55(2):111-5.

175. Cuvillon P, Nouvellon E, Marret E, Albaladejo P, Fortier LP, Fabbro-Perray P, et al. American Society of Anesthesiologists' Physical Status system: a multicentre Francophone study to analyse reasons for classification disagreement. European Journal of Anaesthesiology. 2011;28(10):742-7.

176. Goldman L, Caldera DL, Nussbaum SR, Southwick FS, Krogstad D, Murray B, et al. Multifactorial index of cardiac risk in non-cardiac surgical procedures. New England Journal of Medicine. 1977;297(16):845-50.

177. Copeland GP, Jones D, Walters M. POSSUM: A scoring system for surgical audit. British Journal of Surgery. 1991;78(3):355-60.

178. Brunelli A, Fianchini A, Gesuita R, Carle F. POSSUM scoring system as an instrument of audit in lung resection surgery. Annals of Thoracic Surgery. 1999;67(2):329-31.

179. Haga Y, Ikei S, Ogawa M. Estimation of physiologic ability and surgical stress (E-PASS) as a new prediction scoring system for postoperative morbidity and mortality following elective gastrointestinal surgery. Surgery Today. 1999;29(3):219-25.

180. Yamashita S, Haga Y, Nemoto E, Nagai S, Ohta M. E-PASS (The Estimation of Physiologic Ability and Surgical Stress) scoring system helps the prediction of postoperative morbidity and mortality in thoracic surgery. European Surgical Research. 2004;36(4):249-55.

181. Berrisford R, Brunelli A, Rocco G, Treasure T, Utley M. The European Thoracic Surgery Database project: modelling the risk of in-hospital death following lung resection. European Journal of Cardio-Thoracic Surgery. 2005;28(2):306-11.

182. Brunelli A, Varela G, Van Schil P, Salati M, Novoa N, Hendriks JM, et al. Multicentric analysis of performance after major lung resections by using the European Society Objective Score (ESOS). European Journal of Cardio-Thoracic Surgery. 2008;33(2):284-8.

183. Chamogeorgakis TP, Connery CP, Bhora F, Nabong A, Toumpoulis IK. Thoracoscore predicts midterm mortality in patients undergoing thoracic

surgery. Journal of Thoracic & Cardiovascular Surgery. 2007;134(4):883-7.

184. Bradley A, Marshall A, Abdelaziz M, Hussain K, Agostini P, Bishay E, et al. THORACOSCORE fails to predict complications following elective lung resection. European Respiratory Journal. 2012.

185. Bernard A, Rivera C, Pages PB, Falcoz PE, Vicaut E, Dahan M. Risk model of in-hospital mortality after pulmonary resection for cancer: A national database of the French Society of Thoracic and Cardiovascular Surgery (Epithor). J Thorac Cardiovasc Surg. 2011;141(2):449-58.

186. Stolz A, Pafko P, Harustiak T, Smejkal M, Simonek J, Schutzner J, et al. Risk factor analysis for early mortality and morbidity following pneumonectomy for non-small cell lung cancer. Bratislavske Lekarske Listy. 2011;112(4):165-9.

187. Damhuis R, Coonar A, Plaisier P, Dankers M, Bekkers J, Linklater K, et al. A case-mix model for monitoring of postoperative mortality after surgery for lung cancer. Lung Cancer. 2006;51(1):123-9.

188. Epstein SK, Faling LJ, Daly BD, Celli BR. Predicting complications after pulmonary resection. Preoperative exercise testing vs a multifactorial cardiopulmonary risk index. Chest. 1993;104(3):694-700.

189. Brunelli A, Rocco G, Varela G. Predictive ability of preoperative indices for major pulmonary surgery. Thoracic Surgery Clinics. 2007;17(3):329-36.

190. Bernard A, Ferrand L, Hagry O, Benoit L, Cheynel N, Favre JP. Identification of prognostic factors determining risk groups for lung resection. Annals of Thoracic Surgery. 2000;70(4):1161-7.

191. Brunelli A, Sabbatini A, Xiume F, Borri A, Salati M, Marasco RD, et al. Inability to perform maximal stair climbing test before lung resection: a propensity score analysis on early outcome. European Journal of Cardio-Thoracic Surgery. 2005;27(3):367-72.

192. Brunelli A, Al Refai M, Salati M, Sabbatini A, Morgan-Hughes NJ, Rocco G. Carbon monoxide lung diffusion capacity improves risk stratification in patients without airflow limitation: evidence for systematic measurement before lung resection. European Journal of Cardio-Thoracic Surgery. 2006;29(4):567-70.

193. Amar D, Munoz D, Shi W, Zhang H, Thaler HT. A clinical prediction rule for pulmonary complications after thoracic surgery for primary lung cancer. Anesthesia & Analgesia. 2010;110(5):1343-8.

194. Balduyck B, Hendriks J, Lauwers P, Van Schil P. Quality of life evolution after lung cancer surgery: A prospective study in 100 patients. Lung Cancer. 2007;56(3):423-31.

195. Goldstraw P, Ball D, Jett JR, Le Chevalier T, Lim E, Nicholson AG, et al. Non-small-cell lung cancer. The Lancet. 2011;378(9804):1727-40.

196. Vallières E, Peters S, Van Houtte P, Dalal P, Lim E. Therapeutic advances in non-small cell lung cancer. Thorax. 2011.

197. Damhuis RAM, Wijnhoven BPL, Plaisier PW, Kirkels WJ, Kranse R, van Lanschot JJ. Comparison of 30-day, 90-day and in-hospital postoperative mortality for eight different cancer types. British Journal of Surgery. 2012;99(8):1149-54.

198. Bryant AS, Rudemiller K, Cerfolio RJ. The 30- Versus 90-Day Operative Mortality After Pulmonary Resection. The Annals of Thoracic Surgery. 2010;89(6):1717-23.

199. Schmidt-Hansen M, Page R, Hasler E. The Effect of Preoperative Smoking Cessation or Preoperative Pulmonary Rehabilitation on Outcomes After Lung Cancer Surgery: A Systematic Review. Clinical Lung Cancer. 2013;14(2):96-102.

200. Powell HA, Tata LJ, Baldwin DR, Stanley RA, Khakwani A, Hubbard RB. Early mortality after surgical resection for lung cancer: an analysis of the English National Lung cancer audit. Thorax. 2013;68(9):826-34.

201. Waller DA. Whose risk is it anyway? Thorax. 2013;68(9):799-800.

202. McCracken JD, Janaki LM, Crowley JJ, Taylor SA, Giri PG, Weiss GB, et al. Concurrent chemotherapy/radiotherapy for limited small-cell lung carcinoma: a Southwest Oncology Group Study. Journal of Clinical Oncology. 1990;8(5):892-8.

203. Warde P, Payne D. Does thoracic irradiation improve survival and local control in limited-stage small-cell carcinoma of the lung? A meta-analysis. Journal of Clinical Oncology. 1992;10(6):890-5.

204. Powell HA, Tata LJ, Baldwin DR, Stanley RA, Hubbard RB. P15 Identifying patients who receive chemotherapy for small-cell lung cancer using large datasets. Thorax. 2013;68(Suppl 3):A81.

205. Singh S, Parulekar W, Murray N, Feld R, Evans WK, Tu D, et al. Influence of sex on toxicity and treatment outcome in small-cell lung cancer. Journal of Clinical Oncology. 2005;23(4):850-6.

206. Burgers JA, Arance A, Ashcroft L, Hodgetts J, Lomax L, Thatcher N. Identical chemotherapy schedules given on and off trial protocol in small cell lung cancer response and survival results. British Journal of Cancer. 2002;87(5):562-6.

207. Elliss-Brookes L, McPhail S, Ives A, Greenslade M, Shelton J, Hiom S, et al. Routes to diagnosis for cancer - determining the patient journey using multiple routine data sets. Br J Cancer. 2012;107(8):1220-6.

208. McPhail S, Elliss-Brookes L, Shelton J, Ives A, Greenslade M, Vernon S, et al. Emergency presentation of cancer and short-term mortality. Br J Cancer. 2013.

209. Fountain SW, Armstrong P, Congleton J, Jagoe T, McAuley DF, MacMahon J, et al. Guidelines on the selection of patients with lung cancer for surgery. Thorax. 2001;56(2):89-108.

210. Cykert S. Risk acceptance and risk aversion: patients' perspectives on lung surgery. Thoracic Surgery Clinics. 2004;14(3):287-93.

211. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. Bmc Medical Research Methodology. 2013;13.

212. National Cancer Intelligence Network (hosted by Public Health England). SACT Systemic Anti-cancer Therapy - Chemotherapy dataset. 2011; Available from: http://www.chemodataset.nhs.uk/home.

213. Shaw AT, Kim D-W, Nakagawa K, Seto T, Crinó L, Ahn M-J, et al. Crizotinib versus Chemotherapy in Advanced ALK-Positive Lung Cancer. New England Journal of Medicine. 2013;368(25):2385-94.