

TITLE

A population-based study to quantify the association between patient's age, deprivation and specific comorbid conditions, and the receipt of major surgery in non-small cell lung cancer patients in England

RUNNING TITLE

Characteristics of patients undergoing surgery for lung cancer

AUTHORS

Aurélien Belot¹, Helen Fowler¹, Edmund Njeru Njagi¹, Miguel-Angel Luque-Fernandez¹, Camille Maringe¹, Winnie Magadi¹, Aimilia Exarchakou¹, Manuela Quaresma¹, Adrian Turculet¹, Michael D Peake^{2,3,4}, Neal Navani⁵, Bernard Rachet¹

Authors' Affiliation

¹ Cancer Survival Group, Department of Non-Communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

² Department of Respiratory Medicine, University Hospitals of Leicester

³ National Cancer Registration and Analysis Service, Public Health England

⁴ Centre for Cancer Outcomes, University College London Hospitals

⁵ UCL Respiratory, University College London, 5 University Street, London WC1E 6JF and Department of Thoracic Medicine, University College London Hospital, 250 Euston Road, London NW1 2PG

*Corresponding author:

Aurélien Belot

Cancer Survival Group

Department of Non-Communicable Disease Epidemiology

London School of Hygiene & Tropical Medicine

Keppel Street

London WC1E 7HT

Email: aurelien.belot@lshtm.ac.uk

Article category: Original Research article

Total number of pages (including title page, references, figure legends and tables): 25;

Tables: 3; Figures: 3; Supplementary Data: 1 appendix; References: 40

ABSTRACT

Introduction

We investigated socioeconomic disparities and the role of the main prognostic factors in receiving major surgical treatment in lung cancer patients in England.

Methods

Our study comprised of 31351 patients diagnosed with Non-Small Cell Lung Cancer in England in 2012. Data from the national population-based cancer registry were linked to Hospital Episode Statistics and National Lung Cancer Audit data to obtain information on stage, performance status and comorbidities, and to identify patients receiving major surgical treatment. To describe the association between prognostic factors and surgery, we performed two different analyses: one using multivariable logistic regression, and one estimating cause-specific hazards for death and surgery. In both analyses, we used multiple imputation to deal with missing data.

Results

We showed strong evidence that the comorbidities “congestive heart failure”, “cerebrovascular disease” and “chronic obstructive pulmonary disease” reduced the receipt of surgery in early-stage patients. We also observed gender differences and substantial age differences in the receipt of surgery. Despite accounting for sex, age at diagnosis, comorbidities, stage at diagnosis, performance status and indication of having had a PET-CT scan, the socioeconomic differences persisted in both analyses: more deprived people had lower odds and lower rates of receiving surgery in early-stage lung cancer.

Discussion

Comorbidities play an important role in whether patients undergo surgery, but do not completely explain the socioeconomic difference observed in early stage patients. Future work investigating access to and distance from specialist hospitals, as well as patient perceptions and patient choice in receiving surgery, could help disentangle these persistent socioeconomic inequalities.

Keywords

Lung cancer; surgical treatment; population-based data; socioeconomic inequalities; comorbidities;

Abbreviations used

COPD: Chronic Obstructive Pulmonary Disease

HES: Hospital Episode Statistics

HR: Hazard Ratio

MI: Multiple Imputation

NLCA: National Lung Cancer Audit

NOS: Not Otherwise Specified

NSCLC: Non-Small Cell Lung Cancer

OPCS: Office of Population Censuses and Surveys, Classification of Interventions and Procedures

OR: Odds Ratio

PET-CT: Positron Emission Tomography - Computed Tomography

PS: performance status

SCLC: Small Cell Lung Cancer

SMC-FCS: Substantive-Model Compatible Fully Conditional Specification

Funding

This research was supported by Cancer Research UK grant number C7923/A18525. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of Cancer Research UK.

Authors' contributions

AB and BR developed the concept and design of the study. AB, HF, MALF, CM, WM and AT were involved in the data preparation and the data linkage. AB carried out the data analysis and wrote the manuscript. All authors interpreted the data, drafted and critically revised the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Ethics approval

We obtained the ethical and statutory approvals required for this research (PIAG 1-05(c)/2007; ECC 1-05(a)/2010); ethical approval updated 6 April 2017 (REC 13/LO/0610). We attest that we have obtained appropriate permissions and paid any required fees for use of copyright protected materials.

Acknowledgment

The authors thank three anonymous reviewers for their comments, which helped to improve the manuscript. The authors also thank Francisco Javier Rubio from the Cancer Survival Group of the London School of Hygiene & Tropical Medicine for interesting discussion on the topic.

Short summary ‘box’

- What is the key question?

What is the association between specific comorbid conditions and the receipt of major surgical treatment in Non-Small Cell Lung Cancer (NSCLC) patients, and does it explain socio-economic inequalities in the receipt of treatment?

- What is the bottom line?

We showed strong evidence that (i) the comorbidities “congestive heart failure”, “cerebrovascular disease” and “chronic obstructive pulmonary disease” reduced the rate of surgery in early stage NSCLC patients, and (ii), after adjusting for covariates such as age, stage and performance status, socio-economic inequalities in receiving surgical treatment remain.

- Why read on?

Using population-based data, we quantified the association between specific comorbid conditions and the receipt of surgery, and showed that socio-economic inequalities in receiving major surgical treatment persist, even after accounting for the main prognostic factors (stage, performance status, specific comorbid conditions).

INTRODUCTION

Lung cancer is the leading cause of cancer death in the United-Kingdom for both men and women: in 2014, more than 19000 deaths were observed in men and over 16000 in women, corresponding to age-standardised mortality rates of 89.69 and 60.52 per 100 000 person-years, respectively.¹ Despite encouraging decreasing age-standardised incidence in lung cancer,¹ it remains a major public health issue given its poor prognosis; for patients diagnosed in England between 2010 and 2014, the one-year age-standardised net survival probability was estimated at 34% in men and 40% in women, those figures being 11% and 16% at 5 years.²

Different studies have reported socio-economic inequalities, with a consistently lower survival for the more deprived patients.³⁻⁵ Disentangling the reasons behind these socioeconomic inequalities remains a topic of active research. One aspect to consider is any variation in patient characteristics between socio-economic groups. For example, studies that have examined the prevalence of comorbidity according to socio-economic group found that while comorbidity, or indeed multi-morbidity, was most prevalent in the elderly, it was also found in younger patients of lower socio-economic status.⁶⁻⁷ Moreover, some authors have suggested that variations in surgical resection rates and more generally in access to treatment could partly explain those inequalities.⁸⁻¹² Understanding which factors are associated with the probability to receive major surgical treatment could help further improve the surgical resection rate, and thus lung cancer survival.⁸ In addition to the central role of tumour stage at diagnosis, the role of comorbidities is often studied using the summarised Charlson score.⁹⁻¹¹ Thus, the question regarding the association between specific individual comorbid conditions and the probability to receive major surgery remains unresolved. We hypothesize that patients with comorbidities have lower probability of receiving surgery than patients without comorbidities. Conversely, we hypothesize that patients diagnosed with an early-stage tumour and whose health status is

good should be very likely to receive a surgical treatment with curative intent. Moreover, controlling for these specific comorbidities will allow assessment of whether the association between socioeconomic deprivation and treatment remains.

We aim here to quantify the association between patient and tumour factors and the probability to receive major surgical treatment for patients diagnosed with Non-Small Cell Lung Cancer (NSCLC) in England in 2012, with a particular emphasis on the roles played by specific comorbid conditions and deprivation.

METHODS

Data

Data were obtained from the National Cancer Registry at the Office for National Statistics¹³ for all adults (aged 15-99) diagnosed with a first primary malignant NSCLC (all ICD-O-3 morphology codes except those between 8041 and 8045, behaviour /3) in 2012 in England. Patients diagnosed with a lung cancer for which the morphology was malignant neoplasm Not Otherwise Specified (NOS) or Epithelial neoplasms NOS, were also included. All patients have been followed-up until the 31 December 2015 by the National Health Service Central Register, to update each patient's vital status (alive, emigrated, dead or not traced). Data were further linked to the National Lung Cancer Audit (NLCA) data,¹⁴ using successively different combinations of patient identifier variables to ensure successful matching of same patients in the two datasets.¹⁵ The NLCA data captured information on several tumour and patient-related factors, such as (i) the tumour's T, N, M (coding the size, the node and the metastasis components of stage, respectively), (ii) the patient's performance status (PS), (iii) whether a PET-CT scan was performed (yes/no) and (iv) the surgical procedure undertaken with the

intention of curing the patient.¹⁴ PET-CT scan is recommended prior to surgery as it provides more accurate staging information, in particular detecting metastatic disease unidentified by other imaging investigations. We applied a previously developed algorithm to derive for each tumour a summarised 4-category stage at diagnosis variable using information from the NLCA in the first instance, and from the national cancer registry data if no information were present in the NLCA data.¹⁶

For each NSCLC patient, we used the inpatient, outpatient, and Accident & Emergency Hospital Episode Statistics (HES) datasets to derive prevalence indicators of the 17 comorbid conditions of the Charlson Comorbidity Index, in addition to obesity.¹⁷ We focused on comorbidities that would influence fitness to undergo major surgery, namely myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease and chronic obstructive pulmonary disease (COPD).¹⁸ To retrieve information on comorbidity, we used a 6-year window up to 6 months before diagnosis, checking the diagnostic fields of the HES data to identify whether any of the aforementioned comorbidities had been recorded during this period.¹⁷

The primary source of information for the major surgical treatment was the NLCA data, supplemented with information from the HES data where the former was unavailable. In both datasets, information on surgery (with the corresponding date) was recorded using the Office of Population Censuses and Surveys (OPCS) Classification of Interventions and Procedures (fourth version, 'OPCS-4').¹⁹ We pre-set a time window of 30 days before diagnosis and 180 days after diagnosis to identify the first major surgical treatment received for each patient. The categorisation of surgery as 'major' was made according to the OPCS-4 codes recorded for the procedure, and was based on definitions of "surgery with curative intent" devised from extensive input from clinicians and oncologists and endorsed by the Site-Specific Clinical Reference Groups of the National Cancer Intelligence Network (Appendix 1).²⁰ Where patients

had surgery before the recorded cancer diagnosis date, we recoded their time from diagnosis to surgery as 1 day for the time-to-event analysis (see Analysis section).

We measured deprivation using the Income Domain from the 2010 England Indices of Multiple Deprivation (IMD) defined at the Lower Super Output Area level (average population of 1,500).²¹ The Income Domain score is a combination of five indicators, and it measures the proportion of the population in an area experiencing deprivation related to low income. It ranges from 0 to 77%,²¹ (i.e. the higher the more deprived, see figure S2 in the appendix) and is named “deprivation score” hereafter. More details on the different deprivation measures and the way they are calculated could be found here (<https://www.gov.uk/government/statistics/english-indices-of-deprivation-2010>).

Analysis

The studied prognosis factors were sex, age at diagnosis, deprivation score, PS, stage at diagnosis, a binary variable indicating if a PET/PET-CT scan was performed, and five binary variables indicating the presence of specific comorbid conditions.

We performed two types of analysis to describe the receipt of surgery while controlling for the prognostic factors. Firstly, we used a logistic regression model with major surgery (yes/no) as the outcome (*Analysis 1*). However, some patients may have died before having had a chance to receive surgery. Secondly, we modelled the two cause-specific hazards (i.e. one for surgery and one for death) using the semi-parametric Cox model (*Analysis 2*). We analysed the first event with the corresponding time-to-event, thus accounting for the competing risk of death:²² patients who died without receiving surgery were censored for the surgery-specific rate. Conversely, patients who received surgery were censored at their time of surgery when

estimating the death-specific rate. As the time window for the surgery went up to a maximum of 180 days, patients alive without surgery after 180 days were censored at that time.

In each analysis, the predictors were defined as follows: for the continuous variables age at diagnosis and deprivation score, we assumed non-linear associations using flexible functional forms (quadratic splines with one inflection point around the median, i.e. at 70 years for the age at diagnosis, and at 0.1 for the deprivation score); from the original PS¹⁴ coded according to the World Health Organisation classification (indicating a person's status relating to activity/disability), we re-coded this variable in 2 groups: good (0 or 1) vs. poor (2, 3 and 4); the 4-category variable for stage at diagnosis was dichotomised in 2 groups: early stage (stage 1 and 2) vs. late stage (3 and 4). It would be expected that patients with good PS and early tumour stage would be offered surgery as the first-choice treatment. Finally, we assumed an interaction between all the variables and the binary indicator defining early vs. late stage.

To deal with the issue of missing data, we performed multiple imputation (MI) using the "Substantive-Model Compatible Fully Conditional Specification" (SMC-FCS) method,²³ under the Missing At Random assumption²⁴ (meaning that the probability of missing data in the particular covariate is dependent on the observed data, i.e. the other covariates and the outcome, but not additionally on the unobserved value of the particular covariate). The variables PET-CT scan, stage and PS had missing information, up to 30% of missing values for the PET-CT scan variable. These 3 variables being binary, we used a logistic regression model for the imputation including as variables all the prognostic factors defining the linear predictor (see above) in addition to the outcome (i.e. the binary surgery variable for *Analysis 1* or the time and the type of event for *Analysis 2*). We generated 30 imputed datasets.²⁵

To assess the strength of the association between the outcome and the deprivation score, we performed a Likelihood Ratio Test (LRT) on each imputed dataset by comparing a model that included the deprivation score (in its non-linear form and with the interaction) with one that did not. Each of the 30 LRT statistics were compared to a chi square distribution with 6 degrees of freedom. We also used a joint Wald test from the estimated parameters obtained after MI to assess the strength of the association between the outcome (either binary or time-to-event) and the deprivation score.

We repeated the steps detailed above on the complete case data (using only the data from patients having no missing information in any of the variables), which is usually called a “sensitivity analysis”; it allows to assess how variables the results are to the missing data mechanism assumed, and thus how much caution may be needed in interpreting them.²⁶

RESULTS

Study population

The study population were 31351 patients diagnosed with NSCLC in England in 2012 (11030 adenocarcinomas -35%-, 3442 large cells -11%-, 7058 squamous cells -23%- and 9821 other histological subtypes -31%-). The percentage of successful linkage between the cancer registry data (the core data) and the NLCA data was 83%, and it was higher than 99% for linkage with the HES data. From these patients, 4850 received surgery (Table 1). The mean age was 72.8 years overall, the group who received surgery were younger on average (68.4 compared to 73.6 years). Sixty eight percent of NSCLC were diagnosed with an advanced stage, 9% stage being missing. Information of PET-CT scan and performance status was also missing for 30% and 24%

of patients, respectively. COPD was the most prevalent co-morbidity (22%) with much lower prevalence observed for all other comorbidities (between 5.2% and 6.4%).

Table 1: Characteristics of patients diagnosed with Non-Small Cell Lung Cancer in England in 2012

N	No Surgery N=26501	Surgery N=4850	p ^a	Total N=31351
Age at diagnosis (mean (sd))	73.62 (10.89)	68.42 (9.81)	<0.001	72.81 (10.90)
Deprivation score (mean (sd))	0.17 (0.11)	0.16 (0.12)	0.016	0.17 (0.12)
Sex (%)				
Men	14729 (55.6)	2463 (50.8)	<0.001	17192 (54.8)
Women	11772 (44.4)	2387 (49.2)		14159 (45.2)
PET-CT scan (%)				
No	12334 (46.5)	391 (8.1)	<0.001	12725 (40.6)
Yes	5428 (20.5)	3801 (78.4)		9229 (29.4)
Missing	8739 (33.0)	658 (13.6)		9397 (30.0)
Stage at diagnosis (%)				
Stage I-II (Early stage)	3339 (12.6)	3823 (78.8)	<0.001	7162 (22.8)
Stage III-IV (Advanced stage)	20488 (77.3)	878 (18.1)		21366 (68.2)
Missing	2674 (10.1)	149 (3.1)		2823 (9.0)
PS (%)				
Good (0 or 1)	9332 (35.2)	3728 (76.9)	<0.001	13060 (41.7)
Poor (2, 3 or 4)	10309 (38.9)	340 (7.0)		10649 (34.0)
Missing	6860 (25.9)	782 (16.1)		7642 (24.4)
Myocardial Infarction (%)				
No	25027 (94.4)	4633 (95.5)	0.002	29660 (94.6)
Yes	1474 (5.6)	217 (4.5)		1691 (5.4)
Congestive Heart Failure (%)				
No	24955 (94.2)	4753 (98.0)	<0.001	29708 (94.8)
Yes	1546 (5.8)	97 (2.0)		1643 (5.2)
Peripheral Vascular Disease (%)				
No	24771 (93.5)	4574 (94.3)	0.031	29345 (93.6)
Yes	1730 (6.5)	276 (5.7)		2006 (6.4)
Cerebrovascular disease (%)				
No	24726 (93.3)	4704 (97.0)	<0.001	29430 (93.9)
Yes	1775 (6.7)	146 (3.0)		1921 (6.1)
COPD (%)				
No	20671 (78.0)	3949 (81.4)	<0.001	24620 (78.5)
Yes	5830 (22.0)	901 (18.6)		6731 (21.5)

^a: p-value of the test comparing the groups with and without surgery (t-test for continuous covariables and Chi² test for categorical covariables)

PS: Performance Score; COPD: Chronic Obstructive Pulmonary Disease

Results from Analysis 1: the multivariable logistic regression

The strongest predictor of the odds of receiving surgery was stage: the adjusted odds of receiving surgery for advanced stage patients was 0.01 times that of early stage patients (OR: 0.01, with 95% confidence interval [0.007;0.014]) (Table 2). This was expected because surgery is rarely offered to advanced stage patients. However, it indicates evidence of the quality and the validity of the dataset. Interactions between binary stage (early/advanced) and all other factors included in the logistic model enabled the parameter estimates for early and advanced stage to be interpreted separately, and we focus on interpreting the results among early stage patients unless otherwise stated. A poor performance status (2, 3 or 4) dramatically reduced the odds of receiving surgery (OR: 0.15 [0.13;0.17]). A recorded diagnosis of congestive heart failure, cerebrovascular disease or COPD was strongly associated with reduced odds of receiving surgery (ORs 0.55 [0.40;0.74], 0.58 [0.45;0.76] and 0.50 [0.43;0.57], respectively). By contrast, having had a PET-CT scan was associated with an increased odds of receiving surgery (3.65 [3.08;4.33]). Women had higher odds of having surgery compared to men (OR 1.22 [1.08;1.38]).

The probability to receive surgery for both early and advanced stage patients without any comorbidities, with good PS and who had a PET-CT scan performed is provided for men with low deprivation (score of 0.1) by age at diagnosis (Figure 1a). We observed a clear association between a patients' age and the probability of receiving surgery with a dramatic decrease after 75 years: the probability to receive surgery for a patient aged 50 years was 87.9% [85.4;90.0] compared to 65.2% [61.9;68.3] at 80 years (Figure 1a).

The probability to receive surgery for both early and advanced stage patients aged 70 years at diagnosis by deprivation score is provided for men (Figure 1b).The joint Wald test showed evidence of differences on the probability to receive surgery by deprivation (p -value<0.01 of

the Chi² test comparing models with and without deprivation score). This was also confirmed by the 30 LRT, all showing strong association between surgery and deprivation score. Overall, we observed decreasing probability of surgery from least to most deprived patients in early-stage patients (figure 1b): the probability to receive surgery was 86.0% [83.1;88.4] for patients living in areas with 3% of people with low income (i.e. deprivation score of 0.03, less deprived) compared to 78.6% [75.7;81.3] for patients living in areas with 40% of people with low income (i.e. a deprivation score of 0.4, more deprived). Such difference was not observed in advanced-stage patients for whom the probability to receive surgery was estimated around 16% for all deprivation scores, Figure 1b.

The multivariable regression model does not include interactions between sex and age at diagnosis nor between sex and deprivation score. This leads to patterns of predicted probabilities of surgery for women similar to men, while at a different absolute level, by age and deprivation; we reported the predicted probability to receive surgery for both early and advanced stage patients without any comorbidities, with good PS and who had a PET-CT scan performed for women with a deprivation score of 0.1 according to age at diagnosis (Figure S1a, Appendix 1), and for women of 70 years old at diagnosis according to deprivation score (Figure S1b, Appendix 1).

Comparing the results to those obtained with the complete case analysis, we observed a similar overall pattern of results, with slight differences in some point estimates (Table 2). These differences were expected as when we explored the missing data mechanism, we found that the probability of missing data for PS and for PET-CT scan was strongly associated with the outcome variable after controlling for other covariables. Such an observation of the missing data mechanism implied that the analysis based on complete cases would be expected to be biased.²⁴

Table 2: Adjusted^a odds ratios of receiving surgery with corresponding 95% Confidence Intervals (CI) within each stage-specific stratum^b, and estimated i) after multiple imputation and ii) only using complete cases.

	Multiple Imputation			Complete Cases		
	OR	95% CI		OR	95% CI	
In stratum with Early stage						
Sex (ref: Males)	1.22	1.08	1.38	1.18	1.03	1.35
PET-CT scan (ref: No PET-CT Scan)	3.65	3.08	4.33	4.2	3.47	5.08
PS (Ref PS=good)	0.15	0.13	0.17	0.16	0.13	0.18
Myocardial Infarction (ref: No)	0.94	0.73	1.21	0.91	0.68	1.21
Congestive Heart Failure (ref: No)	0.55	0.4	0.74	0.63	0.45	0.9
Peripheral Vascular Disease (ref: No)	0.91	0.73	1.14	0.9	0.7	1.15
Cerebrovascular disease (ref: No)	0.58	0.45	0.76	0.61	0.46	0.82
COPD (ref: No)	0.5	0.43	0.57	0.49	0.42	0.57
In stratum with Advanced stage						
Sex (ref: Males)	1.26	1.09	1.46	1.23	1.05	1.45
PET-CT scan (ref: No PET-CT Scan)	14.53	11.53	18.3	16.64	13.21	20.97
PS (Ref PS=good)	0.2	0.15	0.26	0.21	0.16	0.27
Myocardial Infarction (ref: No)	1.18	0.81	1.74	1.07	0.73	1.58
Congestive Heart Failure (ref: No)	0.55	0.31	0.99	0.47	0.26	0.86
Peripheral Vascular Disease (ref: No)	1.04	0.73	1.48	0.92	0.65	1.32
Cerebrovascular disease (ref: No)	0.65	0.39	1.08	0.82	0.49	1.37
COPD (ref: No)	1.11	0.9	1.36	1.09	0.89	1.34

^a: the odds ratios are adjusted on the covariables reported in the table and also on age at diagnosis and deprivation score, those variables being modelled using a non-linear functional form (see methods section), and both with a main effect in addition to an interaction with the indicator variable “advanced stage”.

^b: the adjusted OR of receiving surgery for advanced stage patients compared to early stage patients was 0.01, 95% CI [0.007;0.014] after multiple imputation (OR=0.01 [0.008;0.016] for the complete cases).

PS: Performance Score; COPD: Chronic Obstructive Pulmonary Disease

Results of Analysis 2: competing risks analysis

The competing risk analysis corroborated the results obtained with the multivariable logistic regression, with a low Hazard Ratio (HR) of receiving surgery for all advanced stage patients and for patients with poor performance status in both early and advanced stage of diagnosis (Table 3). The estimated HRs associated with comorbidities also showed strong evidence of a lower rate of receiving surgery for patients with congestive heart failure, cerebrovascular disease and COPD in early-stage patients. The mortality rate was clearly associated with stage: advanced stage patients had a mortality rate 3.34 [2.60;4.27] times higher than patients with an early stage diagnosis. A poor performance status increased the mortality rate in both early and advanced stage patients. None of the comorbid conditions showed any strong association with mortality.

The rate of receiving surgery decreased dramatically for older patients (Figure 2a): early-stage patient aged 80 years at diagnosis had around half the rate of receiving surgery of early stage patients aged 70 years (HR 0.57 [0.54;0.60]). Furthermore, we found strong evidence of an association between the rate of receiving surgery and deprivation score from both the joint Wald test and the 30 LR Tests. The rate of receiving surgery for an early stage deprived patient (deprivation score of 0.4) was 12% lower than that of a less deprived patient (deprivation score of 0.1, HR 0.88 [0.80;0.97]) (Figure 2b). We observed a clear association with age for both stage groups (Figure 3a), but no association with deprivation score (Figure 3b) on the hazard of death without surgery. From the complete case analysis, we reported similar patterns with slight differences in some point estimates (Table 3).

Table 3: Adjusted^a hazard ratios of receiving surgery and of death with corresponding 95% Confidence Intervals (CI) within each stage-specific stratum^b, and estimated i) after multiple imputation and ii) only using complete cases.

	Multiple Imputation						Complete Cases					
	Surgery			Death			Surgery			Death		
	HR	95% CI		HR	95% CI		HR	95% CI		HR	95% CI	
In stratum with Early stage												
Sex (ref: Males)	1.14	1.07	1.22	0.79	0.69	0.91	1.13	1.05	1.21	0.79	0.66	0.93
PET-CT scan (ref: No PET-CT Scan)	2.03	1.78	2.31	0.32	0.27	0.4	2.27	1.98	2.61	0.31	0.25	0.37
PS (Ref PS=good)	0.25	0.22	0.29	2.36	1.93	2.89	0.27	0.24	0.31	2.45	1.98	3.03
Myocardial Infarction (ref: No)	0.96	0.82	1.13	0.96	0.76	1.23	0.93	0.78	1.1	0.82	0.6	1.14
Congestive Heart Failure (ref: No)	0.61	0.49	0.77	1.14	0.92	1.42	0.66	0.51	0.85	1.06	0.79	1.4
Peripheral Vascular Disease (ref: No)	0.94	0.81	1.08	1.05	0.84	1.32	0.94	0.81	1.09	1.04	0.78	1.38
Cerebrovascular disease (ref: No)	0.73	0.61	0.88	1.01	0.81	1.27	0.76	0.63	0.93	1.05	0.79	1.39
COPD (ref: No)	0.71	0.65	0.77	1.06	0.92	1.24	0.71	0.65	0.78	1.1	0.91	1.32
In stratum with Advanced stage												
Sex (ref: Males)	1.22	1.07	1.39	0.88	0.85	0.91	1.21	1.04	1.4	0.88	0.84	0.92
PET-CT scan (ref: No PET-CT Scan)	12.02	9.58	15.09	0.33	0.31	0.36	13.49	10.48	17.37	0.33	0.3	0.35
PS (Ref PS=good)	0.23	0.17	0.3	2.7	2.58	2.83	0.25	0.18	0.33	2.77	2.63	2.91
Myocardial Infarction (ref: No)	1.12	0.78	1.6	1.05	0.97	1.13	1.03	0.69	1.53	1.04	0.93	1.15
Congestive Heart Failure (ref: No)	0.59	0.34	1.03	0.99	0.92	1.07	0.49	0.25	0.97	1	0.9	1.11
Peripheral Vascular Disease (ref: No)	1.03	0.74	1.44	1.07	1	1.15	0.93	0.64	1.37	1.09	1	1.2
Cerebrovascular disease (ref: No)	0.68	0.42	1.11	0.96	0.9	1.03	0.84	0.5	1.4	0.96	0.87	1.05
COPD (ref: No)	1.1	0.91	1.33	0.94	0.9	0.99	1.08	0.88	1.33	0.95	0.9	1.01

^a: the rate ratios are adjusted on the covariables reported in the table and also on age at diagnosis and deprivation score, those variables being modelled using a non-linear functional form (see methods section), and both with a main effect in addition to an interaction with the indicator variable “advanced stage”.

^b: the adjusted HR of receiving surgery for advanced stage patients compared to early stage patients was 0.02, 95% CI [0.01;0.03] after multiple imputation (HR=0.02 [0.01;0.03] for the complete cases), and adjusted HR of death for advanced stage patients compared to early stage patients was 3.34 [2.60;4.27] after multiple imputation (HR=3.32 [2.51;4.38] for the complete cases).

PS: Performance Score; COPD: Chronic Obstructive Pulmonary Disease

DISCUSSION

This study found strong evidence of an inverse association between the receipt of surgery and the comorbidities congestive heart failure, cerebrovascular disease and COPD in patients with early stage. We also observed an important fall in the receipt of surgery with increasing age at diagnosis. Despite adjusting for these prognostic factors, the receipt of surgery still also depended upon the deprivation score: more deprived patients have a lower probability of receiving surgical treatment compared to less deprived patients. Stage at diagnosis and the patient's performance status were confirmed as the strongest predictors for receiving major surgery while there was a clear association between having had a PET-CT scan and the receipt of major surgery.

One major strength of the study is the use of population-based cancer registry data linked to NLCA data and HES data, which provided an enriched source of information describing the probability to receive surgery among NSCLC patients diagnosed in England in 2012. We focused on NSCLC patients as surgery is the recommended treatment in contrast to small-cell lung carcinomas where the primary treatment is usually chemotherapy. We benefitted from algorithms that made the best use of our data,^{16 17} and enabled information on stage at diagnosis and specific type of comorbidities to be included as predictors for the probability to receive surgery. We assumed that the measurement of detailed information on important prognostic factors such as comorbidities, stage, PET-CT scan and PS were either properly recorded (i.e. no measurement error nor misclassification) or properly retrieved using the developed algorithms. We analysed an area-based measure for deprivation;²¹ it has the advantage of encapsulating both the individual SES and the greater environment where the patient lives.²⁷

The study dealt with the problem of missing data using a MI technique which properly accounts for the presence of non-linear functional forms for the continuous variable, as well as interactions between variables.²³ This is another strength of this work as we found other studies that either discarded missing data^{28 29} or treated them as another category,^{10 11 30} although neither of these approaches are recommended. As we were expecting a strong association between stage at diagnosis and the outcome, we assumed an interaction between stage and all the other prognostic factors, which was particularly insightful for interpreting results among early-stage lung cancer. We analysed continuous variables (age and deprivation score) in their existing form, allowing the depiction of their continuous gradual association with the receipt of surgery through flexible functional forms (such as splines, as used here). Under the assumption of the Missing At Random missingness mechanism, results from the multiple imputation are unbiased.²⁴

In the multivariable logistic regression, the outcome was receiving surgery yes/no. The group of patients who did not receive surgery within the first 6 months following diagnosis are a mixture of (i) patients who were alive and not surgically treated and (ii) patients who died during the first 6 months without receiving surgery. This last group being not “at risk” of receiving surgery complicates the interpretation.³¹ The competing risk analysis allowed us to distinguish the two different event types (surgery or death), whichever came first.²² In both analyses, we used a model parametrisation that allows for interpretation of the parameter estimates for the early and the advanced stage groups separately, due to the interaction between early/advanced stage and all the other variables. The results from both analyses led to the same findings, which were consistent with results previously published. Firstly, the stage at diagnosis was the strongest predictor of receiving surgery.¹⁰ The patient’s PS and the indication of a PET-CT scan performed³² were also strongly associated with surgery. The persisting differences

observed between the less and more deprived patients in their receipt of major surgery was also found in some other studies,^{8 11 33-36} with the exception of Rich *et al.*³⁰ However, in this latter study, the authors dealt with the missing data problem differently from us, as they created an additional category “missing”, which may lead to biased estimation.^{37 38} We also observed an important decrease in the probability to receive surgery in older patients,^{39 40} even if their benefit in receiving surgical treatment was equivalent to that of their younger counterparts.⁴¹ Indeed, as stressed by Chambers *et al.*⁴¹ “several prospective and large population studies have shown unanimously, that patients >70 years of age respond as well as younger patients in all outcome measures pertaining to morbidity, mortality and QoL [*Quality of Life*] postoperatively, and should receive aggressive surgical management if considered fit for surgery, in accordance with the British Thoracic Society guidelines.”

We showed strong evidence of lower odds of having major surgery in patients with Congestive Heart Failure, Cerebrovascular disease and COPD in patients with early stage, consistently with a population-based study conducted in the United States,²⁹ while another study conducted in the Netherlands did not find an association between comorbidity conditions and treatment received for NSCLC patients.⁴⁰ However, a limitation of the study is related to the information used for comorbidity, which is derived from the HES datasets.¹⁷ Thus we do not know the comorbidity status for patients who never attended secondary care, and assume they do not have any comorbidities. This could also be the case in an instance where comorbid patients attend hospital, but their comorbidities are erroneously not disclosed in hospital records. So, our measure of comorbidity-specific conditions could underestimate the level of comorbidity at a population level. However, we focused on the more severe comorbidities that are likely to limit fitness for surgery and thus, we would expect the majority of patients with these more serious ‘grades’ of illnesses to have been admitted to hospital at some stage. Having had a PET-

CT scan could be considered as a possible mediator (thus being on the causal pathway) between deprivation and receipt of surgery and between comorbidities and receipt of surgery. We therefore performed a supplementary sensitivity analysis (using a competing-risks method of analysis and focusing on the receipt of surgery, see method) without adjusting on the PET-CT scan variable to evaluate the consistency of the results. The adjusted hazard ratios of receiving surgery were quite similar to those presented in Table 3 and in Figure 2, and the main conclusions remained unchanged (data not shown). Those additional results suggest that having had a PET-CT scan performed played a minor role on the socioeconomic differences to receive surgery.

It appears that further investigation is warranted to understand why the more deprived patients have a lower probability of receiving surgery compared with the less deprived patients, and also why there is an extremely large gradient in receiving surgery according to the age at diagnosis. An approach would be to describe the patient choice, as this is always a major factor in whether treatment, especially surgical treatment, is carried out. It may well be that older and more deprived patients are less likely to agree to surgical treatment. Perceptions of treatments, wishes, levels of engagement with healthcare and understanding of risks of benefits of surgery - all of which affect patient choice - may vary by socio-economic level. Travel issues (e.g. costs and time to travel) may well also come into play in these groups of patients. A study conducted in Australia showed that the longer the distance to a specialist hospital the lower the chance of receiving surgery.⁴² A comparable investigation would be to assess urban-rural difference in access to treatment and receipt of surgery, as reported in the US and in Ireland,^{28 43} as well as whether a patient was first seen at a thoracic surgical centre.³⁰ Crawford *et al.* showed distance from surgical centre and socio-economic deprivation were additive in their association with lower resection rates in a study from Yorkshire in the UK.³⁵ Therefore, an interesting step

forward would be to describe patient choice according to the main patient characteristics (including age and deprivation level). This could be informative in identifying key learning points, such as highlighting the need to focus on either improving access to tertiary services, providing clear information to help inform patient choice and evaluating financial barriers to undergoing surgery (e.g. transport costs/leave from work).

Figure legends

Figure 1: Predicted probability of receiving major surgery with 95% confidence interval for early stage (solid line) and advanced stage (dashed line) in male NSCLC patients without any comorbidities, with good performance status, having had a PET-CT scan: (a) according to age at diagnosis with a deprivation score of 0.1 or (b) according to deprivation score with an age at diagnosis of 70 years. Results are based on the multivariable logistic regression. The ranges of age at diagnosis and deprivation score were limited to the 2.5th and 97.5th percentile of the observed distribution.

Figure 2: Adjusted Hazard ratio of receiving major surgery with 95% confidence interval for early stage (solid line) and advanced stage (dashed line) in male NSCLC patients without any comorbidities, with good performance status, having had a PET-CT scan: (a) according to age at diagnosis with a deprivation score of 0.1 or (b) according to deprivation score with an age at diagnosis of 70 years. Results are based on the competing risks analysis using semi-parametric Cox models. The ranges of age at diagnosis and deprivation score were limited to the 2.5th and 97.5th percentile of the observed distribution.

Figure 3: Adjusted Hazard ratio of death with 95% confidence interval for early stage (solid line) and advanced stage (dashed line) in male NSCLC patients without any comorbidities, with good performance status, having had a PET-CT scan: (a) according to age at diagnosis with a deprivation score of 0.1 or (b) according to deprivation score with an age at diagnosis of 70 years. Results are based on the competing risks analysis using semi-parametric Cox models. The ranges of age at diagnosis and deprivation score were limited to the 2.5th and 97.5th percentile of the observed distribution.

References

1. Smittenaar CR, Petersen KA, Stewart K, et al. Cancer incidence and mortality projections in the UK until 2035. *Br J Cancer* 2016;115(9):1147-55. doi: 10.1038/bjc.2016.304
2. Magadi W, Exarchakou A, Rachet B, et al. Cancer survival in England: patients diagnosed between 2010 and 2014 and followed up to 2015. Newport, Wales, 2016.
3. Rachet B, Ellis L, Maringe C, et al. Socioeconomic inequalities in cancer survival in England after the NHS Cancer Plan. *British Journal of Cancer* 2010;103:446-53.
4. Riaz SP, Horton M, Kang J, et al. Lung cancer incidence and survival in England: an analysis by socioeconomic deprivation and urbanization. *J Thorac Oncol* 2011;6(12):2005-10. doi: 10.1097/JTO.0b013e31822b02db
5. Ellis L, Coleman MP, Rachet B. How many deaths would be avoidable if socioeconomic inequalities in cancer survival in England were eliminated? A national population-based study, 1996-2006. *European Journal of Cancer* 2012;48:270-78.
6. Macleod U, Mitchell E, Black M, et al. Comorbidity and socioeconomic deprivation: an observational study of the prevalence of comorbidity in general practice. *Eur J Gen Pract* 2004;10(1):24-6.
7. McLean G, Gunn J, Wyke S, et al. The influence of socioeconomic deprivation on multimorbidity at different ages: a cross-sectional study. *Br J Gen Pract* 2014;64(624):e440-7. doi: 10.3399/bjgp14X680545
8. Riaz SP, Luchtenborg M, Jack RH, et al. Variation in surgical resection for lung cancer in relation to survival: population-based study in England 2004-2006. *Eur J Cancer* 2012;48(1):54-60. doi: 10.1016/j.ejca.2011.07.012
9. Berglund A, Lambe M, Luchtenborg M, et al. Social differences in lung cancer management and survival in South East England: a cohort study. *BMJ Open* 2012;2(3) doi: 10.1136/bmjopen-2012-001048
10. Forrest LF, White M, Rubin G, et al. The role of patient, tumour and system factors in socioeconomic inequalities in lung cancer treatment: population-based study. *Br J Cancer* 2014;111(3):608-18. doi: 10.1038/bjc.2014.310
11. Forrest LF, Adams J, Rubin G, et al. The role of receipt and timeliness of treatment in socioeconomic inequalities in lung cancer survival: population-based, data-linkage study. *Thorax* 2015;70(2):138-45. doi: 10.1136/thoraxjnl-2014-205517
12. Nur U, Quaresma M, De Stavola B, et al. Inequalities in non-small cell lung cancer treatment and mortality. *J Epidemiol Community Health* 2015;69(10):985-92. doi: 10.1136/jech-2014-205309
13. Office for National Statistics. Cancer Registration Statistics, England: first release, 2014. Newport: Office for National Statistics, 2014.
14. Centre HaSCI. The National Clinical Lung Cancer Audit (LUCADA) Data Manual v3.1.5, 2013.
15. Shack LG. What factors influence socio-economic inequalities in colorectal cancer survival? [PhD Thesis]. London School of Hygiene and Tropical Medicine, 2009.
16. Benitez-Majano S, Fowler H, Maringe C, et al. Deriving stage at diagnosis from multiple population-based sources: colorectal and lung cancer in England. *Br J Cancer* 2016;115(3):391-400. doi: 10.1038/bjc.2016.177 [published Online First: 2016/06/22]
17. Maringe C, Fowler H, Rachet B, et al. Reproducibility, reliability and validity of population-based administrative health data for the assessment of cancer non-related comorbidities. *PLoS ONE* 2017;12(3) doi: 10.1371/journal.pone.0172814
18. Lim E, Baldwin D, Beckles M, et al. Guidelines on the radical management of patients with lung cancer. *Thorax* 2010;65 Suppl 3:iii1-27. doi: 10.1136/thx.2010.145938
19. Health and Social Care Information Centre. OPCS-4 Classification [Available from: <http://systems.hscic.gov.uk/data/clinicalcoding/codingstandards/opcs4>].
20. National Cancer Intelligence Network. Site Specific Clinical Reference Groups (SSCRGs) 2013 [Available from: http://www.ncin.org.uk/cancer_type_and_topic_specific_work/cancer_type_specific_work/sscrgs accessed 14/10/2014].
21. Department for Communities and Local Government. The English Indices of Deprivation 2010. London, 2011.

22. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007;26(11):2389-430. doi: 10.1002/sim.2712
23. Bartlett JW, Seaman SR, White IR, et al. Multiple imputation of covariates by fully conditional specification: Accommodating the substantive model. *Statistical methods in medical research* 2015;24(4):462-87. doi: 10.1177/0962280214521348 [published Online First: 2014/02/15]
24. Carpenter JR, Kenward MG. Multiple imputation and its application. Chichester, West Sussex: Wiley 2013.
25. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30(4):377-99. doi: 10.1002/sim.4067 [published Online First: 2011/01/13]
26. Molenberghs G, Verbeke G. Models for discrete longitudinal data. New York: Springer Science+Business Media, Inc 2005.
27. Krieger N. Socioeconomic data in cancer registries. *American journal of public health* 2001;91(1):156-7. [published Online First: 2001/02/24]
28. Johnson AM, Hines RB, Johnson JA, 3rd, et al. Treatment and survival disparities in lung cancer: the effect of social environment and place of residence. *Lung Cancer* 2014;83(3):401-7. doi: 10.1016/j.lungcan.2014.01.008
29. Jiang X, Lin G, Islam KM. Socioeconomic factors related to surgical treatment for localized, non-small cell lung cancer. *Social science & medicine* 2017;175:52-57. doi: 10.1016/j.socscimed.2016.12.042
30. Rich AL, Tata LJ, Free CM, 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. doi: 10.1136/thx.2011.158972
31. Haneuse S, Lee KH. Semi-Competing Risks Data Analysis: Accounting for Death as a Competing Risk When the Outcome of Interest Is Nonterminal. *Circ Cardiovasc Qual Outcomes* 2016;9(3):322-31. doi: 10.1161/CIRCOUTCOMES.115.001841
32. Fischer B, Lassen U, Mortensen J, et al. Preoperative Staging of Lung Cancer with Combined PET-CT. *New England Journal of Medicine* 2009;361(1):32-39. doi: 10.1056/NEJMoa0900043
33. Berglund A, Holmberg L, Tishelman C, et al. Social inequalities in non-small cell lung cancer management and survival: a population-based study in central Sweden. *Thorax* 2010;65(4):327-33. doi: 10.1136/thx.2009.125914
34. Peake MD. Deprivation, distance and death in lung cancer. *Thorax* 2015;70(2):108-9. doi: 10.1136/thoraxjnl-2014-206153
35. Crawford SM, Sauerzapf V, Haynes R, et al. Social and geographical factors affecting access to treatment of lung cancer. *Br J Cancer* 2009;101(6):897-901. doi: 10.1038/sj.bjc.6605257 [published Online First: 2009/08/20]
36. Raine R, Wong W, Scholes S, et al. Social variations in access to hospital care for patients with colorectal, breast, and lung cancer between 1999 and 2006: retrospective analysis of hospital episode statistics. *BMJ* 2010;340:b5479. doi: 10.1136/bmj.b5479 [published Online First: 2010/01/16]
37. Vach W, Blettner M. Biased estimation of the odds ratio in case-control studies due to the use of ad hoc methods of correcting for missing values for confounding variables. *American journal of epidemiology* 1991;134(8):895-907.
38. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393. doi: 10.1136/bmj.b2393
39. Peake MD, Thompson S, Lowe D, et al. Ageism in the management of lung cancer. *Age Ageing* 2003;32(2):171-7. [published Online First: 2003/03/05]
40. Janssen-Heijnen MLG, Smulders S, Lemmens VEPP, et al. Effect of comorbidity on the treatment and prognosis of elderly patients with non-small cell lung cancer. *Thorax* 2004;59(7):602-07. doi: 10.1136/thx.2003.018044
41. Chambers A, Routledge T, Pilling J, et al. In elderly patients with lung cancer is resection justified in terms of morbidity, mortality and residual quality of life? *Interact Cardiovasc Thorac Surg* 2010;10(6):1015-21. doi: 10.1510/icvts.2010.233189 [published Online First: 2010/04/01]
42. Tracey E, McCaughan B, Badgery-Parker T, et al. Survival of Australian lung cancer patients and the impact of distance from and attendance at a thoracic specialist centre: a data linkage study. *Thorax* 2015;70(2):152-60. doi: 10.1136/thoraxjnl-2014-205554

43. Thomas AA, Pearce A, O'Neill C, et al. Urban–rural differences in cancer-directed surgery and survival of patients with non-small cell lung cancer. *Journal of Epidemiology and Community Health* 2017;71(5):468-74. doi: 10.1136/jech-2016-208113

Appendix 1:

A population-based study to quantify the association between patient's age, deprivation and specific comorbid conditions, and the receipt of surgery in non-small cell lung cancer patients in England

Aurélien Belot¹, Helen Fowler¹, Edmund Njeru Njagi¹, Miguel-Angel Luque-Fernandez¹, Camille Maringe¹, Winnie Magadi¹, Aimilia Exarchakou¹, Manuela Quaresma¹, Adrian Turculet¹, Michael D Peake^{2,3,4}, Neal Navani⁵, Bernard Rchet¹

¹ Cancer Survival Group, Department of Non-Communicable Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

² Department of Respiratory Medicine, University Hospitals of Leicester

³ National Cancer Registration and Analysis Service, Public Health England

⁴ Centre for Cancer Outcomes, University College London Hospitals

⁵ UCL Respiratory, University College London, 5 University Street, London WC1E 6JF and Department of Thoracic Medicine, University College London Hospital, 250 Euston Road, London NW1 2PG

Table S1: OPCS-4 codes used to identify a major surgical procedure (National Cancer Intelligence Network)

Code	Description	Number identified in 2012 in England in NSCLC patients
E391	Open excision of lesion of trachea	5
E398	Other specified partial excision of trachea	0
E399	Unspecified partial excision of trachea	0
E441	Excision of carina	2
E461	Sleeve resection of bronchus and anastomosis HFQ	9
E541	Total pneumonectomy, total removal of lung, Pneumonectomy NEC	292
E542	Bilobectomy of lung	141
E543	Lobectomy of lung	3384
E544	Excision of segment of lung	816
E545	Partial lobectomy of lung NEC	75
E548	Excision of lung, other specified	75
E549	Excision of lung, Unspecified	1
E552	Open excision of lesion of lung	40
E559	Open removal of lesion of lung, unspecified	0
T013	Excision of lesion of chest wall	10
T023	Insertion of prosthesis into chest wall NEC	0

Figure S1: Predicted probability of receiving major surgery with 95% confidence interval for early stage (solid line) and advanced stage (dashed line) in female NSCLC patients without any comorbidities, with good performance status, having had a PET-CT scan: (a) according to age at diagnosis with a deprivation score of 0.1 or (b) according to deprivation score with an age at diagnosis of 70 years. Results are based on the multivariable logistic regression. The ranges of age at diagnosis and deprivation score were limited to the 2.5th and 97.5th percentile of the observed distribution.

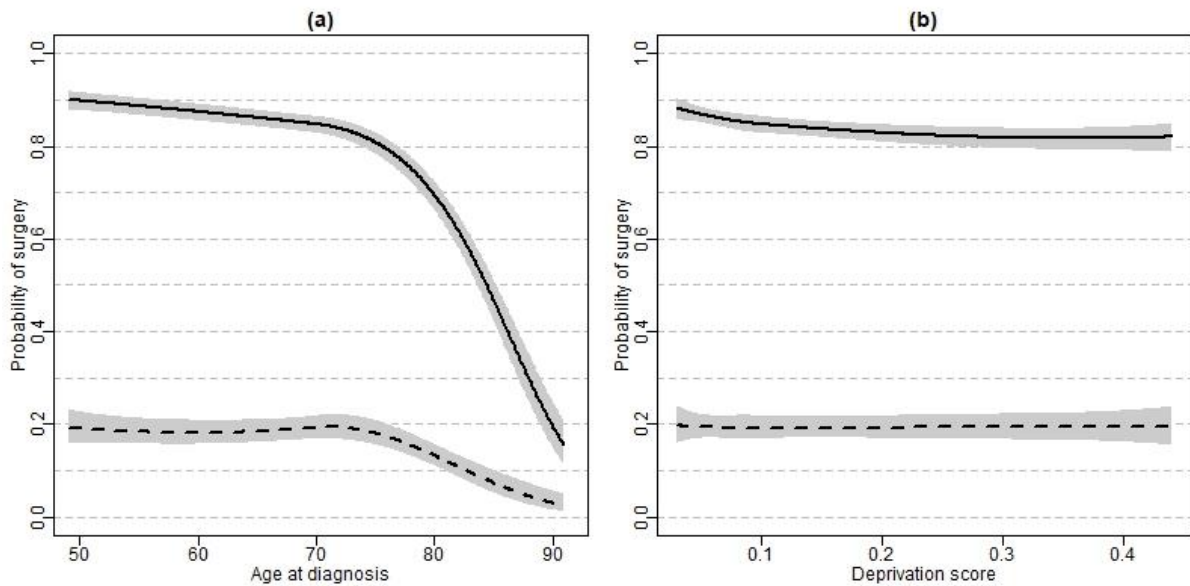
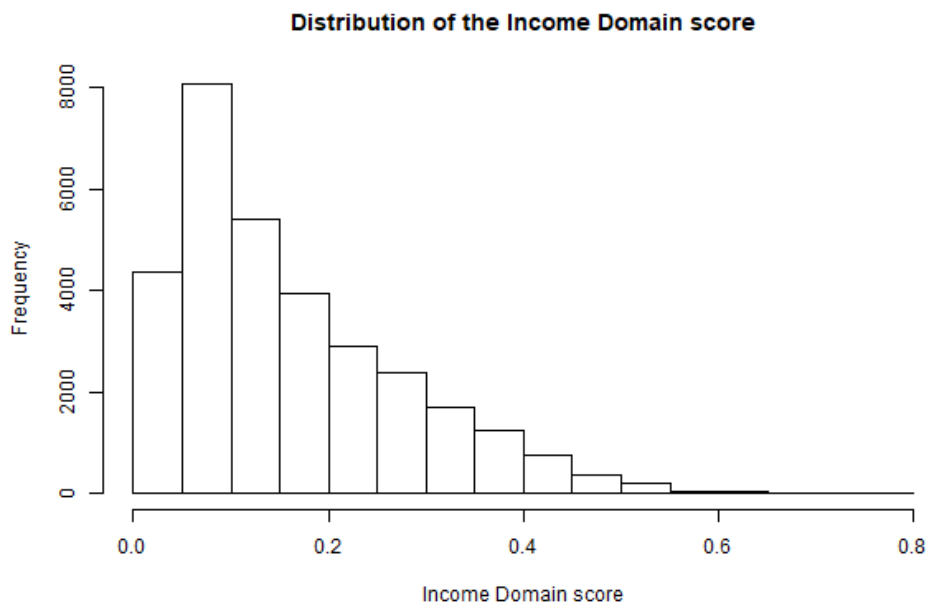
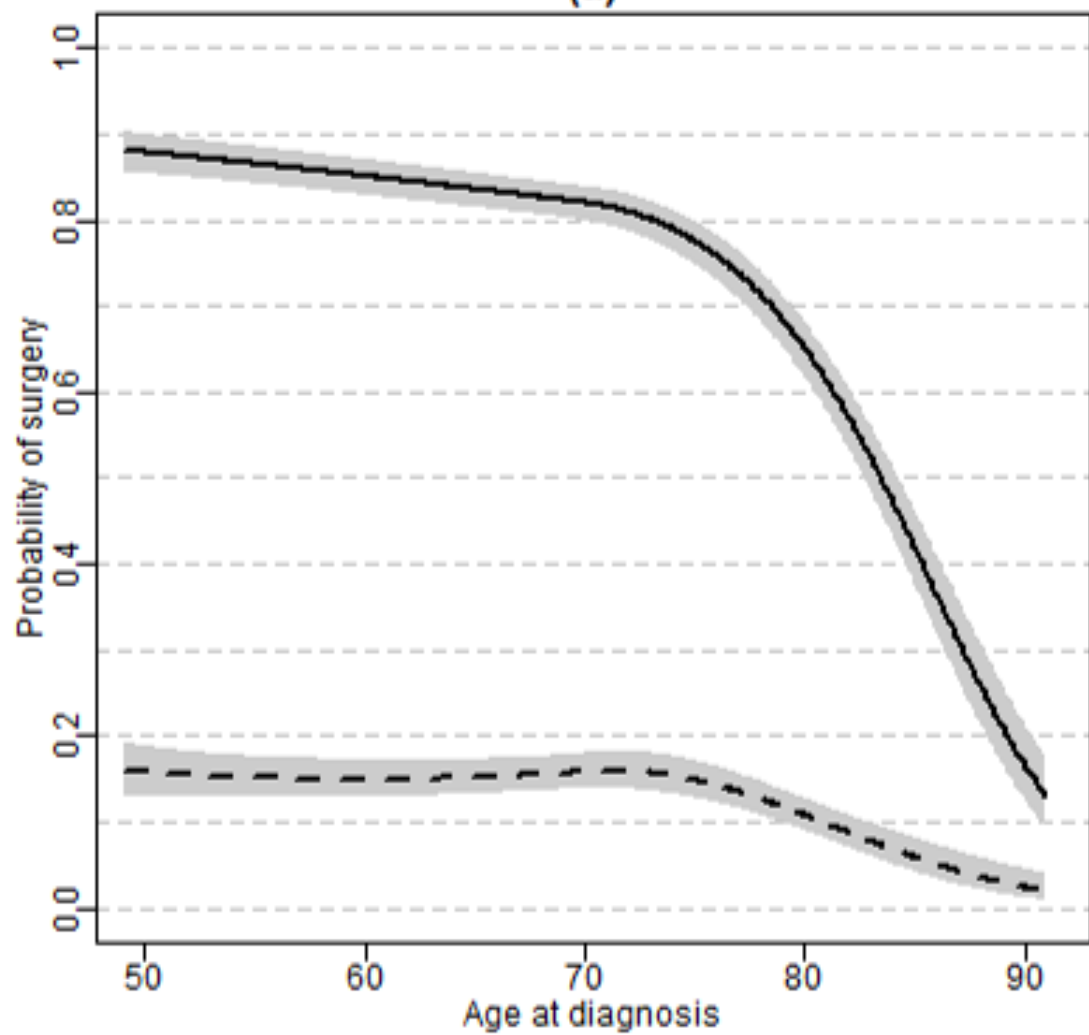


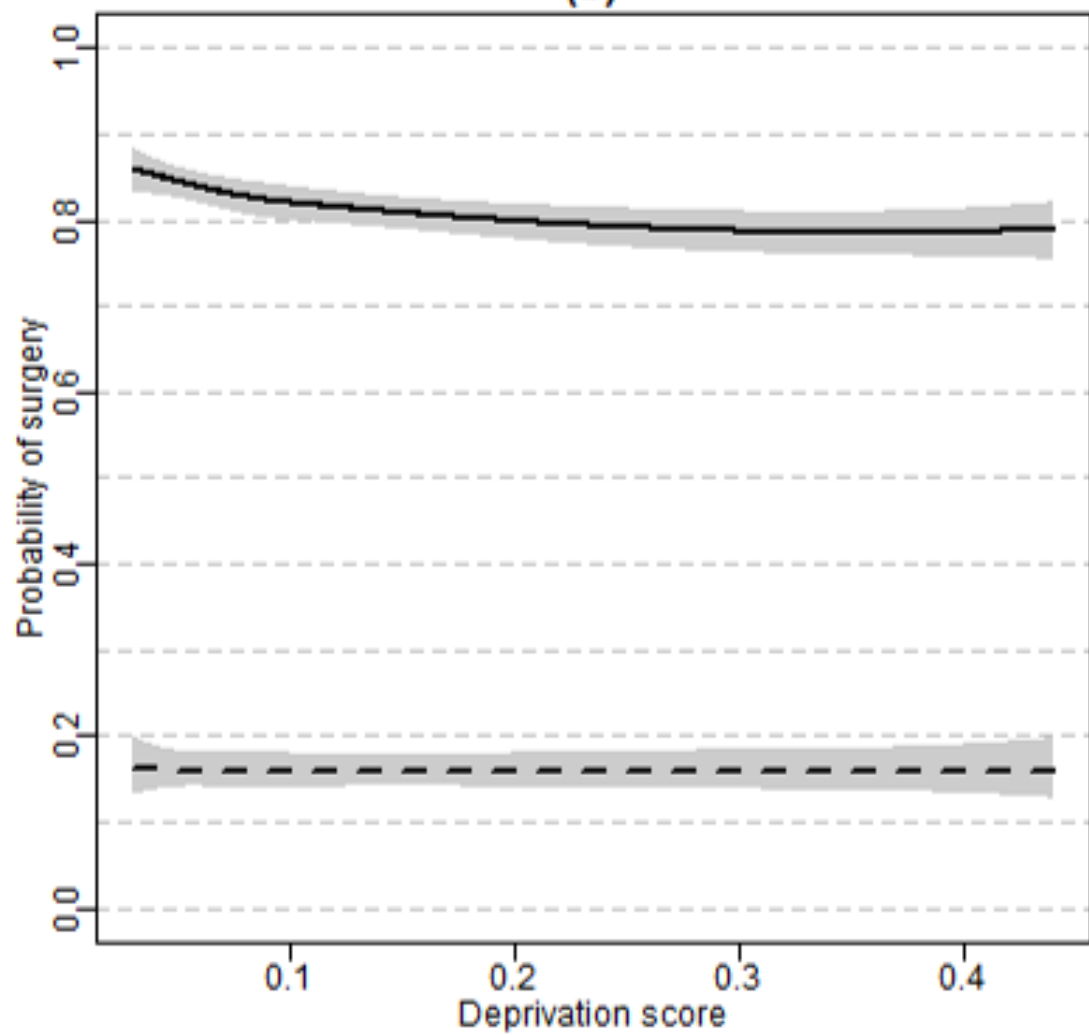
Figure S2: Distribution of the Income Domain score among NSCLC patients diagnosed in 2012 in England (mean=0.17, 1st quartile=0.07, 3rd quartile=0.23)

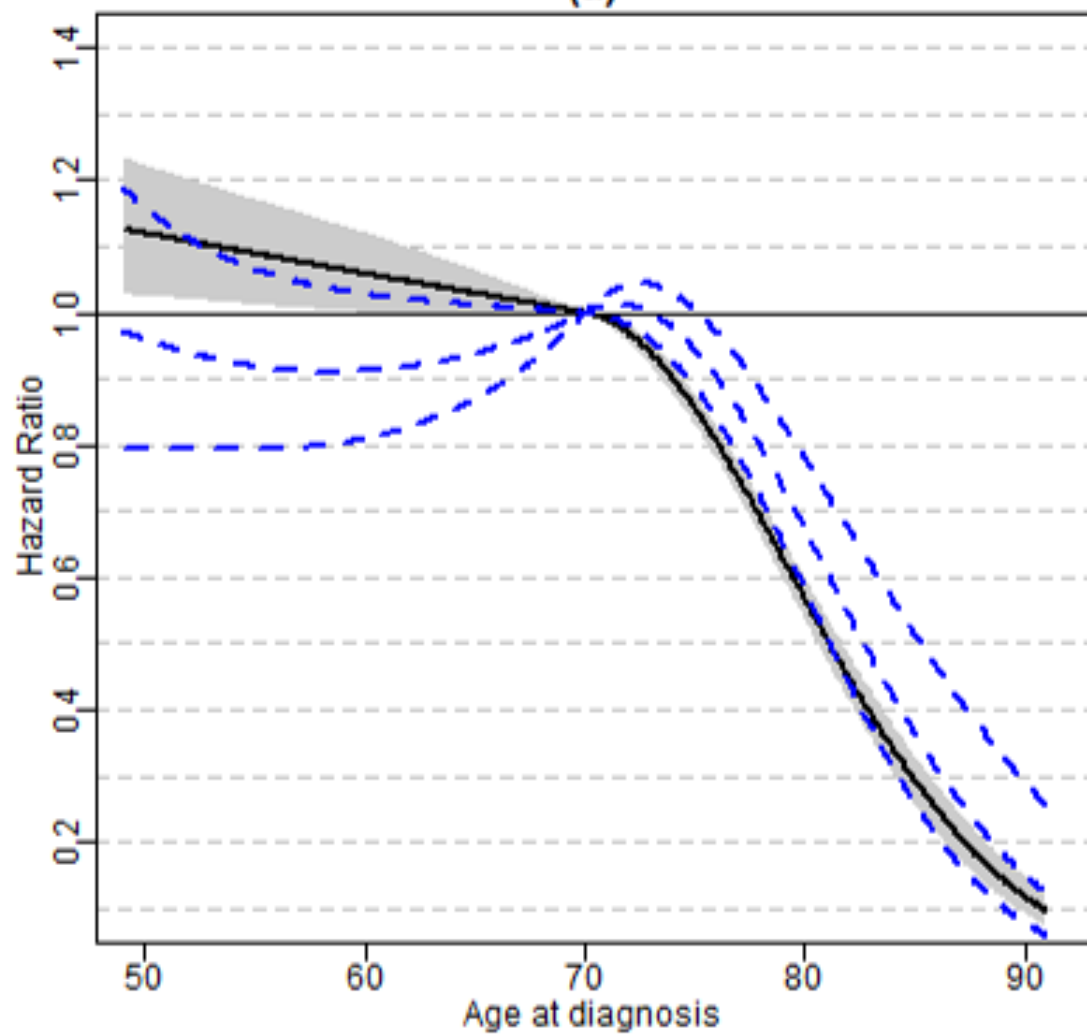
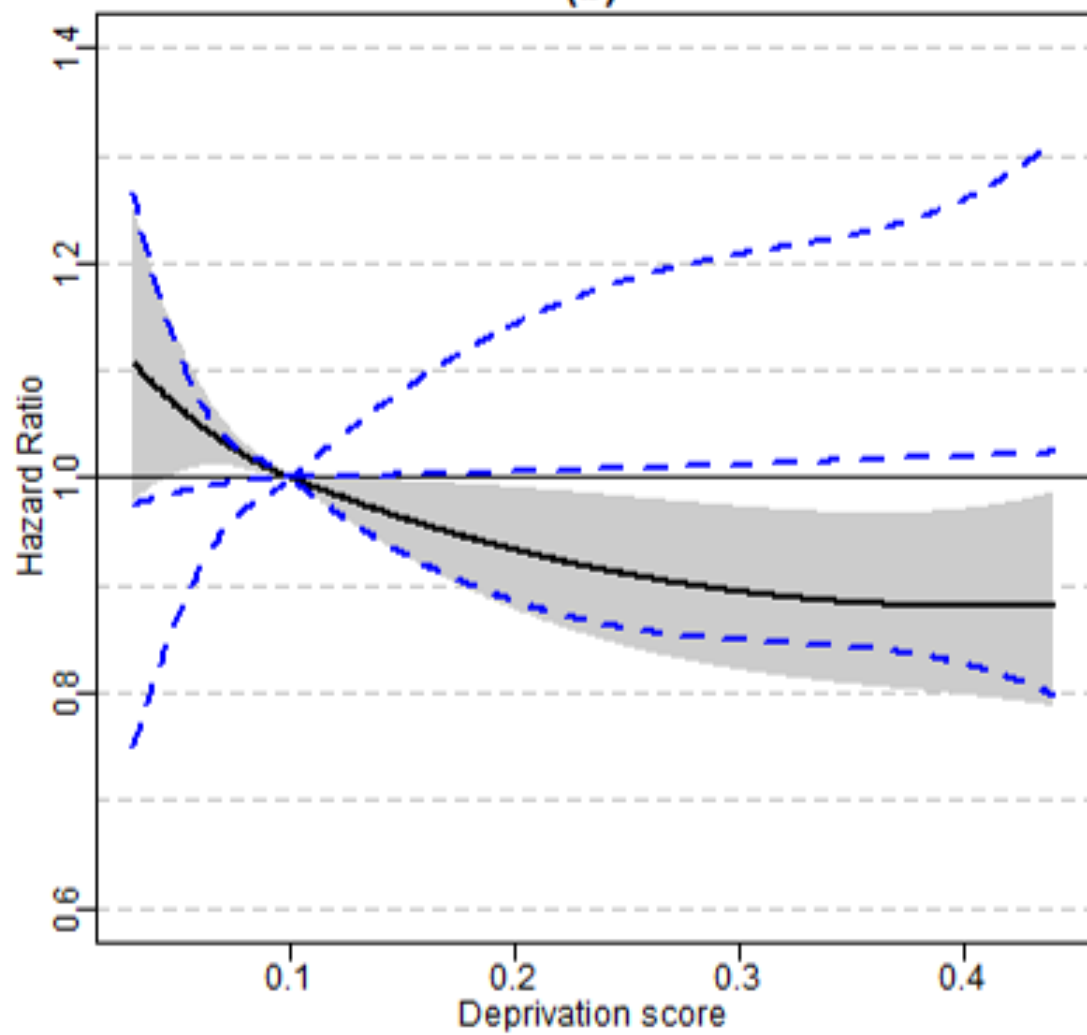


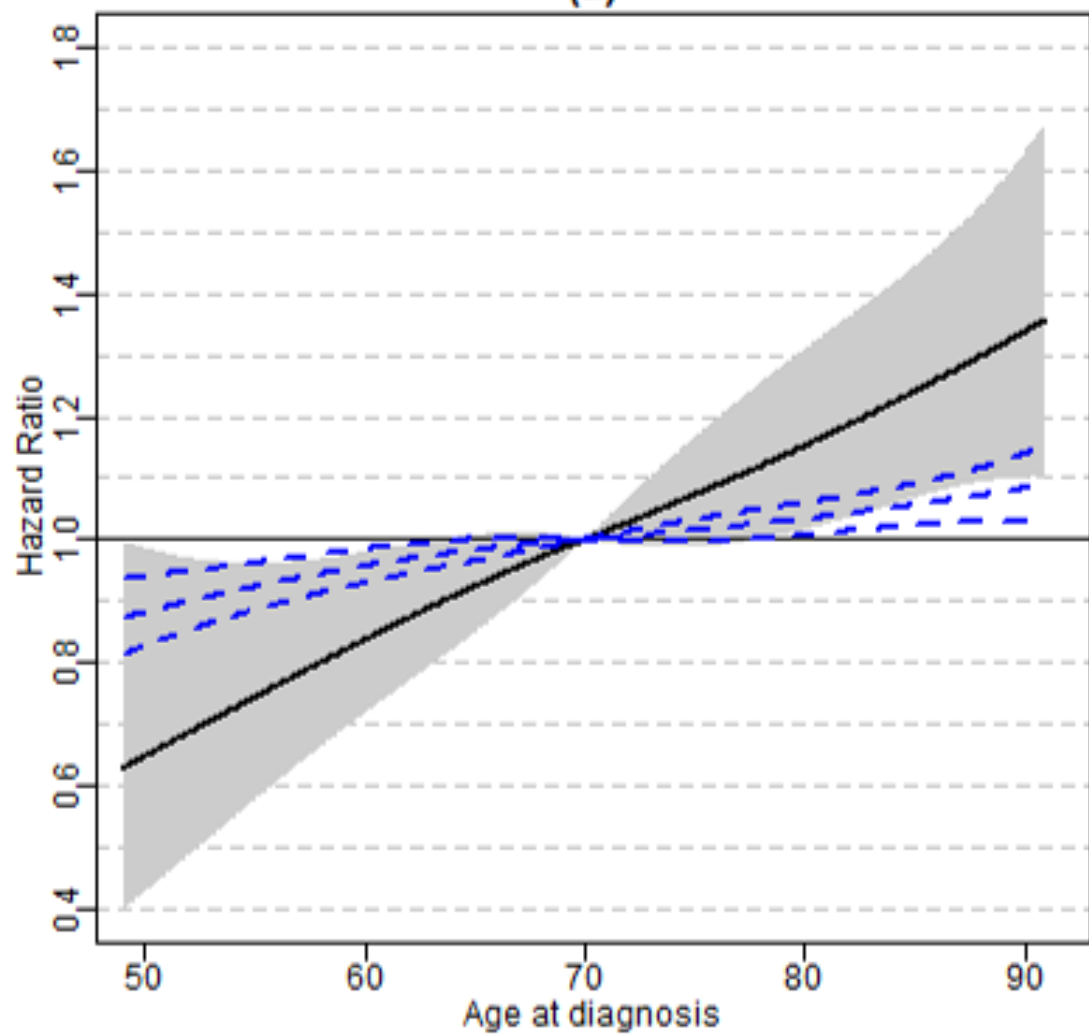
(a)



(b)



(a)**(b)**

(a)**(b)**